CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 3 Evaluation & Monitoring Annual Report

Prepared for:

CMS

Prepared by:

The Lewin Group

October 2017 – Revised October 2018

CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 3 Evaluation & Monitoring Annual Report

October 2017 – Revised October 2018

The Lewin Group

Authors:

Laura Dummit, Grecia Marrufo, Jaclyn Marshall, Brandon Maughan, Tristan Ackerman, Sarah Bergman, Aylin Bradley, Rebecca Cherry, Inna Cintina, Daniel Cooper, Syvart Dennen, Philip Dishuck, Ayah Fannoun, Megan Hyland, Gina Gerding, Dan Gregory, Catherine Griffin, Ashley Johnson, Susan Joy, Jordan Kahn, Jon Kelly, Alex Lampert, Karla Lopez de Nava, Court Melin, Sebastian Negrusa, Julie Somers, Dan Strubler, Katarina Swanson, Ellen Tan, Kyi-Sin Than, Peter Weidner, Ayanna Williams, Grace Yang, Mark Zezza, David Zhang, Dean Farley, Ian Breunig, Andrea Hassol, Qian Gu, Sean McClellan, Matt Trombley, Christine LaRocca, Lindsay Kirsch, Colleen Kummet, Michelle Roozeboom

Lewin's address:

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

Federal Project Officer:

Daver Kahvecioglu
Division of Data, Research, and Analytic Methods (DRAM)
Research and Rapid Cycle Evaluation Group (RREG),
Center for Medicare and Medicaid Innovation (CMMI),
Centers for Medicare and Medicaid Services (CMS)

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2011-00001I Task Order HHSM-500-T0007.

Table of Contents

List of Appendices

Executive Summary

The Centers for Medicare & Medicaid Services (CMS) launched the Bundled Payments for Care Improvement (BPCI) initiative under the authority of the Center for Medicare and Medicaid Innovation (CMMI). The BPCI initiative tests four Models for linking provider payments for a clinical episode of care to determine whether bundled payments can reduce Medicare payments while maintaining or improving quality of care. The voluntary initiative is designed to allow participants to choose among several key options such as payment approach, type of clinical episode, and episode definitions. This design implicitly recognizes the variability across health care markets, providers, and episodes of care. The resulting diversity in responses and impacts will provide CMMI with information on the approaches that show the most promise in achieving payment reductions while maintaining or improving quality.

This annual report reflects the early stages of our evaluation of the impact of three of the four BPCI Models based on the first two years of the BPCI initiative.[1] Although episode-based payments began as early as October 2013 for Model 2, 3, and 4 participants, organizations were able to join and add clinical episodes over an extended period. For this reason, this report includes an average of three quarters of data across participating providers. The resulting short intervention periods for many participants may have contributed to the lack of significant changes across many of the key outcomes. Nevertheless, this third annual report provides insights into responses to bundled payment incentives.

A. Structure of the Initiative

The BPCI initiative rewards participants financially for reducing Medicare payments for a clinical episode of care relative to a target price. BPCI Awardees, which can be health care providers or other entities that convene health care organizations, entered into agreements with CMS to be held accountable for total Medicare episode payments. Awardees' agreements with CMS specified their Model choice as well as choices among 48 clinical episodes, other episode characteristics, and multiple options for program rule waivers and financial arrangements with other parties. The clinical episodes are defined by the Medicare Severity Diagnosis Related Group (MS-DRG) assigned to the anchor or qualifying hospitalization. Providers and other organizations voluntarily participate in BPCI. They could enter into the risk-bearing phase of the initiative during a 2-year period, from October 2013 through October 2015, and enter additional clinical episodes into the risk-bearing phase throughout that time.

The three BPCI Models evaluated in this report are:

- **Model 2** – This Model has the most comprehensive bundle, which includes the anchor inpatient hospital stay and all concurrent professional services and other Medicare Part A- and Part B-covered services (with certain exclusions) furnished within the chosen episode length of 30, 60, or 90 days post-discharge. Individual providers continue to be paid Medicare fee-for-service amounts and aggregated episode payments are reconciled retrospectively against a target price, which CMS determined based on discounted, historical Medicare payments. When Awardees' episode payments are less than the target price, Awardees may receive the difference, termed the net payment reconciliation

[1] Model 1 began earlier than Models 2, 3, and 4 and was evaluated separately.

amount (NPRA), which they can keep or share with their partnering providers. When Awardees' episode payments are greater than the target price, they may have to pay amounts to CMS. The episode initiator (EI), that is, the provider associated with the start of the episode, can be a hospital or a physician group practice (PGP).

- **Model 3** – 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) within 30 days following a hospital discharge from an MS-DRG in the participant's chosen clinical episode, or when a beneficiary is admitted to a SNF, HHA, IRF, or LTCH within 30 days following a hospitalization for an MS-DRG in which the attending or operating physician is a member of a PGP participating in a clinical episode that includes that MS-DRG. The bundle includes all Part A- and Part B-covered services (with certain exclusions) within the chosen episode length of 30, 60, or 90 days. Individual providers continue to be paid Medicare fee-for-service amounts and aggregated episode payments are reconciled retrospectively against the target price. When Awardees' episode payments are less than the target price, Awardees may receive NPRA, which they can keep or share with their partnering providers. When Awardees' episode payments are greater the target price, they may have to pay amounts to CMS. The EI can be a SNF, HHA, IRF, LTCH, or PGP.
- **Model 4** – The episode starts when the beneficiary is admitted to the participating hospital for the chosen clinical episode. The bundle includes the anchor hospitalization, all professional services furnished during the anchor hospitalization, and any readmissions and associated professional services that occur within 30 days of hospital 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. The EI is a hospital.

B. Evaluation Design

The BPCI evaluation is based on a mixed methods approach designed to incorporate multiple data sources to generate valid information on the three main research questions:

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?

The evaluation relies on Medicare claims and enrollment data and post-acute care (PAC) provider patient assessments for episodes initiated through September 2015, patient surveys through December 2015, participant interviews through June 2016, and participant site visits through August 2016. The non-experimental quantitative design uses 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 rely on data available only for BPCI participants, we use pre- and post-comparisons to estimate changes

attributed to BPCI. The qualitative data provide more detailed information on BPCI participant approaches and experiences as well as context for the quantitative results.

Because of the unique design of the BPCI initiative, with multiple start dates for participants and various combinations of Models, clinical episodes, and EIs, the results across the three main evaluation questions are first differentiated by Model and EI provider type. They are further stratified by clinical episode because of differences in the underlying cost and utilization patterns across them. The sample sizes are not sufficient to examine every Model, EI, and clinical episode combination. Because of data limitations, this report does not include analyses of episodes initiated by PGP EIs, which are expected in future reports. To discern trends or emerging patterns that may suggest opportunities for achieving CMS's objectives or raise potential concerns, we also examined outcomes across clinical episodes with similar characteristics under the same Model, and outcomes across certain clinical episodes regardless of Model or EI.

C. Results

Key findings are summarized under the three main evaluation questions.

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

- **Model 3 had the most participants but the vast majority of episodes were initiated under Model 2**

Between October 2013 and September 2015, Model 2 was selected by 215 Awardees that represented 422 acute care hospitals and 277 PGP EIs. These participants initiated more than 242,000 episodes, or over 85% of all episodes across the three Models, during the first two years of the initiative. All 48 clinical episodes were chosen among the Model 2 hospital and PGP EIs, although few EIs participated in all of the clinical episodes. The major joint replacement of the lower extremity (MJRLE) clinical episode was selected by 60% of EIs, making it the most popular, followed by congestive heart failure (CHF) with 27 % of EIs; 28 clinical episodes were chosen by fewer than 10% of EIs. During the first two years of the BPCI initiative, 27 Model 2 hospitals (6%) withdrew from BPCI, as did 1 PGP.

Under Model 3, between October 2013 and September 2015, there were 136 Awardees and 873 SNF, 144 PGP, and 116 HHA EIs, with little participation by IRFs and LTCHs. There were 35,000 episodes initiated under Model 3. About one-third of EIs participated in MJRLE, one-quarter participated in simple pneumonia and respiratory infection (SPRI), CHF, or chronic obstructive pulmonary disease (COPD) clinical episodes, and 10% or less participated in 13 of the remaining 48 clinical episodes. During the first two years of the initiative, 44 SNF EIs (5%) and 15 HHA EIs (13%) withdrew from BPCI.

- **Model 4 was chosen by few and experienced substantial withdrawals**

Participation in Model 4 during the first two years of the initiative was characterized by low and falling interest among providers. From a high of 23 episode-initiating hospitals, nine were participating by the end of the second performance year. A total of 7,682 episodes, predominantly for MJRLE, were initiated under Model 4 through September 2015. Model 4 participants indicated that several BPCI design features made it difficult for them to succeed and contributed to their decision to withdraw. They cited difficulties in developing and maintaining

the infrastructure required to pay participating providers. Some also mentioned that they were disadvantaged because episodes with short inpatient stays were excluded from the Model, limiting their ability to reduce average episode payments.

- **Providers that chose to participate in BPCI are larger and appear to have more resources than providers that did not participate**

The typical BPCI-participating hospital EI under Model 2 was a large, non-profit, urban facility with a teaching program. BPCI-participating SNF EIs and HHA EIs under Model 3 were larger than non-participants and more likely to be for-profit and part of a chain. Participants under Model 2 and 3 tended to be located in competitive health care markets, meaning that there was not a dominant provider. Their markets contained fewer SNF beds and more physicians, particularly specialists. Markets with BPCI participants also tended to have larger, more affluent populations. The resources available to these providers owing to their size, markets, or other characteristics may have contributed to their ability or willingness to engage in the initiative.

- **Participating providers had higher 2011 standardized Part A payments than non-participants**

Hospital EIs under Model 2 chose to participate in clinical episodes for which their average 2011 standardized Part A payments were higher than average payments for non-participating hospitals. In 2011, BPCI-participating hospital EIs had standardized Part A payments that averaged $1,366 (7%) higher than payments for non-participating hospitals (among clinical episodes with sufficient sample size for evaluation). The same was true for Model 3 SNFs and HHAs (among clinical episodes with sufficient sample size for evaluation). BPCI-participating SNFs had 2011 standardized Part A payments that were $742 (3%) higher than payments for the same episodes for non-participating SNFs. HHAs had 2011 standardized Part A payments that were $263 (3%) higher than non-participants. These data are consistent with what participants told us in interviews, that they chose clinical episodes that gave them the most opportunities to reduce payments.

- **Use of beneficiary incentives, program rule waivers, and gainsharing was limited**

The initiative includes certain Medicare program flexibilities through the use of beneficiary incentives, program rule waivers, and gainsharing. Although the majority of Awardees indicated in their agreements with CMS that they wanted these options for their EIs, few EIs used them. Under Model 2, 49% of EIs could provide beneficiary incentives, but only 18% actually did so through the second year of the initiative. Under Model 3, 71% of EIs could provide beneficiary incentives, but only 5% actually did so. Although 63% of Model 2 EIs could use a program rule waiver that allows Medicare coverage of a SNF stay following a hospitalization of less than three days, it was used for only 1,846 episodes (approximately 5% of all episodes discharged to a SNF). Similarly, there was limited use of the telehealth or the home visit program rule waivers.

Awardees that described their gainsharing plans in their CMS agreement are allowed to share NPRA or internal cost savings (ICS) with partnering providers. Gainsharing is intended to help participants engage their partnering providers. Gainsharing plans were included in the agreements of 61% of Model 2 Awardees. Gainsharing was used by 18% of these Awardees to distribute approximately $13.5 million in NPRA and by 8% of Awardees to distribute $3.1 million in ICS. Physicians were the most likely recipients of gainsharing payments. Gainsharing plans were

included in the agreements of 15% of the Model 3 Awardees and 43% of these Awardees distributed approximately $4.6 million in NPRA to provider partners, mostly physicians.

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

We had sufficient sample to examine 23 clinical episodes under Model 2, 14 episodes under Model 3, and 2 episodes under Model 4.

- **Medicare payments for the anchor or qualifying[2] hospitalization plus 90 days post discharge for MJRLE episodes declined under Model 2 hospital EIs by 4.5% and under Model 3 SNF EIs by 7.1% relative to the comparison group**

Medicare payments for the MJRLE anchor hospitalization plus bundled services furnished during the 90 days post discharge declined by $1,273 ($p<0.01$) from the baseline to the intervention period for BPCI episodes relative to the comparison group under Model 2 hospital EIs. This is 4.5% less than what payments would have been without BPCI. The relative decline in Model 2 episode payments was due to lower payments for PAC. SNF payments declined by $711 ($p< 0.01$) and IRF payments declined by $435 ($p<0.01$) relative to the comparison group. Model 2 hospital EIs reduced the share of their MJRLE patients who were discharged to institutional PAC from 62.5% to 52.3% while the comparison group reduced the share discharged to institutional PAC from 61.2% to 56.9% (-5.9 percentage points, $p<0.01$). For those patients who had any SNF use during the episode, the average number of SNF days decreased 1.5 days ($p<0.01$) relative to the comparison group. Participants we interviewed indicated that they reduced PAC use by changing patient and physician expectations about the need for PAC through education and consistent messaging by hospital staff.

Under Model 3, payments for the qualifying hospitalization plus bundled services furnished during the 90 days post-discharge declined for SNF EIs with MJRLE episodes by $2,568 ($p<0.01$) relative to comparison providers, resulting in episode payments that were 7.1% lower than what they would have been absent BPCI. The lower payments were due primarily to reduced SNF days. SNF payments declined $2,255 ($p<0.01$) for BPCI episodes than comparison episodes.

Model 3 HHA EIs achieved a $970 ($p=0.09$) relative reduction in CHF payments for the qualifying hospitalization plus 90 days post-hospital discharge, resulting in payments that were 3.6% lower than what they would have been absent BPCI.

There were no other clinical episodes under Model 2, 3 or 4 among those we evaluated that had statistically significant differences in payments relative to the comparison group for the inpatient stay plus bundled services furnished during the 90-day post-hospital discharge period under BPCI.

- **Under Model 2, participants reduced institutional PAC**

Although for the most part the proportion of patients in Model 2 episodes discharged to PAC did not change, for many clinical episodes SNF and IRF payments declined and HHA payments increased. Among patients who received any PAC, the percentage discharged to institutional PAC declined in about two-thirds (61%) of clinical episodes. The decline was statistically significant for

[2] The inpatient hospital stay is defined as the anchor hospitalization in Model 2 and the qualifying hospitalization in Model 3.

MJRLE ($p<0.01$), cardiac valve ($p=0.10$), and other respiratory ($p=0.08$) clinical episodes relative to the comparison group. HHA payments increased in 19 of the 23 Model 2 clinical episodes we examined. The increase was statistically significant relative to the comparison group for 6 of the clinical episodes. All 10 of the Model 2 clinical episodes with the highest proportion of total episode payments attributable to PAC during the baseline period had relative increases in HHA payments and 9 had relative declines in SNF payments. These patterns were associated with relative reductions in total per episode payments, although the decline in per episode payments was statistically significant only for MJRLE clinical episodes. These patterns suggest that BPCI participants shifted PAC use from institutional settings to HHAs.

Consistent with the PAC provider payment trends, there was an overall decline in the use of SNF care and an increase in HHA care. The number of SNF days declined for BPCI patients who used SNF care relative to the SNF days for patients in the comparison group in 16 of the 23 clinical episodes, although the difference was statistically significant only for MJRLE and SPRI clinical episodes. The number of home health visits increased from the baseline to the intervention period among BPCI patients who had at least one visit relative to the change for the comparison group in 20 of the 23 clinical episodes. This relative increase was statistically significant for CHF (0.8 visits, $p=0.05$) and hip and femur procedures except major joint (1.9 visits, $p<0.01$) clinical episodes. The shift from SNF to HHA care was more notable in the 10 clinical episodes with the highest proportion of baseline episode payments due to PAC.

- **There were few statistically significant differences in quality of care under Model 2**

Across all Model 2 clinical episodes with sufficient sample size, there were few statistically significant relative changes in quality of care during the first two years of the BPCI initiative.

Medicare claims data were used to compare changes in mortality, emergency department use, and readmission rates between BPCI and comparison episodes. Only six of the 63 claim-based quality outcomes measured for the 23 clinical episodes were statistically significant at the 10% level; four of which indicated a relative improvement in quality for BPCI patients relative to comparison group patients. BPCI-participating hospitals experienced a statistically significant relative decline in mortality for renal failure (2.4 percentage points, $p=0.03$), nutritional and metabolic disorders (3.6 percentage points, $p=0.07$), and acute myocardial infarction (2.5 percentage points, $p=0.05$) clinical episodes. Emergency department use declined in stroke clinical episodes (2.3 percentage points, $p=0.07$). At the same time, more than half of the clinical episodes had increases in emergency department use and readmission rates, although the magnitude of the increases tended to be small and the differences between BPCI and comparison episodes were not statistically significant. There was a 2.7 percentage point ($p=0.06$) relative increase in unplanned readmissions for hip and femur procedures except major joint clinical episodes. Spinal fusion (non-cervical) clinical episodes had a 4.5 percentage point ($p<0.01$) increase in emergency department use relative to the change in the comparison group.

Beneficiaries with one of eight clinical episodes were surveyed on their satisfaction, improvement in functional status, health status, and care experience. BPCI and comparison respondents generally reported similar changes across the three domains from before their anchor hospitalization to after their episode ended. Moreover, few of the differences were consistent from one survey wave to the next. The two clinical episodes with the most consistent differences were MJRLE and sepsis. For MJRLE episodes, which exhibited declines in episode payments

and PAC use, BPCI respondents reported greater improvement across all three domains than comparison beneficiaries. BPCI respondents with sepsis episodes, which had an increase in anchor hospital length of stay and HHA payments, reported declines across the three domains, relative to the comparison group.

- **Unplanned readmissions within 90 days of SNF admission did not decline as much in BPCI participating-SNFs as in comparison SNFs under Model 3**

There were few clinical episodes in which the change in the 90-day unplanned readmission rate or readmission payments for BPCI SNFs under Model 3 was statistically different from those of comparison SNFs. However, for 8 of 11 clinical episodes we examined, unplanned readmission rates increased between 0.2 and 6.3 percentage points more for BPCI SNFs. In most instances, unplanned readmission rates fell from the baseline to the intervention period for BPCI SNFs under Model 3, just not as much as they fell for the comparison providers. At the same time, the standardized allowed amount for readmissions increased in 9 of 11 (82%) clinical episodes. Even though the preponderance of these relative differences were not statistically different from zero, these findings may suggest that participating SNFs have reduced their ability to lower unplanned readmissions rates further. However, there were large differences in baseline readmission rates between BPCI and comparison SNFs, which contributed to these findings and indicate the need for further analysis.

- **Assessment-based measures for PAC users in Model 2 episodes did not indicate systematic quality issues relative to the comparison group**

We examined changes in activities of daily living (ADLs) for patients who were discharged to SNF, IRF, or HHA care among the approximately three-quarters of patients who were in the PAC setting long enough for two assessments. We compared initial assessments with discharge assessments to measure the change in ADLs during patients' initial PAC stay. There were no patterns in the change in ADLs between the Model 2 BPCI and comparison patients across 211 ADL measures in the three PAC settings. For eight of the measures, there was a statistically significant relative increase in the proportion of BPCI patients with improvement. For 14 measures, there was a statistically significant relative decline in the proportion of BPCI patients with improvement.

We measured the change in three ADL measures for patients who received care in a SNF. There were two clinical episodes (nutritional and metabolic disorders and cardiac value) with relative improvements in at least one ADL measure. There were five clinical episodes (MJRLE, revision of the hip or knee, sepsis, non-cervical spinal fusion, and stroke) with at least one statistically significant relative decline. The relative decline in SNF length of stay would reduce the time to achieve ADL improvements. However, there was not a strong relationship between the change in SNF length of stay and improvement in ADLs during the SNF stay.

We measured the change in five ADL measures among patients discharged to HHA. There were four clinical episodes (cardiac arrhythmia, cellulitis, esophagitis, gastroenteritis and other digestive disorders, and SPRI) with at least one measure that had a statistically significant relative improvement. There were two clinical episodes (revision of the hip or knee and cardiac value) with at least one measure that had a statistically significant relative decline.

- **Quality of care outcomes under Model 3 show mixed impact of BPCI**

For Model 3 SNF episodes, the claim-based quality measures (90-day readmission rates, emergency department use, and mortality) indicated some areas of concern, although the assessment-based measures were generally positive. Our analysis indicated relative declines in some of the quality measures for renal failure, stroke, sepsis, and COPD clinical episodes in BPCI participants. There was a relative increase in mortality (3.8 percentage points, p=0.09) for SPRI clinical episodes at the same time that other evidence suggested that there was a relative increase in the severity of patients with SPRI episodes in BPCI-participating SNFs from the baseline to the intervention period relative to comparison episodes. We will continue to explore whether adverse quality outcomes may be due to unmeasured patient severity differences, random fluctuations in relatively rare events, or an outcome of the initiative.

The assessment-based quality outcomes, based on the approximately three-quarters of patients who were in the PAC setting long enough for two assessments, indicate quality improvements for Model 3 SNF episodes. There were ADL improvements for CHF, medical non-infectious orthopedic, other respiratory, renal failure, SPRI, sepsis, and urinary tract infection clinical episodes relative to the comparison episodes. Only one ADL decline was noted, which was for MJRLE clinical episodes. Beneficiary surveys of patients with a MJRLE episode initiated by Model 3 SNF EIs indicated similar changes in functional status and mental and physical health outcomes from before to after the episode ended between the BPCI and comparison respondents. However, BPCI respondents reported worse care experiences on three of ten measures of their health care experiences.

For Model 3 HHA episodes, BPCI did not appear to have a systematic effect, either positive or negative, on the quality of care. There were no statistically significant declines in the five ADL measures for BPCI episodes relative to comparison episodes. There were, however, statistically significant relative improvements in one measure for MJRLE episodes and one for CHF episodes.

- **There were few instances in which patient complexity appeared to change for BPCI participants from baseline to intervention period**

We examined patient characteristics for indications of participants shifting their mix of patients within clinical episodes. While the impact analysis on payment, utilization, and quality outcomes controls for differences in patient characteristics, it does not directly examine any changes in patient mix. For MJRLE non-fracture patients under Model 2, there was a statistically significant decline in prior health services use, which raises concern that BPCI participants may be selecting healthier patients in this planned clinical episode. For other Model 2 clinical episodes there were no clear patterns in changes in patient severity.

Under Model 3, there may be more opportunities across clinical episodes for participants to select a healthier mix of patients because PAC providers choose who they will treat. BPCI-participating SNFs had a less severe non-fracture and fracture MJRLE patient mix from the baseline to the intervention period relative to the change for the comparison group, based on differences in demographic characteristics, prior health care use, and functional status. CHF patients in BPCI-participating HHAs also appeared to be less severe during the intervention, relative to the change in patient mix in the comparison group. In contrast, there was a significant

increase in the severity of the patients in SPRI episodes in Model 3 SNFs relative to the change for the comparison group.

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

Our analysis of variations in the impact of BPCI focused on clinical episodes with sufficient volume to analyze on a provider level, which were Model 2 COPD, MJRLE, and CHF clinical episodes. We compared participants that achieved relatively high NPRA and those that had low, or even negative, NPRA.

- **The difference between the target price and aggregate episode payments, or NPRA, varied widely across Model 2 hospital EIs in COPD, MJRLE, and CHF clinical episodes**

There was substantial variation in the NPRA earned by Model 2 hospital EIs for COPD, MJRLE, and CHF clinical episodes. NPRA per episode for MJRLE clinical episodes ranged from -$6,053 to $7,117. For COPD clinical episodes, hospital EIs achieved per-episode NPRA that ranged from -$4,580 to $2,377. For CHF episodes, per-episode NPRA ranged from -$3,917 to $5,172.

- **Reduced institutional PAC and unplanned readmissions under Model 2 was associated with greater per-episode NPRA**

Hospitals that had the highest average NPRA per episode relative to their target price decreased their use of institutional PAC in COPD, MJRLE, and CHF clinical episodes. They also increased the share of their patients discharged home without HHA. Reductions in unplanned readmissions were also associated with higher NPRA.

- **Hospital EIs with greater use of PAC and healthier patients in the baseline achieved higher per-episode NPRA under Model 2**

Hospital EIs with the highest NPRA per episode relative to their target price discharged a higher share of MJRLE and CHF patients to institutional PAC during the baseline period than hospital EIs with lower NPRA for these episodes. Hospital EIs with higher per-episode NPRA for CHF clinical episodes had healthier CHF patients during the baseline than hospital EIs with lower per-episode NPRA for CHF clinical episodes. Because the target prices were usually based on the participants' historical episode payment amounts, hospital episodes with greater use of PAC would generally have higher target prices, which could make it easier for these participants to achieve greater reductions in episode payment amounts during the intervention. When they cared for healthier patients, it may have been easier to reduce PAC without detrimental effects on quality.

- **Under Model 2, hospital EIs with higher per-episode NPRA for MJRLE clinical episodes had greater increases in non-fracture MJRLE discharges, which were likely matched by volume increases for non-BPCI hospitals in their markets**

Hospital EIs with the highest NPRA per MJRLE episode relative to their target price had an average increase of 25.4% in MJRLE patients from baseline to intervention, from 91.7 per quarter in the baseline to 115.0 per quarter in the intervention period. This compares with an average increase of 1.4% for the hospital EIs with the lowest NPRA per MJRLE episode, where their volume increased from 62.9 per quarter in the baseline to 63.8 per intervention quarter. The

larger increase in MJRLE volume among EIs with the highest NPRA per MJRLE episode was driven by an increase in non-fracture procedures. Please note that the source of the increase in volume among BPCI hospitals (e.g., increases due to general market trends, shifts to BPCI hospitals, or potential induced demand) cannot be derived from these results alone as this analysis did not consider overall volume within markets. Our separate analysis of the impact of BPCI on hospital market share indicated that Model 2 hospital EIs did not change their market share of MJRLE clinical episodes under BPCI. Taken together, these two analyses suggest that the increase in volume among BPCI hospitals was seemingly matched by an increase among non-BPCI hospitals.

D. Discussion

In this third annual report on the evaluation of Models 2, 3, and 4 of the BPCI initiative, we are beginning to see the impact of changes providers made in response to the BPCI incentives. The effects are far from clear or straightforward, however. The lack of consistent or significant results may be partly due to the short average tenure of participants in the initiative. The data in this report are based on EIs with an average of three quarters of BPCI experience, which may not be enough time to see results on payments and quality from care redesign. Even so, patterns across Models and episodes indicate that participants are reducing the intensity of PAC use.

The providers that chose to participate in BPCI are a self-selected group that are generally larger, urban, and in more affluent areas. Providers also were allowed to choose among alternative ways to participate in BPCI. Their choices of specific Model, episode length, and clinical episodes were based on their assessment of where their organization could make the most advantageous changes and have the most impact. In particular, there is a consistent pattern of participants entering into clinical episodes in which they had higher than average baseline payments, which may indicate that they had the most inefficient patterns of care that would be easier to change. The implication of these choices, which are consistent with expectations in a voluntary model, is that the impacts of BPCI are likely to be in part limited to higher cost, less efficient episodes or providers.

It is not surprising that the initiative has seen the most notable impacts on episode payments for MJRLE under Models 2 and 3. MJRLE is a relatively high volume procedure with substantial PAC use. It is usually an elective surgery, so providers can plan the episode of care and prepare the patient. In addition, MJRLE patients can generally be identified at the beginning of or prior to the hospital admission. This is important to ensuring that care protocols can be implemented in a timely manner, which has been an issue for other clinical episodes because providers indicated that they often did not know for days after admission whether a patient was in a BPCI episode or not. Interestingly, reductions in episode payments were even more pronounced for MJRLE due to fracture, which is a higher cost episode with greater PAC use, but is generally not a planned or an elective surgery.

The BPCI initiative has had its greatest impact on PAC payments and use. Medicare spending for PAC is quite variable, even within the same clinical episode. The variability may be due to inefficiency in PAC use and the lack of clinical consensus over what constitutes appropriate care for rehabilitation or recovery following a hospitalization. Therefore, reducing PAC spending may be the most viable approach to reducing total episode payments. We have seen reductions in hospital discharges to SNF and particularly IRF, both of which typically have higher Medicare

payments than HHA. We have also seen declines in SNF length of stay. Under Model 2, reducing PAC spending is particularly important in achieving positive NPRA because the hospital payment, which is often the largest component of the episode payment, is a per discharge amount. Reducing resources used during the hospital stay can contribute to internal cost savings for the hospital, but is unlikely to affect Medicare's payment (unless the hospital length of stay falls below a limit that triggers a per diem payment). The same is true for HHA payments. Because Medicare pays HHAs for a 60-day episode of care, reducing HHA visits may generate internal cost savings for the agency, but is unlikely to reduce Medicare payments (unless the number of visits falls below a low volume episode limit). SNFs, by contrast, are paid a daily amount, so the decline in SNF length of stay that we have observed in many clinical episodes translates into lower Medicare payments.

While our analysis shows some significant changes in quality of care outcomes for a few Model 2 clinical episodes, the lack of consistency for any clinical episode across the claims, assessment, and patient survey measures reduces concerns about systematic problems. In addition, there have been few statistically significant changes in service use among clinical episodes with declines in quality measures, so it is unclear how these outcomes could be due to BPCI. That said, MJRLE clinical episodes under Model 2, which showed the most significant changes in service use, also showed significant relative improvements in patient-reported outcomes.[3]

For several SNF Model 3 clinical episodes, readmissions, mortality, and emergency department use increased relative to the comparison group. It is unclear what is causing these statistically significant changes in quality outcomes, but it is important to keep in mind that the results are based on a small number of EIs and patients and they were not consistently attributed to the same clinical episodes. We will continue to monitor these outcomes, track whether they are consistent over time and across clinical episodes, and evaluate them in the next annual report with an additional four quarters of experience. At the same time, however, among the subgroup of patients who remained in the SNF long enough to receive two patient assessments, the functional assessments administered in the SNFs indicated improvements for patients in BPCI episodes, relative to the change for comparison patients.

Differences between hospital EIs that achieved the highest and the lowest NPRA for their MJRLE episodes provide further insights into how participants respond to BPCI. Hospitals that received the highest relative NPRA had higher baseline institutional PAC use in their MJRLE episodes, which may have facilitated greater PAC payment reductions under the initiative. This is consistent with evidence from site visits and interviews in which representatives of hospitals with the highest NPRA discussed their strategies for reducing SNF use by substituting HHA care. Further, from the baseline to the intervention period, MJRLE patients in hospitals with the highest NPRA became less severe across several measures, and the share of fracture patients in these hospitals declined compared with hospital EIs with the lowest relative NPRA. This may have been because their volume of MJRLE discharges increased, primarily because of an increase in non-fracture patients who are typically less severe than fracture patients.

[3] Subsequent results indicated a reversal of the positive changes in functional status among Model 2 BPCI respondents with MJRLE episodes.

There are several limitations with this evaluation. The primary analytic approach relies on the differential change in claim-based and patient assessment-based measures between the clinical episodes of BPCI participants and a comparison group to infer the impact of BPCI. The strength of these results therefore is dependent on how well the changes in the comparison group episodes represent what would have happened absent the BPCI initiative. Some provider and patient characteristics, however, cannot be captured through administrative data, which limits our ability to match on all factors that may have influenced participation in BPCI and success under the initiative. Further, the evaluation estimates the average impact of BPCI, so the range of effects of the initiative across participants are not considered. In addition, although the BPCI impact estimates account for differences in patient mix and provider and market characteristics, we were unable to control for factors that were unobservable with existing data. Also, the changes in assessment-based quality outcomes were limited to those patients who were in the PAC setting long enough to have two assessments and, therefore, they are not representative of all patients discharged to PAC. This evaluation also does not include analyses of the impact of BPCI on the volume of clinical episodes, nor does it include an analysis of the change in Medicare program outlays, which would have required accounting for any volume change due to BPCI and NPRA. Finally, it is important to keep in mind that providers are responding to multiple initiatives and incentives in the changing health care environment. This context makes it difficult to isolate the effects of BPCI.

The BPCI initiative was designed with many options to encourage participation and also to allow CMS to relatively quickly assess responses to payment incentives across a range of situations. This strength of the initiative, however, contributes to the main limitation of its evaluation. Because of the vast range of situations encompassed under the initiative, including the selective and heterogeneous group of participants and their limited and varied experience, it is challenging to reach conclusions about the overall impact of BPCI. Of equal concern, however, is that because we are measuring multiple outcomes across the range of Model, participant, and episode combinations, by chance alone some results will appear significant, although in reality they are not true effects of the initiative. This increases the importance of relying on our mixed methods approach and triangulating results across analyses. Quantitative results from claims, patient assessments, and beneficiary surveys, combined with information gleaned from site visits, interviews, and insights from clinical experts, together provide a strong evaluation of BPCI. Consistency across findings lends strength to our conclusions, while inconsistencies raise questions for further inquiry. This year we have been able to compare results across similar episodes and across Models for the same clinical episodes. These comparisons likewise add strength to conclusions and illustrate the variations in impact. Limitations related to sample size and tenure will continue to be mitigated as the BPCI participants have more time under the initiative.

E. Conclusions

The mixed methods evaluation we have employed indicates that BPCI participants have responded to BPCI incentives, but there are relatively few instances in which these responses significantly changed key outcomes. Because of the large number of situations encompassed under the initiative, including the selective and heterogeneous group of participants and limited and varied experience of participants, it is challenging to reach conclusions about the overall impact of BPCI. It is also important to keep in mind that the kind of changes envisioned under the initiative often

need to occur within complex organizations and require collaboration across organizations that may have differing objectives.

The evaluation of the BPCI initiative is far from complete, but this third annual report identifies BPCI design features that affect initiative results and conclusions. One key feature is the target price used in calculating NPRA. Because the target price is fixed and based on historical payments, providers tended to choose clinical episodes in which they have exhibited higher than average payments and, therefore, have the greatest potential for efficiency gains. A second feature is that quality of care is not directly tied to financial incentives under BPCI. While this evaluation has not identified systematic declines in quality of care that are related to reduced services or payments, there have not been systematic improvements as might be expected with incentives to coordinate care across an episode. A third consideration has to do with the hospital infrastructure needed under Model 4 for managing physician payment, which limited its success. Another issue is that the reliance on MS-DRG assignment to determine the clinical episode has hampered some providers in Model 2 and Model 3 in determining which patients are in BPCI episodes, particularly patients who present with multiple diagnoses that complicate the ultimate assignment of the MS-DRG for the anchor or qualifying hospitalization. The delay in identifying the patients in BPCI episodes, particularly those who might benefit the most from episode-based care coordination, may diminish the impact of BPCI.

In future annual reports there will be more information about the care redesign participants are able to achieve and its impact on the cost and quality of care across clinical episodes. Future reports will also include information on PGP-initiated episodes under Models 2 and 3. PGPs are significant participants in BPCI and their results may differ from results achieved by institutional providers. Future evaluation work will focus on clarifying the effects of the BPCI initiative with additional data and experience and understanding the provider strategies that are most effective in achieving success under the initiative.

I. Introduction

The Bundled Payments for Care Improvement (BPCI) initiative is designed to test whether linking the payments for all providers involved in furnishing Medicare-covered items and services during an episode of care that is related to an inpatient hospitalization can reduce Medicare expenditures 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 that convene health care organizations, 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, three risk tracks, and implementation plans that require waivers of Medicare rules or fraud and abuse law.

BPCI Awardees can choose from four different payment models, which differ in the services included in the episode bundle and in payment method. This report describes the evaluation of Models 2, 3 and 4; Model 1 is evaluated separately.[1] Awardees in Models 2 and 3 are rewarded for reducing Medicare payments for the bundle of services in the episode relative to a target price. The target price is determined by CMS and generally based on historical payments for episodes attributed to the episode initiating provider for the same type of episode. When aggregate Medicare episode payments are less than the target price, Awardees may receive net payment reconciliation amounts (NPRA) that reflect this difference, which they can keep or share with their partnering providers. When aggregate episode payments are higher than the target price, Awardees may have to pay amounts to CMS. Thus, to obtain positive NPRA, Awardees have incentives to reduce aggregate episode payments. In Model 4, CMS makes a single, prospectively determined payment to the hospital where the episode is initiated for all services furnished by the hospital, physicians, and non-physician practitioners during the episode of care, which lasts the entire inpatient stay and any readmissions that occur within 30 days of discharge that are not explicitly excluded from the episode.

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 third 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 a Medicare Severity Diagnosis Related Group (MS-DRG) contained in one of

[1] Model 1, which concluded on December 31, 2016, defined an episode of care as the inpatient stay in the acute care hospital. Under Model 1, Medicare paid hospitals discounted payments — based on the payment rates established under the Inpatient Prospective Payment System — but continued to pay physicians separately for their services under the Medicare Physician Fee Schedule.

48 clinical episodes (see **Appendix A** for a list of the 48 clinical episodes and associated MS-DRGs).[2]

The services provided during the clinical episode are bundled for payment purposes. Hospice and certain Part A and Part B-covered services unrelated to the triggering hospitalization are excluded from the bundle, such as readmissions for certain MS-DRGs and some Part B services. The bundle and payment approach vary by Model as follows:

- **Model 2** has the most comprehensive bundle, which includes the triggering hospital stay (i.e., the anchor hospitalization) and all professional items and services (with certain exclusions) furnished within the chosen episode length of 30, 60, or 90 days post-discharge. The episode starts when a beneficiary is admitted to a participating acute care hospital (ACH) or when the attending or operating physician for the beneficiary's hospitalization is in a participating physician group practice (PGP). Individual providers are paid on a fee-for-service basis throughout the episode and aggregate episode payments are reconciled retrospectively against the target price.
- The **Model 3** bundle includes items and services furnished after the anchor hospital discharge, within the chosen episode length of 30, 60, or 90 days. 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) within 30 days of discharge from a hospitalization for a chosen clinical episode. In the case of PGP episode initiators, the episode starts when a beneficiary is admitted to a PAC setting within 30 days of discharge from a hospitalization where the attending or operating physician for the beneficiary's hospitalization is in a participating PGP. Individual providers are paid on a fee-for-service basis throughout the episode and aggregate episode payments are reconciled retrospectively against the target price.
- The **Model 4** bundle includes the anchor hospitalization, all professional services during the anchor hospitalization, and any readmissions and associated professional services that occur within 30 days of discharge that are not explicitly excluded from the bundle. The admitting hospital is paid a prospectively determined amount and it, in turn, pays the providers furnishing services included in the episode. There is no NPRA for Model 4 because participants keep any difference between the prospectively determined amount and their payments to other providers for services furnished during the episode.

There are 384 possible unique combinations of Model, clinical episodes, and participant type in BPCI. During the first two years of the initiative, patient episodes were initiated in 264 of the possible combinations (See **Appendix C** for count of patient episodes by Model and clinical episode during the first two years of the initiative). Of these combinations, only 39 had large enough participation and volume to warrant a regression-based difference-in-differences (DID) evaluation using a matched comparison group.[3] 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 level of winsorization, that is, the outliers that are excluded from the reconciliation payment calculation

[2] **Appendix B** includes an acronym list and glossary for common terms used through this report.
[3] An additional 88 combinations were for PGPs, which we were unable to analyze at the time of our analyses.

(Risk track A includes episodes whose costs fall between the 1st and 99th percentile of national payments for that DRG; B: 5th to 95th percentile; and C: 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. Phase 1 ended on June 30, 2015. All participants must have transitioned all of their chosen clinical episodes to Phase 2 by October 2015. In May 2016, CMS offered Awardees the opportunity to extend their participation in BPCI past the original three year period of performance and continue in BPCI until September 30, 2018, if they chose to sign the Extension Amendment. In August 2016, CMS announced plans to build upon the BPCI initiative and implement a new voluntary bundled payment Model for calendar year 2018. The proposed Model would be designed to meet the criteria to be an Advanced Alternative Payment Model.[4]

1. Participant roles

Organizations may participate in BPCI in several ways, depending on whether the participant is risk bearing, can initiate episodes under BPCI, and/or serves as an administrator or convener. An Awardee is a provider or other entity that has entered into the BPCI agreement with CMS and accepts financial risk. An episode initiator (EI) may or may not be an Awardee, but it is the care by the EI provider that starts 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.[5] SAs are also EIs.
- **Designated Awardee (DA) –** A DA is an entity that functions as an SA but joins the initiative under a Facilitator Convener (FC). The DA would have an agreement with CMS and assume financial risk under the Model for episodes initiated at its institution.
- **Awardee Convener (AC) –** An AC is a parent company, health system, or other organization that assumes financial risk under the Model for Medicare beneficiaries that initiate episodes at their respective Episode Initiating Bundled Payments Provider Organization (EI-BPPO). An AC may or may not be a Medicare provider or initiate episodes.

[4] Centers for Medicare & Medicaid Services. (2016). *Medicare Program; Advancing Care Coordination Through Episode Payment Models (EPMs); Cardiac Rehabilitation Incentive Payment Model; and Changes to the Comprehensive Care for Joint Replacement Model (CJR)*, 81 Fed. Reg. 148. Federal Register: The Daily Journal of the United States. Web. 10 August 2016.

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

- **Designated Awardee Convener (DAC) –** A DAC is an entity that functions as an AC but joins the initiative under a FC. The DAC would have an agreement with CMS and assume financial risk under the Model for episodes initiated at its institution.
- **Facilitator Convener (FC) –** An FC is an entity that submits a BPCI application and serves an administrative and technical assistance function on behalf of one or more DAs or DACs. FCs do not have an agreement with CMS, nor do they bear financial risk under the Model.
- **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 EIs associated with an AC or DAC and initiate episodes. EI-BPPOs do not bear financial risk directly with CMS.
- **Episode Initiator (EI) –** Under Model 2, an EI is the participating hospital where the BPCI episode begins or a participating 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 an MS-DRG for the relevant clinical episodes (anchor hospitalization). Under Model 4, an EI is the participating hospital where the BPCI episode begins. SAs and DAs are EIs. ACs and DACs may or may not be EIs themselves and also have one or more EIs under their Awardee structure.

2. BPCI programmatic flexibilities, including program rule waivers, beneficiary incentives, and gainsharing

The design of the BPCI initiative allows Awardees to utilize several programmatic flexibilities to facilitate the implementation of care redesign interventions. An EI may or may not elect to use a programmatic flexibility chosen by its Awardee. Fraud and abuse law waivers, specifically the waivers that permit gainsharing and beneficiary incentives subject to conditions specified in the BPCI participant agreement, allow BPCI Awardees to engage in certain types of financial relationships that are not allowed under existing law and regulation. The program rule waivers for three-day hospital stay, telehealth, and post-discharge home visits permit BPCI beneficiaries in episodes under the responsibility of BPCI Awardees to receive certain services under circumstances where such services would not otherwise be covered (and therefore, paid) by the Medicare program.

- **Gainsharing –** A gainsharing waiver under Models 2, 3, or 4 allows BPCI participants to share incentive payments with gainsharing partners. The gainsharing partners may include an Awardee's EIs and other providers with a gainsharing agreement with the Awardee. Gainsharing is used to offer financial incentives to providers to support Awardees' care redesign initiatives.

  Awardees must describe in their Implementation Protocol (IP) the specific methods for calculating and distributing incentive payments. Awardees have many options for customizing their gainsharing methodology. Awardees can share savings generated internally, termed 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 determine the manner in which the savings are contributed to various savings pools. The gainsharing calculations, 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 an in kind item or service to a beneficiary that is related to the episode. There must be a reasonable connection between the item or service 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.
- **Three-day hospital stay waiver –** As a general policy under Medicare rules, 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 they plan 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 according to 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 Medicare-covered home visits (1 in a 30-day episode, 2 in a 60-day episode, and 3 in a 90-day episode) by licensed clinical staff. These visits are paid under the Physician Fee Schedule.

B. Research questions

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 three years of the contract. It is organized by the three major research questions that provide the framework for our analytic approach.

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 the costs of the episodes and the quality of care for Medicare beneficiaries?

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

Each major research question is addressed separately for the three BPCI Models under this evaluation.[6] Under each major question are more detailed research questions that are addressed in this Annual Report and provide additional structure to the information presented.

Question A: 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 analyzed data submitted by Awardees in their IPs and quarterly data submissions. This information was supplemented by site visits, focus groups, and quarterly interviews with BPCI participants. We developed market profiles with data from the Provider of Service (POS) files, 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.

- **Participant characteristics –** We described 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 documented characteristics that we hypothesized may affect their ability to redesign processes of care, negotiate with potential provider partners, control costs, or influence admitting or ordering physicians. We documented the choice of Model, episodes, and episode length as well as characteristics that may imply the participants' readiness for BPCI, such as prior experience with care redesign and payment incentives.
- **Market characteristics –** The structure of the health care market may affect participants' ability to develop relationships with other providers or partnerships to deliver care across the episode more efficiently. Market structure may also affect the care redesign opportunities for BPCI participants. We compared characteristics of BPCI-participant markets to characteristics of markets without a BPCI participant. We examined the overall competitiveness among providers, the availability of various types of providers, and Medicare managed care penetration.
- **Characteristics of initiative structure –** We described the various roles of Awardee Conveners within the initiative, the participants' relationships with partners, and the use of program rule waivers, beneficiary incentives, and gainsharing. These design choices may affect participants' ability to achieve the initiative's objectives.
- **Care redesign and cost saving strategies –** BPCI is intended 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. We analyzed the strategies Awardees documented in their IPs, supplemented with data from Awardee interviews and site visits.
- **Implementation challenges –** We analyzed the challenges participants faced in implementing BPCI. We gathered this information through our qualitative data

[6] Research Question C is only addressed for Model 2 in this report.

collection activities, including site visits, Awardee interviews, as well as exit interviews with Awardees who terminated their participation in BPCI.

- **Characteristics of participants that terminated or withdrew from BPCI –** We used exit interviews to obtain information about why Awardees dropped out of the initiative, which may provide insights into the scalability of this initiative. We also used secondary data sources to compare the EIs that withdrew from BPCI to those who remained.

Question B: What is the impact of the BPCI initiative on the costs of episodes 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 BPCI initiative and how those responses affect Medicare 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 of the bundle. We examined changes in the costs of care using Medicare standardized allowed payments.[7] We accounted for differences in patient need for services through risk adjustment based on characteristics available in claims data.
- **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, and functional status. We also assessed the impact of BPCI on beneficiaries' experiences with care and improvements in functional status through the beneficiary survey.
- **Change in Patient Mix –** With respect to beneficiaries historically cared for by BPCI participants, the participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce episode payments below their target price. However, it is unclear, particularly for Model 2 and 4 hospitals, how participants could directly influence any change in patient mix. Regardless of the ability for the participants to select healthier patients after joining BPCI, changes in patient mix could potentially impact the participant's NPRA. To evaluate if there was a change in patient mix, we compared the characteristics of beneficiaries treated by BPCI providers over time relative to the comparison group. In addition to demographic characteristics, we examined the number of Hierarchical Condition Categories (HCC) indicators and utilization prior to the anchor stay. For Model 3, we also examined information from the initial patient assessment, such as conditions at admission and functional status.
- **Market dynamics –** Participation in BPCI might result in a shift in market share for one or more clinical episodes. To measure this phenomenon we calculated the market share of BPCI participants over time to determine whether BPCI participants captured a greater share of BPCI-eligible discharges after starting participation in BPCI. We also examined whether Model 2 hospital EIs tended to discharge their patients to a smaller

[7] 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 considered payments that occur up to 60 days beyond the end of the episode.

set of PAC providers, which may reflect participant care redesign efforts to encourage patients to use higher quality or more efficient PAC providers.

Question C: What initiative, provider, beneficiary, and environmental factors contributed to the various results of the BPCI initiative?

There will be a range of responses to the BPCI initiative and a range of effects across participants. The analyses to address this research question are intended to identify the factors—either in participant characteristics or participant responses—that distinguish among participants that achieved the initiative's objectives and those that did not. We investigated the initiative, provider, beneficiary, and environmental factors that contributed to BPCI initiative success or failure. We assigned EIs to performance categories based on their realized NPRA for select Model 2 clinical episodes. We compared the top and bottom performers on their outcomes during the baseline period, changes in the composition of patient mix before and after BPCI implementation, shifts in payment and utilization patterns before and after BPCI implementation, and contextual factors gathered from interviews and site visits (e.g., relationships with PAC providers, established referral patterns within the market).

II. Methods

A. Data sources

1. Secondary data

Exhibit 1 lists the secondary data sources and their uses for this study. Overall, we used provider-level data sources to identify and describe BPCI participant providers and select comparison providers. Medicare claims and enrollment data were used to construct episodes of care for patients at BPCI-participating sites (BPCI population) and at matched comparison providers. We also used claims and patient assessment data to create outcome measures and beneficiary risk factors associated with the outcomes.

Exhibit 1: 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	2013-2016	Information compiled by CMS on BPCI participants and future participants and their clinical episodes, including participant name, CMS Certification Number, location, type (ACH, SNF, etc.), BPCI "role", Model, clinical episode(s) and length(s), BPCI participation start and end dates, and contact information.	Used to identify Quarter 4 2013 through Quarter 3 2015 BPCI participating providers and clinical episodes. Identified potential future participants, and those participating in Model 1 of BPCI, 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	2011-2015	Information on Medicare-approved institutional providers, including provider number, size, and staffing.	Used within descriptive analysis of BPCI and non-BPCI participants. Used as 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. Used as predictors in provider propensity model on participation in BPCI or characteristics for Mahalonobis matching.
	Implementation Protocols	2013-2016	Information provided by an Awardee to CMS when joining BPCI (may be updated quarterly with any changes). The Awardee describes their care redesign activities, notes whether they will be participating in the OIG or CMS waivers, provides SNF partner lists, and lists beneficiary incentives and gainsharing methodology, if relevant.	Used to identify the count and percentage of Awardees and EIs participating in various care redesign activities and utilizing the waivers. Used as potential characteristics of interest when evaluating what BPCI characteristics are associated with success/failure in the initiative.
	Master Data Management (MDM)	2013-2016	Provider- and beneficiary-level information on participation in CMMI payment demonstration programs.	Used to identify providers who are involved in an Accountable Care Organization (ACO) or other Medicare Shared 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
Transaction-level data sources	Medicare fee-for-service (FFS) Claims	Jan 2010-Dec 2015	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 care setting (e.g., SNF). Also used to create risk factors including HCCs and health care utilization prior to anchor/qualifying hospitalization.
	Medicare standardized payments	Jan 2011-Dec 2015	Medicare standardized payments for 100% of Part A and B claims received via the Integrated Data Repository (IDR) from another CMS contractor.	Used to create Medicare standardized payment amounts (Part A and B) and allowed standardized payment outcomes (including beneficiary out-of-pocket amounts).
	The Master Beneficiary Summary File (MBSF)	Jan 2010-Dec 2015	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	2011-2015	Comprehensive post-acute patient assessments completed by clinicians. Required for residents of Medicare-certified SNF facilities. Administered at entry to the facility, at discharge, days 14, 30, 60, 90, and quarterly thereafter.	Provided functional status outcomes (early-loss, mid-loss, and long-form activities of daily living) for BPCI and comparison groups. Provided conditions and functional status upon admission to SNF within Model 3.
	Outcome and Assessment Information Set (OASIS) patient assessments	2011-2015	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. Provided conditions and functional status upon admission to HHA within Model 3.
	Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI)	2011-2015	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 is not required.)	Provided functional status outcomes (self-care and mobility) for BPCI and comparison groups.

2. Beneficiary survey

The BPCI beneficiary survey was used to explore differences in patient care experiences and functional outcomes between Medicare beneficiaries cared for by BPCI providers and similar beneficiaries whose providers do not participate 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, patient education) or that are not available for patients in all care settings (e.g., functional status measures). In this report, we present results from Waves 2-5 of the BPCI beneficiary surveys, which covered beneficiaries receiving services from participating providers in fall 2014 (Wave 2), spring 2015 (Wave 3), summer 2015 (Wave 4), and fall 2015 (Wave 5), respectively.[8] This section describes the instrument, sampling, and administration of the beneficiary survey.

a. Survey instrument

The survey instrument (included in **Appendix D)** was based on items adapted from validated survey instruments, such as the CARE Tool[9], National Health Interview Survey[10], and Short Form 36 Health Survey[11]. New questions underwent cognitive testing with a small convenience sample of Medicare beneficiaries with recent hospital and PAC experience.

The beneficiary 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 (Exhibit 2). For each of 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 or 4 episode or PAC initiation that starts a Model 3 episode.

[8] We do not include Wave 1 in this report, because we focus only on waves where we collected results for Model 4 overall and Models 2 and 3 at the clinical episode level.

[9] Gage et al. (2012). The Development and Testing of the Continuity Assessment Record and Evaluation (CARE) Item Set.

[10] Centers for Disease Control and Prevention. (2012). National Health Interview Survey.

[11] Brazier et al. (1992). Validating the SF-36 health survey questionnaire: new outcome measure for primary care, *BMJ*, 305(6846), 160-164.

Exhibit 2: Domain and Survey Items for Beneficiary Survey

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	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
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 used 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 and 310 comparison responses, given expected response rates (see Power calculation and response rate assumptions below)**.** Prior to Wave 4, we sampled Model 2 and Model 3 beneficiaries by group of clinical episodes, with the exception of Model 2's major joint replacement of the lower extremity (MJRLE) clinical episode, which had enough cases in a single wave to support a separate stratum in Waves 2 and 3. For Wave 4 and beyond, we sampled Model 2 and 3 beneficiaries at the clinical episode level only, and we also surveyed all Model 4 BPCI beneficiaries. For strata that did not have enough episodes to complete in a single wave, we drew a census of BPCI beneficiaries in each wave and continued to do so until the stratum achieved at least 310 BPCI and 310 comparison responses. Strata that did not complete in a single wave are referred to as "open tab" strata in this report. For these strata we "opened a tab," taking a census of BPCI beneficiaries, and then continued sampling across waves until the target number of completes

was reached. The combined results for all waves that make up a stratum are presented on the last wave that completed the desired sample size.

The Wave 2 survey had four sampling strata:

- Model 2 MJRLE episodes
- Model 2 "non-surgical: cardiovascular" episodes [12]
- Model 2 "non-surgical: respiratory" episodes [13]
- A single stratum including all Model 3 episodes

The Wave 3 survey had four sampling strata:

- Model 2 MJRLE episodes with an ACH EI
- Model 2 ACH "non-surgical: cardiovascular and neurovascular" episodes[14]
- Model 2 ACH "non-surgical: respiratory" episodes[15]
- A single stratum including all Model 3 episodes from SNFs and HHAs

The Wave 4 survey included 20 Model 2 sampling strata at the clinical episode level (17 for ACH EIs and 3 for PGP EIs), 29 Model 3 sampling strata at the clinical episode level (25 for Model 3 SNF EIs, 3 for HHA EIs, and 1 for PGP EIs), and one aggregate Model 4 sampling stratum, most of which were open tabs. Below we report the strata that were completed and analyzed at the close of Wave 4:

- Model 2 ACH MJRLE episodes
- Model 2 ACH congestive heart failure (CHF) episodes
- Model 2 ACH chronic obstructive pulmonary disease, bronchitis, asthma (COPD) episodes
- Model 2 ACH simple pneumonia and respiratory infections (SPRI) episodes
- Model 2 ACH sepsis episodes

The Wave 5 survey included 28 Model 2 ACH sampling strata at the clinical episode level, 6 Model 3 SNF sampling strata at the clinical episode level, 2 Model 3 HHA sampling strata at the

[12] This clinical episode group includes the following episodes: acute myocardial infarction, cardiac arrhythmia, atherosclerosis, chest pain, medical peripheral vascular disorders, syncope & collapse, and congestive heart failure.
[13] This clinical episode group includes the following episodes: chronic obstructive pulmonary disease, bronchitis, asthma; other respiratory; and simple pneumonia and respiratory infections.
[14] This clinical episode group includes the following episode groups: acute myocardial infarction, cardiac arrhythmia, atherosclerosis, chest pain, medical peripheral vascular disorders, syncope & collapse, congestive heart failure, and stroke and transient ischemia.
[15] This clinical episode group includes the following episode groups: chronic obstructive pulmonary disease, bronchitis, asthma; other respiratory; and simple pneumonia and respiratory infections.

clinical episode level, and one aggregate Model 4 sampling stratum, most of which were open tabs.[16] Strata that were completed and analyzed at the close of Wave 5 were:

- Model 2 ACH MJRLE episodes
- Model 2 ACH CHF episodes
- Model 2 ACH COPD episodes
- Model 2 ACH SPRI episodes
- Model 2 ACH sepsis episodes
- Model 2 ACH stroke episodes
- Model 2 ACH cardiac arrhythmia episodes
- Model 2 ACH urinary tract infection (UTI) episodes
- Model 3 SNF MJRLE episodes
- Model 4 all episodes combined

Within each stratum, BPCI and comparison beneficiaries were matched within cells. Matching was based on beneficiary and provider-level characteristics (see section B.3 for additional detail).

c. 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 to 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% 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 from prior waves to determine the size of the initial sample required to yield 310 completed surveys in each group and to estimate the number of waves it would take to complete each of the "open tab" strata. Beneficiaries sampled in the first two Waves were used to estimate the response rates for each stratum in Wave 3 and beyond. Estimated response rates used to determine the initial sample size were calculated as the actual observed response rate minus the margin of error. For example, if Model 2 ACH MJRLE respondents had a response rate of 74% through Wave 2, with a 5% margin of error, we estimated a 69% response rate for Wave 2 to err on the conservative side. After each wave concluded, we updated the estimated response rate for each stratum.

[16] Nineteen Model 3 SNF episodes and 1 Model 3 – HHA episode were dropped after Wave 4 because it was determined they would never reach completion by the end of the evaluation.

d. 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 Models 2 and 4, and within about 120 after PAC episode initiation for Model 3. In Wave 2, we mailed surveys to a total of 6,162 beneficiaries and 3,008 surveys were returned with at least one question answered. For Wave 3, we mailed surveys to a total of 7,341 beneficiaries and 3,651 surveys were returned with at least one question answered. In Wave 4, we mailed surveys to a total of 22,677 beneficiaries and 10,110 surveys were returned with at least one question answered. In Wave 5, we mailed surveys to a total of 23,275 beneficiaries and 10,900 were returned with at least one question answered.

3. Case study site visits

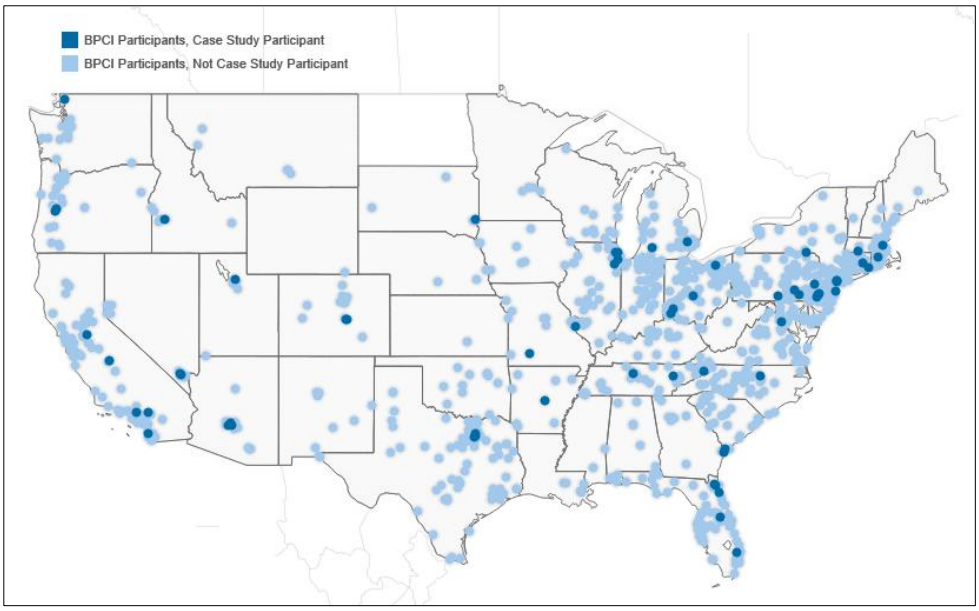
Case studies were based on two-day, in-person site visits that involved interviews with key individuals responsible for different aspects of BPCI implementation and management, including clinical and administrative leaders and operational staff, at episode-initiating sites (both Awardees and EIs under Awardees). The information collected during each site visit complements data submitted by Awardees through their IP and quarterly Awardee data submissions.

During our site visits conducted during Q4 2013 through Q3 2016 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, reconciliation results, implementation of beneficiary incentives, and both the perceived challenges and successes. We continued to ask these questions during the third year of the initiative and for some sites conducted episode-specific interviews to learn more about sites' experience with CHF and MJRLE.

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 included: Model, Awardee size, convener approach, and tenure in the initiative. Exhibit 3 displays the 66 selected case study sites on a national map with all BPCI participants. The selected sites included 32 hospitals, 21 SNFs, 5 HHAs, 6 PGPs, and 2 IRFs. Thirty-eight sites had an Awardee Convener, 17 had a Facilitator Convener, and 11 were Single Awardees. **Appendix E** includes an exhibit that compares the characteristics of the case study sites to all BPCI EIs.

Exhibit 3: Geographic Location of Case Study Participants among all Q4 2013 - Q3 2015 BPCI Episode Initiators



Note: Data points may overlap in locations with multiple participant EIs.
Source: Lewin analysis of CMS's BPCI database, as of November 2016, on BPCI participants from Q4 2013 through Q3 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 (**Appendix E**). 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. Separate interview protocols were tailored to and used for each type of respondent, ensuring consistency and appropriateness in question presentation.

During site visits, participant 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., PAC providers, physician groups), care redesign approaches, gainsharing, and reasoning behind their decisions 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. 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 management at each site through several planning calls in advance of the site visit. Convener staff were included, where appropriate, although all site visits focused on the experience of the EI. Interviewees were asked about their expected goals for their tasks related to the initiative, how their efforts differed from prior practice, how their jobs have changed, the types of materials (e.g., educational materials, care protocols, risk stratification tools) or practice programs they put in place to affect 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

We conducted focus groups for seven Model 2 participants and six Model 3 participants to collect data on care redesign across EIs under the same Awardee Convener and relationships between EIs and PAC partners. 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.

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. Information on the topic and number of participants is included in Exhibit 4.

Exhibit 4: Focus Groups conducted in 2014-2016

Topic	Points for Discussion	Model	Participants
Care redesign under the same Awardee Convener in local region	• Relationships between EIs and ACs • AC services provided to EIs • How ACs adapt their approach to address needs of the distinct EIs	2	We held three focus groups, each for a different AC. Two focus groups had three EI participants and one focus group had five EIs.
		3	We held four focus groups, each for a different AC. One focus group had four EI participants, one focus group had five EIs, and two focus groups had six EIs.
Relationship between PAC providers and EIs working together in same local region	• How EIs and PAC providers work together to change processes for 1) care delivery, 2) hospital discharge and 3) PAC admission • PAC providers' views of BPCI	2	We held five focus groups, each with PAC providers working with unique BPCI EIs. One focus group had three PAC providers, three focus groups had four PAC providers, and one focus group had seven PAC providers.
		3	We held one focus group with PAC providers working with one BPCI EI. Seven PAC providers participated in this focus group.

The Lewin team, in consultation with CMS, selected the overall topics and questions for these focus groups. A unique protocol was created for each focus group topic and split into two sections. The first section was designed to elicit important information for the evaluation related to the focus group topic and the second section addressed lessons learned. Focus groups were 90 minutes in length, were conducted in-person or virtually, and introductions were scripted to explain the goals of the focus group to participants. The protocols used for these focus groups are included in **Appendix F**.

5. Interviews with Awardees that terminated BPCI participation

Upon receiving notice of their termination from CMS, we reached out to BPCI Awardees that terminated their participation. A letter to the Awardees requested a 60-minute call with their key staff involved in the BPCI initiative to discuss their reasons for termination. Thirteen of 25 Awardees that were contacted participated in interviews.

Designed with input from CMS, the interview protocol elicited information on potential challenges that may affect the ability to scale this initiative to a broader group of providers. Awardee respondents described various topics: key factors in the decision to terminate from BPCI, suggestions for initiative improvement for future participants, and lessons learned for future success in BPCI. Interviews were typically conducted with two to three individuals and lasted between 30 to 60 minutes. The sample of Awardees interviewed and the interview protocols are included in **Appendix G**.

6. Awardee interviews

We conducted interviews with 20 to 25 Awardees or EIs each quarter to collect qualitative and quantitative data for analysis. Semi-structured interviews lasted up to one hour with the Awardees' or EIs' choice of representatives. Interviews were conducted over a period of two to three weeks prior to the end of the calendar quarter. Each quarter's sample was designed to inform a specific domain (e.g., waiver use, PAC use, care redesign). In this report, we summarize findings from the quarterly interviews conducted from Q2 2015 through Q2 2016, during which we conducted 85 quarterly interviews with a total of 108 Awardees or EIs (Exhibit 5).

Exhibit 5: Characteristics of Awardees and EIs Interviewed, Q2 2015 – Q2 2016

		Q2 2015 – Q2 2016 (N=108)	
		N	%
Model	2	62	57
	3	42	39
	4	4	4
Role	DA	23	21
	AC	8	7
	SA	14	13
	DAC	1	1
	EI-BPPO	62	57
Organization Type	ACH	33	31
	HHA	8	7
	IRF	1	1
	SNF	16	15
	PGP	50	46
Episode Initiating?	Yes	99	92
	No	9	8

Source: Lewin quarterly analysis of CMS's BPCI database for Phase 2 Awardees between Q2 2015 and Q2 2016.

The Lewin team selected interview topics, in consultation with CMS, based on themes identified through other BPCI qualitative and quantitative analyses. Exhibit 6 displays the topic, Model, and number of Awardees and EIs with whom we held interviews by quarter. Protocols for quarterly interviews are included in **Appendix H.**

Exhibit 6: Number of Awardees and Episode Initiators Interviewed by Quarter, Q2 2015 – Q2 2016

Quarter	Topic	Model(s)	Number of Interviewees
Q2 2015	SNF waiver, relationships with SNFs, and PAC utilization	2	18
Q3 2015	PAC utilization	2	12
		4	4
Q4 2015 – Q1 2016	Reasons for entry into the initiative and types of organizational arrangements and partnerships PGPs have with other entities	2	32
		3	17
Q2 2016	Care redesign efforts and cost saving strategies	3	25

7. Technical expert panels

Information obtained from three Technical Expert Panels (TEPs) provided clinician insights into patterns of care and changes in care for BPCI beneficiaries identified through case studies or quantitative data analysis. Each TEP focused on a single clinical episode; therefore, the TEP panelists represented the range of clinicians and specialists that care for that type of case. The first TEP, conducted in 2015, focused on beneficiaries with MJRLE under BPCI Models 2 and 3.[17] After consultation with CMS, CHF and coronary artery bypass graft (CABG) surgery were selected as the conditions for the two TEPs conducted in 2016 for several reasons, including their prevalence in the Medicare population. For each TEP, eight panelists of various backgrounds and expertise were identified through professional contacts and vetted by CMS.

The objectives of the TEPs were to:

- Identify potential outcomes associated with the observed care patterns.
- Identify care patterns that may suggest questionable utilization and signal potentially lower quality care.
- Identify markers of appropriate versus inadequate care.
- Identify beneficiary populations that may be susceptible to poor quality care.
- Identify outcomes measures to include in quarterly rapid cycle evaluation reports and other quantitative analyses.

a. TEP administration

All TEPs were administered in the same manner and convened via webinar. We created a pre-work packet consisting of relevant BPCI background information, panelist biographies, an agenda, general

[17] See BPCI Models 2 through 4 Evaluation Option Year 1 Annual report for additional details on the MJRLE TEP methods.

expectations for the TEP, and presentation slides that included evaluation results and the probing questions for discussion. CMS approved the materials prior to distribution to the panelists.

Through webinars, Dr. Christine LaRocca, geriatric medicine physician and medical director at Telligen, led a discussion structured on questions based on the evaluation results to date. Each question was discussed for approximately 20 minutes and all participants were given an opportunity to answer. The meetings were recorded and transcribed to ensure accurate records of the discussions.

Technical Expert Panel focusing on beneficiaries with CHF

On April 12, 2016, the TEP discussion focused on beneficiaries with CHF under BPCI Models 2 and 3; Model 4 was not discussed due to insufficient sample size to risk adjust outcomes. Experts were asked to comment on specific BPCI findings from the first six quarters of BPCI (episodes initiated between October 2013 to March 2015) to interpret data and develop new directions for analysis.[18] See **Appendix I** for a summary of the evaluation results shared with the TEP, the probing questions, the findings from the TEP discussion, and the list of panelists.

Technical Expert Panel on CABG Surgery

On July 8, 2016, a TEP panel focused on CABG under BPCI Model 2; Models 3 and 4 were excluded due to insufficient sample size to risk adjust outcomes. Experts interpreted and commented upon BPCI data from the first seven quarters (episodes initiated between October 2013 to June 2015) to develop new directions for analysis. Results were stratified to distinguish "urgent/emergent" from "elective" CABG surgeries in response to Awardee concerns for differences in quality and utilization outcomes between these groups. The MS-DRGs defining these two groups outlined in Exhibit 7 guided the TEP discussion. See **Appendix I** for a summary of the evaluation results shared with the CABG TEP, the probing questions, the findings based on the CABG TEP discussion, and the list of panelists.

Exhibit 7: MS-DRGs to Distinguish "Urgent/ Emergent" from "Elective" CABG

"Urgent/ Emergent" CABG (with PTCA* or Cardiac Catheterization)	"Elective" CABG
DRG 231 CORONARY BYPASS W PTCA W MCC	DRG 235 CORONARY BYPASS W/O CARDIAC CATH W MCC
DRG 232 CORONARY BYPASS W PTCA W/O MCC	DRG 236 CORONARY BYPASS W/O CARDIAC CATH W/O MCC
DRG 233 CORONARY BYPASS W CARDIAC CATH W MCC	
DRG 234 CORONARY BYPASS W CARDIAC CATH W/O MCC	

*Percutaneous Transluminal Coronary Angioplasty

8. Awardee-submitted data

Awardees submit data quarterly to Lewin through an online reporting platform for the following purposes: 1) track use of program rule waivers and adherence to program requirements; 2) measure quality with data not available through secondary sources; 3) document

[18] Given the timing of the TEP, this evidence did not include the episode initiators that joined BPCI in April 2015 and July 2015. The full eight quarters presented in the BPCI impact estimates presented within this annual report were not available yet.

participant characteristics; and 4) gather initiative-related information, such as progress towards implementing care redesign.

BPCI participants first submitted data in Q1 2015, describing activities that occurred between Q4 2013 to Q4 2014. [19] Thereafter, participants submitted data on a quarterly basis, with each data submission primarily based on activities that occurred in the prior calendar quarter. Data reported about activities that occurred during the first nine BPCI quarters (from Q4 2013 through Q4 2015) are included in this report.

For each data submission period, Awardees that had active Phase 2 episodes at any point in the reporting quarter 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. Exhibit 8 shows the number of Awardees required to submit data by reporting period.

Exhibit 8: Number of Awardees Required to Submit Data and the Percentage of Awardees that Submitted Data during each Data Submission Period by Model, Q4 2013 - Q4 2015

Reporting Quarter	Model 2		Model 3		Model 4	
	Required (N)	Submitted N (%)	Required (N)	Submitted N (%)	Required (N)	Submitted N (%)
Q4 2013 – Q4 2014	61	52 (85%)	20	19 (95%)	13	11 (85%)
Q1 2015	61	52 (85%)	20	20 (100%)	9	8 (89%)
Q2 2015	106	92 (87%)	55	53 (96%)	9	9 (100%)
Q3 2015	204	193 (95%)	135	127 (94%)	9	9 (100%)
Q4 2015	203	202 (100%)	118	104 (88%)	9	9 (100%)

Note: Reporting quarter is defined as the period of time that participants reported about, not the date in which the information was reported.

Source: Lewin analysis of Awardee-submitted data collected from Q1 2015 through Q1 2016 for Model 2, 3 and 4 Awardees participating in BPCI between Q4 2013 – Q4 2015.

Exhibit 9 includes all of the data elements that we collected from the Awardees.

[19] The gainsharing section in the Awardee submission was implemented during Q1 2016, at which time participants provided information about their gainsharing activities from the time they joined the initiative through Q4 2015.

Exhibit 9: Awardee-provided Data Elements

Domain	Data element	Definition
Gainsharing waiver use	Number and types of gainsharing partners	The number of ACH, SNF, HHA, PGP, physician, and other types of partners listed on the Awardee's most recent gainsharing screening list
	NPRA distribution	Indication of whether NPRA was distributed in the reporting quarter or reasons why NRPA was not distributed, the NPRA dollars distributed to gainsharing partners, and the calendar quarters in which the distributed NPRA was earned
	ICS distribution	Indication of whether ICS was distributed in the reporting quarter or reasons why ICS was not distributed, the amount of ICS realized, the ICS dollars distributed to gainsharing partners, the calendar quarters in which the distributed ICS was earned, and the source of the ICS
	Distributions to gainsharing partners	List of entities receiving gainsharing distributions, the amount of NPRA and ICS dollars received, and the approximate number of BPCI patients cared for by the gainsharing partner
Participant baseline characteristics	Patient mix	The proportion of unique patients by payer
	Prior care redesign experience	Prior experience with care redesign initiatives
	Payment incentives experience	Prior experience with payment incentives
	Electronic Health Record (EHR) use	Indication of whether the EI uses an electronic health record/electronic medical record system or whether the EI has plans to implement an EHR system in the future, and the estimated month and year of implementation
	Meaningful use functionalities	Meaningful use functionalities that are available and used through the EHR system
	Health information exchange capabilities	Availability and use of health information exchange capabilities through the EHR system, and whether those capabilities are used to exchange information with providers or beneficiaries
	Health information exchange method	Indication of the methods used to exchange information with beneficiaries and providers if the EI does not use an EHR system, or if the EHR system does not have health information exchange capabilities
BPCI-related activities	Status of care redesign	Progress in implementing care redesign activities
Quality monitoring measures	Medication reconciliation at discharge	The percentage of BPCI patient discharges for beneficiaries 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 beneficiaries 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

Domain	Data element	Definition
Beneficiary incentive waiver use	Beneficiary incentive list	List of beneficiaries receiving incentives, the person/entity administering the item/service, date the item/service was provided, and 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 creating comparison groups for each combination of Model, clinical episode, and provider type ("stratum") analyzed in this report. We also specify the BPCI population and comparison group, if any, used in each analysis.

1. BPCI study population

The set of BPCI participants that were covered in the analyses presented in this report varied depending on the type and complexity of the analysis (Exhibit 10). The most comprehensive BPCI study population included all Phase 2 EIs that were active between Q4 2013 and Q3 2015, and it was used to analyze the characteristics of the initiative. However, for other types of analyses (i.e., impact, beneficiary survey, provider referral and market share, and differences across providers) BPCI EIs in strata that did not meet minimum sample sizes for the analysis to be informative were excluded.[20]

Exhibit 10: Model and EI[21] Types Included in each Analysis

	EI Type	Number of BPCI EIs, Q4 2013 – Q3 2015*	Characteristics of the Initiative	Impact of BPCI (DiD estimates)	Beneficiary survey	Provider referral & market share	Factors Contributing to Differences across BPCI Providers
Model 2	Hospital	422	X	X	X	X	X
	PGP	277	X				
Model 3	SNF	873	X	X	X	X	
	HHA	116	X	X	X	X	
	IRF	9	X				
	LTCH	1	X				
	PGP	144	X				
Model 4	Hospital	23	X	X	X		

*The count of EIs includes all EIs that participated in BPCI, regardless of length of time in the initiative or whether any BPCI episodes were initiated at the provider.

[20] In addition, we did not conduct any impact analysis on Model 2 or 3 PGP EIs. The validated list of physicians for BPCI-participating TINs was not available at the time the analysis for this report was conducted.

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

2. Selection of providers in comparison group

The difference-in-differences (DiD) approach requires a comparison group of non-BPCI providers ("non-participants"). This comparison group needs to be similar to BPCI providers with respect to baseline characteristics that could affect their decision to participate and could be related to their performance under BPCI. Such characteristics include market-level and provider specific attributes. Because providers voluntarily enroll in BPCI, participants were likely to be different than non-participants in ways that could be correlated with patient outcomes. For example, BPCI participants may have had less efficient care in the pre-intervention period and consequently had more room for improvement relative to non-participants. If not accounted for in the analysis, this self-selection could bias the measured impact of BPCI on outcomes. Moreover, program evaluation literature indicates that treatment effect estimates can be very sensitive to untestable model assumptions when the intervention and comparison group are dissimilar in one or more dimensions.[22, 23, 24]

We constructed comparison groups for 39 Model, provider type and clinical episode strata from the universe of Medicare providers that had not signed up for BPCI in Phase 2. Strata included were generally considered to have a sufficient sample size for meaningful analysis if there were 20 EIs with 1,000 clinically relevant episodes.[25] A few strata with lower sample sizes were included if a strata came close to those criteria and had policy relevance (e.g., Model 4 episodes and episodes associated with conditions considered for additional episode payment Model initiatives by CMMI).

Comparison providers and episodes were selected in four steps. First, potential comparison providers were identified if they were eligible to be used as a comparison provider: (i) shared certain key characteristics with participating providers, (ii) would be eligible to participate in the BPCI initiative (e.g., in Model 2, ACHs that were paid under Medicare's inpatient prospective payment system), and (iii) were not affiliated with BPCI participants. Second, each BPCI-participating provider was matched with up to 15 comparison providers using a statistical matching technique to minimize the differences in the distributions of characteristics between BPCI and providers in the comparison group (See section 2 below for details on the various matching approaches that were used). Third, episodes were constructed for beneficiaries treated by matched comparison providers by following the BPCI program rules. Finally, a sample of episodes was drawn from among those identified in the previous step to match the distribution of BPCI episodes by MS-DRG and date of service. Below, we describe these steps in more detail.

[22] Dehejia, R.H. & Wabha, S. (2002). Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and Statistics, 84*(1), 151-161.

[23] Zhao, Z. (2004). Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence. *Review of Economics and Statistics, 86*(1), 91-107.

[24] Smith, J.A. & Todd, P.E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics, 125*(1-2), 305-353.

[25] Strata was considered meaningful for the analysis if there was enough participation in BPCI, but no formal power calculation was conducted to assess minimum size.

Step 1: Exclude ineligible non-participating providers

The exclusions were applied for each Model, EI type, and clinical episode separately. Providers were excluded if they met any of the following criteria:

- Would be ineligible to participate in BPCI (e.g., in Model 2, ACHs that were not paid under Medicare's inpatient prospective payment system).
- Were owned by a BPCI-participating organization.
- Participated in any of the BPCI Models (Model 1 to Model 4).
- Did not have an ownership status (i.e., government, non-profit, for-profit) or location (i.e., rural/urban) represented in the BPCI participant Model, EI type, and clinical episode combination.
- Were located in markets where BPCI participants have over half of the discharges associated with any of the 48 BPCI clinical episodes. This exclusion avoids including providers that may be exposed to "spillover effects" of BPCI in those locations, which could cause changes in utilization for other local providers that may confound the results. Such changes include the following: (i) non-BPCI beneficiaries receiving some care from BPCI participants, (ii) comparison providers adopting practices similar to BPCI participants, or (iii) BPCI affecting referral patterns in the market.
- Had fewer than five clinically relevant discharges during the matching period. These providers are excluded in order to remove providers that did not have meaningful utilization in the episode domain. This exclusion applies to participants as well as non-participants. For some strata, in which we matched on baseline variables measuring the change from 2011 to 2012, a minimum of ten discharges were required (at least five in each year). More discussion on which strata used change variables is provided below and in **Appendix J**.

Step 2: Use matching algorithms to select close matches

For each strata, we assessed the relative performance of Propensity Score Matching (PSM), Mahalanobis Distance Matching (MDM), or Coarsened Exact Matching (CEM) methods, and selected the method that performed best, assessed using criteria as described below. PSM tended to work well for Model, provider type and clinical episode combinations meeting the criteria of 20 EIs and 1,000 cases. However, for some of the combinations with a smaller number of participants, MDM or CEM methods performed best and were used (more detail on these instances are provided below).

A *propensity score* is defined as the probability of receiving the "treatment" (in this case, BPCI participation), conditional on a set of characteristics. This probability was estimated using a logistic regression model that included key factors thought to influence both the participation decision and performance in BPCI. These factors included market characteristics (e.g., population size and primary care physician to population ratios), provider characteristics (e.g., ownership status and number of beds), and performance- and practice pattern-related factors (e.g., historical Part A Medicare payments and use of PAC services). The variables considered for matching by provider type for Model 2 and Model 3 are displayed in Exhibit 11.

In some cases, transformations of the variables or a smaller set of variables were used to improve the matching diagnostics (as discussed below).

Exhibit 11: Key Variables used for Matching Provider Type for Model 2 and Model 3

Variable	Model 2: ACH	Model 3: HHA	Model 3: SNF
Ownership - Non-Profit, Government, For-Profit	X	X	X
Urban/Rural Location	X	X	X
Bed Count			X
Number of Nurses Employed by an HHA		X	
Chain Indicator	X		X
SNF in Hospital			X
Medicare Days as a Percent of Total Inpatient Days	X		
Resident-Bed Ratio	X		
Number of points out of 5 in overall rating and in three areas: Quality, Survey/Health Inspections, and Staffing (from Nursing Home Compare)			X
Disproportionate Share Percent	X		
Teaching Status	X		
Population Size of Market Area	X	X	X
Median Household Income	X		X
Medicare Advantage Penetration	X		X
Primary Care Providers per 10,000 in Market	X		
SNF Beds per 10,000 in Market	X		X
Inpatient Rehabilitation Facility in Market	X		X
Provider Market Share of the 48 potential BPCI episodes	X		X
Herfindahl Index of Hospital Market Shares	X		X
Herfindahl Index of SNF Market Shares			X
Percentage of total discharges in the 48 clinical episodes in 2011	X	X	X
Number of discharges for clinical episode in 2011	X	X	X
Number of SNF days per patient within 90 days after an ACH by clinical episode in 2011			X
Number of HHA days per patient within 90 days after an ACH discharge by clinical episode in 2011		X	X
Percent of patients in 2011 that went home with no post-acute care by clinical episode	X		
Percent of patients in 2011 that used an inpatient rehabilitation facility as first post-acute care setting by clinical episode	X		
Percent of patients in 2011 that used a SNF as first post-acute care setting by clinical episode	X		
Percent of patients in 2011 that used a long-term care hospital as first post-acute care setting by clinical episode	X		
Percent of patients in 2011 that went home with HHA services as first post-acute care setting by clinical episode	X		
Unplanned readmission rate by clinical episode in 2011	X	X	X

Variable	Model 2: ACH	Model 3: HHA	Model 3: SNF
Change in unplanned readmission rate by clinical episode from 2011 to 2012*	X	X	X
All-cause mortality rate in 2011 by clinical episode	X	X	X
Change in all-cause mortality rate by clinical episode from 2011 to 2012*	X	X	X
Average 90-day standardized Medicare Part A payment amount by clinical episode in 2011	X	X	X
Change in average 90-day standardized Medicare Part A payment amount by clinical episode from 2011 to 2012*	X	X	X

* Variables indicating the change in outcomes from 2011 to 2012 were added to initial model specifications if it was found that these trends did not match well, when including only the 2011 levels in the matching algorithms.

Using the coefficients from the logistic regression model, we constructed a propensity score as the predicted probability of participating in BPCI. Using propensity scores, each BPCI-participant was matched with up to 15 comparison providers with a propensity score absolute difference below a defined caliper. A caliper acts as a constraint on the "distance" between BPCI and potential comparison providers based on the difference in their estimated propensity scores. Any comparison provider outside of the "caliper" of a BPCI provider would not be matched to that BCPI provider. In some cases, the use of a caliper excluded all potential matches, which excluded the unmatched BPCI provider from the analysis. These providers typically had outliers measured in several of the key factors used for matching, such as the number of discharges for the episode or the share of BPCI episodes in the market. Calipers were chosen based on the standard deviation of the estimated log-odds propensity score. Multiple calipers were tested for each strata to identify the specification that generated the most similar comparison group across all of the attributes considered important for matching. Finally, a comparison provider was allowed to be used as a match for more than one participant.

The key diagnostic used to determine similarity between BPCI and comparison providers was the standardized difference in the mean of each of the matching variables between participants and non-participants. The standardized difference compares the differences in means in relation to the pooled standard deviation. The average standardized difference across all variables was computed to assess overall balance. The method that yields the lowest standardized difference of means across the largest number of covariates and that results in the fewest number of "large" standardized differences (i.e., greater than 0.20) is typically preferred.[26] Particular emphasis was given to matching well on performance-related variables: 90-day standardized Medicare Part A payment, unplanned readmission rates, and mortality rates. Standardized differences below 0.10 were targeted for these variables.

We initially started the propensity score matching with the variables identified in Exhibit 11 above. For certain Model, provider type, and episode combinations, the standardized differences were large based on the previously defined threshold (either across all variables or for specific key variables) regardless of the caliper used. In these cases, we tested alternative specifications of variables used in the model in order to improve matching. For example, a categorical variable (as opposed to a continuous variable) was used to measure baseline Medicare payments relevant

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

to percutaneous coronary intervention and cardiac valve episodes for Model 2. If the alternative specifications using propensity score matching did not result in better matches, we then tried alternative matching methods including MDM and CEM.

We found MDM to yield a better comparison group for Model 3 HHA CHF strata. The Mahalanobis distance between two sets of covariates Xi and Xj, 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. Each BPCI provider was matched to the 15 providers in the non-participant group with which it has the lowest (i.e., nearest) Mahalanobis distance. We excluded one BPCI HHA participant from the Mahalanobis distance matching, as this provider was not matched under any PSM specification and it was an extreme outlier in terms of the number of nurses associated with the HHA and the number of relevant cases treated by the HHA. Given the outlier status, we did not think there was a comparable HHA in the country.

Given that there were only seven hospitals with enough cases to be eligible for the analysis in the Model 4 CABG clinical episode, CEM was used.[27] With so few participants, matching on the wide array of variables typically used for hospital-related strata (see list of variables for Model 2 ACH in Exhibit 11) was challenging. Instead, the matching was focused on the key performance variables, namely the 2011 levels and change from 2011 to 2012 in unplanned readmission rate, mortality rate, and 90-day standardized Medicare Part A payment. The CEM approach allowed for specifying the degree of standardized differences tolerated for each of these variables. Using this approach, the standardized differences were constrained to 0.10 for each of the key variables.

Appendix J shows the calipers chosen for each model as well as the standardized differences of each covariate included in the models between BPCI providers and matched comparison providers for each strata. As shown in **Appendix J,** our ability to construct comparison groups varied across strata.

Step 3: Construct episodes for matched comparison providers

The BPCI episode algorithm rules were applied to construct simulated episodes that would have been assigned to comparison facilities if they were participating in BPCI. We constructed simulated episodes from October 2010 through September 2015.

Step 4: Select random sample of comparison group episodes

Among all episodes identified in the previous step, we drew a random sample of comparison group episodes. Each BPCI episode was randomly matched to one episode from a pool of comparison episodes in the same quarter with the same MS-DRG, originating from the comparison providers that were matched to the BPCI participant. In the case of the MJRLE clinical episode, matches were also constructed to take into account patients with fractures versus those without fractures. The matched comparison was then excluded from the pool of episodes eligible for future matching. In some cases, the comparison pool did not contain enough episodes resulting in unmatched participant episodes. Sensitivity analyses were performed to test the DiD using both the matched and unmatched episodes.

[27] Iacus, S.M., King, G., & Porro, G. (2012). Causal Inference without Balance Checking: Coarsened Exact Matching. *Political Analysis, 20*(1), 1-24.

3. Beneficiary survey sample

This annual report includes the results for four survey waves that were conducted between October 2014 and November 2015. In this section, we describe the creation of the samples used for Wave 2 (October and November), Wave 3 (February and March 2015), Wave 4 (May and June 2015) and Wave 5 (October and November 2015) for both BPCI and comparison survey samples.

a. Sampling frame

For Wave 2, all Model 2 strata were constructed using Medicare FFS claims from two "rolling" one month samples; the beneficiaries in the two rolling one month samples received their surveys one month apart.[28] For example, for the first rolling month of Wave 3, claims for February 2015 were pulled in early March 2015 and surveys were mailed in the first week of May 2015. For the second rolling month of Wave 3, claims for March 2015 were pulled in early April 2015 and surveys were mailed the first week of June 2015. This rapid sampling process was deliberately used to reduce recall bias. It does, however, limit the sample to patients whose claims were filed quickly, within one month of discharge.[29]

In all waves, 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. Two 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 12 summarizes the periods used to create the sample frame of the BPCI beneficiary survey Waves 2-5.

Exhibit 12: Episode period for beneficiary survey sample frame, Models 2 and 3, waves 2 through wave 5

Wave	Model 2	Model 3
Wave 2: Fall 2014	Two "rolling" one-month periods: October & November 2014 hospital discharges	One two-month period: October & November 2014 admissions to PAC
Wave 3: Spring 2015	Two "rolling" one-month periods: February & March 2015 hospital discharges	One two-month period: February & March 2015 admissions to PAC
Wave 4: Summer 2015	Two "rolling" one-month periods: May & June 2015 hospital discharges	One two-month period: May & June 2015 admissions to PAC
Wave 5: Fall 2015	Two "rolling" one-month periods: October & November 2015 hospital discharges	One two-month period: October & November admissions to PAC

Starting with Wave 4 we began defining all Model 2 and 3 strata at the clinical episode level using open tab sampling (see Section II.A.2). To select open tabs, we reviewed episode volume for combinations of all 48 clinical episodes with three EI types (ACH, SNF, HHA) across

[28] One month of claims was not adequate to reach the necessary sample size at the levels of clinical precision used to define the strata.

[29] 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.

Models 2-4.[30] Combined with observed response rates from prior waves, we estimated the number of completed surveys we could obtain in each survey wave, and projected the number of waves necessary to reach our target of 310 BPCI and 310 comparison completed surveys. Any stratum projected to reach the number of completes by the conclusion of the final survey wave was included in the sampling frame.

b. Sample construction

Survey samples were constructed in three steps. First, we excluded any beneficiary who had already responded to one of our BPCI surveys in a previous wave from the sampling universe. Second, we excluded certain BPCI and comparison group providers from the sample frame to improve the comparability of BPCI and comparison samples on the basis of key provider characteristics. 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. These characteristics included provider type (ACH, 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 did not include BPCI or comparison group providers. These strata that included only BPCI or comparison group were dropped if they had enough beneficiaries to complete in a single wave. To avoid losing BPCI episodes in the "open tab" strata, we dropped all provider strata with comparison beneficiaries but no BPCI beneficiaries, but did not drop provider strata with BPCI beneficiaries and no comparison beneficiaries, because comparison beneficiaries similar to those BPCI beneficiaries could be found in later waves.

Third, we created cells within each of the sampling strata (i.e., combinations of Model, EI type, and clinical episode) and then matched beneficiaries within each cell by provider and patient characteristics. When defining the cells, we aimed to strike a balance that: 1) matched on factors that would most affect survey responses; and 2) had a sufficient number of episodes in each cell to support valid comparisons. We ordered the factors that could be used for defining the cells from most to least likely to influence outcomes, and then combined factors to create sample cells according to the two criteria above. In Waves 2 and 3, cells were defined by beneficiary age group, BPCI clinical episode, provider size, and for Model 3, provider type.[31] For Waves 4 and 5, for strata that were large enough to complete in a single wave (i.e., "regular" strata), cells were defined by presence or absence of a major complication or comorbidity (MCC), patient age group, provider size, and provider academic affiliation. And for the "open tab" strata that did not complete in a single wave, cells were defined by MCC, patient age group, and provider size. Since Model 4 was too small for episode level sampling, we aggregated across all episodes and created cells for Model 4 according to MCC and clinical episode only.

[30] PGP EIs were excluded from consideration in Models 2 and 3 because validated list of physicians for BPCI-participating TINs was not available at the time the analysis for this report was conducted.

[31] Waves 2 and 3 had a single stratum at the clinical episode level. In this stratum, cells were defined by beneficiary age, provider size, and provider academic status.

C. Outcome measures

In this section we define the various measurement periods during which we define the outcome measures and summarize the outcome measures presented in the results section.

1. 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 episode length as they are defined relative to the BPCI bundle period (i.e., pre-bundle, post-bundle). In contrast, the patient timeline measurement periods are consistent across Models and episode lengths because they depend on the patient's transition through the episode of care (e.g., post-hospital discharge), allowing us to compare outcomes across Models and episode lengths.

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 two patient timeline measurement periods: within 30 days and within 90 days of hospital discharge. These measurement periods are labeled 30 day post-discharge and 90 day post-discharge. Exhibits 13 and 14 describe the bundle and the patient timeline measurement periods for each Model.

Episodes were dropped from measure denominators when there was not enough claims run-out to cover the measurement period. Specifically, if the end of our current observational period (December 31, 2015) occurred within the *measurement period* for the given episode, we dropped the episode from the denominator. For example, if a beneficiary had a Model 2 episode begin on September 23, 2015 and had a post-discharge period beginning October 6, 2015, we dropped the episode from any 90-day post discharge measures since the 90-day post-discharge period extends beyond December 31, 2015. As a result of these exclusions, the number of episodes included in each measure varies.

Exhibit 13: Definition of Measurement Periods Relative to the Bundle Timeline across Models

Model	Pre-bundle period	Bundle start	Bundle end	Acute services within bundle	Post-acute services within bundle	Post-bundle period
2	Anchor IP stay admission date minus 30 days	Anchor IP stay admission date	Anchor IP stay discharge date plus bundle length (30, 60, or 90 days)	Anchor IP stay from IP admission date to IP discharge date[b]	From IP discharge date to bundle end date	30, 90, 120, and 180 days after the end of the bundle
3	EI PAC admission date minus 30 days	EI PAC admission date	EI PAC admission date plus bundle length (30, 60, or 90 days)	N/A	N/A	30, 90, 120, and 180 days after the end of the bundle
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	30 and 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.

[c] IP = inpatient

Exhibit 14: Definition of Measurement Periods Relative to the Patient Timeline across Models and episode lengths

Model	Pre-admission	Inpatient hospitalization	Post-discharge period (PDP)
2	30 days prior to anchor/qualifying hospital stay	Anchor/qualifying IP stay from IP admission date to IP discharge date	From anchor/qualifying IP discharge date
3			
4			

2. Outcome definitions

In this section we present the outcome measures that were constructed and analyzed to evaluate the impact of BPCI during the first two years of the initiative. Exhibit 15 summarizes the key outcome measures by domain. **Appendix K** provides detailed definitions of each outcome measure.

Exhibit 15: 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[a]	Total Medicare standardized allowed payment for inpatient stay plus 90 days post-discharge	X				
	Total Medicare standardized payment for inpatient stay plus 90 days post-discharge (no beneficiary cost sharing)	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, 60, 90, 120, and 180 day post-bundle period	X				
	Total Medicare Part A standardized allowed payment (by various settings)	X				
	Total Medicare Part B standardized allowed payment (by various settings)	X				
Utilization	Acute inpatient length of stay	X				
	Number of days in PAC setting (total and for SNF)	X				
	Number of home health visits	X				
	First PAC setting following inpatient discharge	X				
	Patients discharged to institution relative to discharged home with home health	X				
	Patients discharged to any PAC	X				

	Domain/Quantitative Outcomes	Medicare Claims	MDS assessments	OASIS assessments	IRF-PAI assessments	Beneficiary Surveys
Quality	Unplanned readmission rate	X				
	Emergency department use without hospitalization	X				
	All-cause mortality rate	X				
Functional status	SNF PAC setting:[b,c]					
	SNF patients who improve status or remain completely independent in long-form ADL function (a measure of overall function)		X			
	SNF patients who improve status or remain completely independent in early-loss ADL function (a measure of self-care function)		X			
	SNF patients who improve status or remain completely independent in mid-loss ADL function (a measure of mobility)		X			
	HHA PAC setting:[b,d]					
	HHA patients who improve status or remain completely independent in bathing (a measure of self-care function)			X		
	HHA patients who improve status or remain completely independent in upper body dressing (a measure of self-care function)			X		
	HHA patients who improve status or remain completely independent in lower body dressing (a measure of self-care function)			X		
	HHA patients who improve status or remain completely independent in ambulation/locomotion (a measure of mobility function)			X		
	HHA patients who improve status or remain completely independent in bed transferring (a measure of mobility function)			X		
	IRF PAC setting:[b]					
	Average change in self-care score				X	
	Average change in mobility score				X	
	Patients with improvement in bathing, dressing, using the toilet, or eating[e]					X
	Patients with improvement in walking without rest[e]					X
	Patients with improvement in use of mobility device (i.e., less frequent) [e]					X
	Patients with improvement in using stairs[e]					X
	Patients with improvement in planning regular tasks[e]					X
	Patients with improvement in physical/emotional problems limiting social activities (i.e., less frequent) [e]					X
	Patients with improvement in pain limiting regular activities (i.e., less frequent) [e]					X

	Domain/Quantitative Outcomes	Medicare Claims	MDS assessments	OASIS assessments	IRF-PAI assessments	Beneficiary Surveys
Patient Experience	Patients who have limited normal activities because of pain					X
	Patients, families, or caregivers that received conflicting advice from medical staff about treatment					X
	Frequency with which the patient received services that were appropriate for the level of care needed					X
	Patients who thought they were discharged at the right time					X
	Patients who had a good understanding of how to take care of themselves before they prepared to leave the hospital					X
	Patients who thought medical staff clearly explained how to take their medications prior to leaving hospital					X
	Patients who thought medical staff clearly explained what follow-up appointments or treatments would be needed					X
	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
	Patients who since having left the hospital thought that they and their caregivers have been able to manage their health needs					X
Market Dynamics	Percent of market share (Models 2/3)	X				
	Number of PAC providers receiving referrals, per EI (Model 2)	X				
	Proportion of patients discharged to PAC providers with high star-ratings (Model 2)	X				

Notes: This table includes quantitative outcomes only. 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.

[e] For each of the seven functional status measures in the beneficiary survey, respondents were asked to recall their functional status before their treatment episode and also to report their current status at the time they completed the survey. We calculated 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 reported to a better functional level 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 when completing the survey (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 reported to a worse functional level after hospitalization, or if the patient recalled having the worst functional status prior to hospitalization and remained in the worst status when completing the survey.

D. Analytical Methods

1. Descriptive analysis

To summarize characteristics of the initiative 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 Awardee IPs, Awardee-submitted data, POS files, and the AHRF. We also run descriptive analyses to examine if there are any differences in the use of the 3-day hospital stay waiver[32] between Medicare-only enrollees and Medicare-Medicaid Enrollees (MMEs) within the Model 2 MJRLE clinical episode.

2. Difference-in-differences analysis

The DiD approach quantifies the impact of BPCI by comparing changes in claim 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.[33] The DiD regression model incorporates data from two periods prior to BPCI implementation (baseline and Phase 1) as well as the intervention period. Phase 1 was initiated the moment BPCI was announced, and encompasses the one year period prior to the BPCI intervention period. 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. Including Phase I in the baseline would likely underestimate the BPCI effect given that participants started to prepare for the intervention during that 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 2015.[34]

Consider the following linear model to illustrate the DiD calculation in a regression framework:

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

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., baseline, transition, or intervention), and $X_{i,k,t}$ are beneficiary demographics, clinical characteristics observed before hospitalization, and provider characteristics. The vector β is a vector of regression coefficients that captures the impact of risk factors $\boldsymbol{X_{i,k,t}}$ on the outcome of interest. The regression coefficient β_1 captures any inherent, time invariant differences between the control

[32] The 3-day hospital stay waiver available to Model 2 BPCI participants allows Medicare coverage of a SNF stay for beneficiaries following a hospitalization of less than 3 days.

[33] 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 Model inducing correlation between BPCI participation and outcomes.

[34] Post-bundle payment outcomes are reported with one quarter delay. The DiD results for these outcomes use October 2013 to June 2015 as an intervention period. All other periods remain the same.

and the treatment groups, 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 $u_{i,k,t}$ represents a random error term. In this linear example, the DiD estimate is 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, clinical characteristics, and prior care use before the hospitalization, along with 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 group, gender, age and gender interactions, Medicaid eligibility status, and disability status. To control for participation in other Medicare initiatives, we used a dummy variable that indicated whether the beneficiary was in the Medicare Shared Savings Program or Pioneer ACO during their BPCI or comparison episode. We also controlled for hospital/SNF bed count and for-profit status. To control for prior health conditions, we used HCC indicators.[35] To further control for case-mix differences, we included measures of prior care use in the following settings: hospital, LTCH, SNF, nursing facility stay, IRF, hospice, HHA, psychiatric facility, and emergency department. In addition, to account for regional differences, we either used regional or state dummies.

While the same demographic and enrollment status indicators were included for all outcomes, we considered alternative aggregation levels to control for service mix, prior health conditions, prior care use, and regional characteristics (see Exhibit 16). To assess different specifications, we split the sample into a model development and a validation sample and estimated each model using data from the model development sample. We then evaluated models' 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.

[35] The CMS-HCC model is a prospective risk adjustment model used by CMS to adjust Medicare Part C capitation payments for beneficiary health spending risk. The model adjusts for demographic and clinical characteristics. The clinical component of the model uses diagnoses from qualifying services grouped into several HCC indicators. Pope, et al. (2004). Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model. *Quantitative Health Sciences Publications and Presentations*, Paper 723.

Exhibit 16: 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
Patient Demographics and Enrollment	• Age (under 65, 65-79, 80+) • Gender • Medicaid status • Disability status • Alignment to Medicare Shared Savings Program or Pioneer ACO during BPCI episode
Prior health conditions	• Alternative specifications ▪ HCC indicators from qualifying services and diagnoses from claims and data for six months preceding the anchor admission or qualifying stay ▪ HCC aggregated to 45 risk variable groups (RV-HCC) according to NQF measure 1789 (**Appendix L** shows a crosswalk from HCC indicators to RV-HCC.) ▪ HCC index, HCC indicators weighted by their relative weight in the CMS-HCC model
Utilization measures preceding the start of the anchor stay/qualifying inpatient stay	• Alternative specifications ▪ Binary indicators for utilization of ED, inpatient, SNF, nursing facility (NF)/SNF, IRF, HHA services 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
Geography	• Alternative specifications ▪ 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 binary quality outcomes (e.g., mortality rate). A Cox proportional hazard model was used to estimate inpatient length of stay. OLS was estimated for the total number of days measures (e.g., 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 skewedness.

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 providers. To control for changes in service and case-mix over time, as well as differences between BPCI and comparison 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 during the

baseline and intervention period. We tested for equality of trends in key outcomes between the BPCI participants and comparisons and found that the baseline trends in the large majority of strata and outcomes were the same.

The DiD estimate was then calculated by first taking the difference between the two scenarios for both BPCI-participating and comparison providers and thereafter taking the difference between BPCI-participating and comparison providers. 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 were actually treated. Standard errors of ETT estimation were computed using the Delta method.[36]

While we made every attempt to construct a comparison group of providers that closely matched BPCI providers in key characteristics, we could not guarantee BPCI and comparison providers would have parallel trends during the baseline period for every outcome since not every outcome could be included in the matching set and some outcomes fluctuated vastly over time. Because it was not feasible to test the null hypothesis that BPCI participants and comparison providers had parallel trends during the baseline for every outcome and every strata, we tested for parallel trends on a subset of the outcomes according to the following criteria. We tested the parallel trends for unplanned readmission rate, emergency department use, and mortality rate outcomes as well as the total payment for the inpatient stay plus 90 days post-discharge outcome in all strata. If we rejected the null hypothesis that there were parallel trends in the baseline (at the .10 level) and the DiD estimate was statistically significant (positive or negative), we attempted to find an alternative risk adjustment model where we failed to reject the null hypothesis of parallel trends. We also tested the null hypothesis of parallel trends in baseline for any additional outcomes when there was visual evidence that the direction of change from baseline to intervention for BPCI differed from the change for the comparison group. In this report we report all DiD estimates, but we include a footnote when we rejected the null hypothesis that there were parallel trends in baseline.

There are some outcomes for which we do not report the DiD estimate. We report DiD estimates for each given outcome if the sample exceeds 30 BPCI episodes during the intervention period for outcomes evaluated using duration, logistic, and OLS models. In contrast, we used a minimum of 100 BPCI episodes during the intervention period to report DiD estimates for outcomes using two-part models. Some outcomes, including IRF and LTCH payments during the 90-day post-discharge period and payment outcomes that are stratified by bundle length, suffer from small sample sizes, and consequently, DiD estimates for these outcomes were largely not reported.

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

This report includes beneficiary survey results from Waves 2 to 5. For each wave, we report results on three sets of survey measures: functional status, health status, and patient experience.

[36] 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 $Var(Y) \approx (2\mu_x)^2 Var(X)$.

We analyzed a total of 14 unique strata across the four waves (Exhibit 17; four strata in Wave 2, four strata in wave 3, five in Wave 4, and 10 in Wave 5). Nine of these strata were "regular" strata, which were sampled and completed in a single Wave. The other five strata were "open tab" strata that required pooling responses from multiple waves. Analyses were conducted for each stratum.

Exhibit 17: Strata in Waves 2 through 5 Beneficiary Survey Data Analysis

Stratum Name	Wave 2	Wave 3	Wave 4	Wave 5
Model 2 ACH – Major joint replacement of the lower extremity (MJRLE)	Regular	Regular	Regular	Regular
Model 2 ACH – Non-surgical: cardiovascular (cardio/neurovascular) episodes[37]	Regular			
Model 2 ACH – Non-surgical: cardio and neurovascular (cardio/ neurovascular) episodes[38]		Regular		
Model 2 ACH – Non-surgical: respiratory episodes[39]	Regular	Regular		
Model 3 – All episodes combined	Regular	Regular		
Model 2 ACH – Congestive heart failure (CHF)			Regular	Regular
Model 2 ACH – Sepsis			Regular	Regular
Model 2 ACH – Chronic obstructive pulmonary disease, bronchitis, asthma (COPD)			Regular	Regular
Model 2 ACH – Simple pneumonia and respiratory infection (Pneumonia)			Regular	Regular
Model 2 ACH – Stroke				Open Tab
Model 2 ACH – Urinary tract infection (UTI)				Open Tab
Model 2 ACH – Cardiac arrhythmia				Open Tab
Model 3 SNF – Major joint replacement of the lower extremity (MJRLE)				Open Tab
Model 4 – All episodes combined				Open Tab

a. Analysis of survey measures

All survey measures were collapsed into a binary indicator as discussed in the outcome definitions section. For all measures, we used logistic regression to calculate risk adjusted outcomes among BPCI and comparison respondents. We then estimated the difference in risk adjusted outcomes between the two groups. All regressions were weighted and risk adjusted as discussed below, and we estimated robust standard errors for all regressions.[40]

[37] The Non-surgical: cardiovascular stratum was an aggregation of the following clinical episodes: acute myocardial infarction, cardiac arrhythmia, atherosclerosis, chest pain, medical peripheral vascular disorders, syncope & collapse, and congestive heart failure.

[38] The Non-surgical: cardio and neurovascular (cardio/neurovascular) stratum was an aggregation of the following clinical episodes: acute myocardial infarction, cardiac arrhythmia, atherosclerosis, chest pain, medical peripheral vascular disorders, syncope & collapse, congestive heart failure, stroke and transient ischemia.

[39] The Non-surgical: respiratory group was an aggregation of the following clinical episodes: chronic obstructive pulmonary disease, bronchitis, asthma; other respiratory; and simple pneumonia and respiratory infections.

[40] We explored the possibility of clustering at the provider level. However, the median provider only contributes a few observations to the sample for nearly all strata, suggesting there is no need to cluster. Additional testing did not reveal any meaningful correlation between the providers and the variance of the outcomes.

b. Weighting

We applied sampling weights and nonresponse weights before analyzing results.[41] The sampling weight is the inverse of the selection probability within each sampling strata in each Wave. The nonresponse weight was calculated for all survey respondents (complete and partial responses) and reflects the inverse of the probability of response among eligible beneficiaries of the sample (with deceased respondents removed) within each of the 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 (e.g., 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 in our regression analyses (e.g., age).

c. Controlling for differences in patient mix

We controlled for important risk factors to ensure comparability, as much as possible, between the BPCI and comparison groups. We performed regression-based risk adjustment for all survey questions,[42] which included the factors listed in Exhibit 18.[43]

[41] For both BPCI and comparison respondents, the sampling weights sum to the population size of the BPCI sample.
[42] Perceptions of care experience were not risk adjusted in Wave 2 because all beneficiaries should receive the same high level of care in the hospital and afterward, regardless of their demographics or clinical risk factors. However to increase the robustness of our estimates we began risk adjusting on these factors beginning in Wave 3.
[43] Risk adjustment variables varied by wave. Result tables indicate which covariates were included in which regressions.

Exhibit 18: Predictive risk factors used to risk adjust survey outcomes

Domain	Variables
Service Mix[44]	• Anchor MS-DRG[45] • Lower body fracture (MJRLE and Model 4 episodes only) • Large vessel ischemic stroke (stroke episodes only) • Intracerebral hemorrhage (stroke episodes only)
Patient Demographics and Enrollment	• Age (under 65, 65-79, 80+) • Gender • Medicaid status • Patients' language was English (for question 23)
Prior health conditions	• HCC index: HCC indicators weighted by their relative weight in the CMS-HCC model • Baseline functional status (for functional assessment measures) • Functional status using three summary measures (for questions 9 through 31) [46]
Prior utilization measures	• Number of SNF and IPPS days in the 90 days prior to the anchor hospitalization • Whether patients were admitted to the anchor hospital from the community
Provider Characteristics	• Ownership status • Academic status (ACH only) • Provider size
Survey Dimensions	• Wave of Survey (open tabs only)

4. Analysis of variation in performance among BPCI participants

To identify what characterizes top and bottom performers under the BPCI initiative, we assigned BPCI EIs into mutually exclusive performance groups based on standardized realized NPRA from Q4 2013 to Q2 2015. We conducted a variety of descriptive analyses with the following aims: a) examine variation in NPRA across providers; b) distinguish the main characteristics of top and bottom BPCI performers and identify the means by which BPCI participants succeeded under the initiative; c) assess whether there were unintended consequences correlated with high performance; and d) generate hypotheses to support and complement the findings obtained from the DiD analysis.

[44] Additional variables for MJRLE, stroke, and Model 4 episodes control for clinical heterogeneity that is not accounted for by MS-DRGs, and which is easily identifiable from ICD-9 and ICD-10 codes.

[45] The exception was Model 4, which included all clinical episodes in the Model. For Model 4, we controlled for four aggregate episode types: (1) non-surgical, (2) surgical: cardiovascular, (3) surgical: orthopedic excluding spine, and (4) surgical: spinal. Additionally, because MS-DRGs incorporate information about MCC episodes versus non-MCC episodes, but aggregate episode types do not, we included an indicator variable for MCC in the Model 4 analyses.

[46] Three of the functional status questions have only three possible responses, two functional status questions have four possible responses, and two have five. For each of the outcomes with less than five possible responses, the best functional status was coded as 1, the middle status (or two statuses) was coded as 2, and the worst functional status was coded as 3. We created a variable summing the number of functional measures with 2, the number with 3, and also a binary indicator for "missing functional status." For the two measures with five possible responses we created binary indicators for "all of the time/most of the time" and created a control variable summing the number of indicators equal to 1, as well as a binary indicator for "missing activity status." For functional status variables with four possible responses, we considered alternative cutoffs for coding responses as 1, 2, or 3; however none of these alternative cutoffs altered the results in any meaningful way.

To conduct these analyses, we compared top and bottom performers along the following dimensions:[47]

- Outcomes during the baseline period (Q4 2011 through Q3 2012)
- Changes in patient mix before and after BPCI implementation
- Shifts in payment and utilization patterns before and after BPCI implementation
- Participation in gainsharing
- Model, hospital, and market characteristics

a. Assigning EIs to performance groups

We assigned EIs to performance groups based on their standardized NPRA from the first quarter they joined BPCI through Q2 2015. Standardized NPRA measures a provider's average NPRA as a percent of the provider's target price for a given Model and clinical episode.[48] We selected this metric, firstly, because it is based on realized NPRA, which allows us to measure the monetary gains received by providers; and, secondly, because it does not favor EIs that had the highest savings opportunities because they had the highest payments during the baseline period.

We calculated standardized NPRA as follows:

$$Standardized\ NPRA = \frac{(Average\ EI\ NPRA)}{(Weighted\ Average\ EI\ Target\ Price)}$$

where *Average EI NPRA* is the average NPRA per episode and it is defined as:

$$\frac{(Q_1 NPRA + Q_2 NPRA + \cdots Q_N NPRA)}{\sum_{i=1}^{N} EI\ Cases\ in\ Q_i}$$

and *Weighted Average EI Target Price* is the average target price, weighted by the number of episodes for each, and it is calculated as:

$$\frac{[(Q_1\ EI\ Target\ Price\ x\ Q_1\ EI\ Cases) + (Q_2\ EI\ Target\ Price\ x\ Q_2\ EI\ Cases) + \cdots\ (Q_N\ EI\ Target\ Price\ x\ Q_N\ EI\ Cases)]}{\sum_{i=1}^{N} EI\ Cases\ in\ Q_i}$$

We classified BPCI providers into three groups, within a clinical episode, based on where they fell in the distribution:

- Top Performers: percentiles 76-100
- Average Performers: percentiles 26-75
- Bottom Performers: percentiles 1-25

[47] When discussing top and bottom performers we are referring to BPCI EIs, regardless of whether or not they were Awardees. No comparison group providers were included in the analysis, as the main goal of this analysis is to understand variation in monetary gains, as measured by standardized NPRA, among BPCI participants.

[48] To account for the changes in target price and variation in the number of episodes each quarter, we estimated a weighted average target price, weighted by the number of episodes.

b. Selection of Model/EI type/clinical episodes

To ensure meaningful results, we only included combinations of Model/EI type/clinical episodes (strata) with a sufficient number of EIs for the analyses. Our inclusion criteria were as follows:

- We included strata that had at least 30 BPCI EIs with at least 50 episodes each during the baseline and intervention periods. We required that each EI had at least 50 episodes to ensure reliable estimates of different outcomes at the EI level. We required at least 30 EIs for each stratum so we could generate a reliable estimate of the average performance of the group.
- Strata that met the first criteria were also required to have at least 30 EIs and 50 episodes each during both baseline and intervention periods with a non-missing value for the total Medicare standardized allowed payment for inpatient stay plus 90 days post-discharge period (PDP) outcome.[49]

The following Model 2 ACH clinical episodes met our selection criteria: MJRLE, COPD, and CHF. No Model 3 or 4 clinical episodes met our selection criteria.

c. Data sources

In addition to NPRA data from the Net Payment Reconciliation Reports (Q4 2013 to Q2 2015), we used the following data sources for our analysis: Claims data (Q4 2011-Q3 2012 for baseline period and Q4 2013-Q2 2015 for intervention period), Awardee-submitted data (Q4 2013-Q4 2015), and Salesforce data (Q4 2013- Q2 2015).

d. Analysis

We conducted descriptive analyses, including frequencies and other univariate statistics, to examine distributions of our performance measure and our independent variables during the baseline and intervention periods. The summary statistics reflect how the performance of each group (i.e., top performers, average performers, bottom performers) correlates with a variety of characteristics of interest. Exhibit 19 lists the characteristics we analyzed.

[49] Although we tested more stringent selection criteria, such as requiring that all strata included in the analyses had a statistically significant decrease in total payments for the inpatient stay plus 90 days PDP relative to the comparison group (at the 10% significance level), we selected the above criteria as these other more stringent criteria resulted in the inclusion of only one clinical episode (MJRLE).

Exhibit 19: Characteristics Compared across Performance Groups

Dimension	Characteristics
Baseline quality, utilization, payments, and case-mix	• Baseline quality, utilization, and payment outcomes: ▪ Total Medicare standardized allowed payment for inpatient stay plus 90 days post-discharge period* ▪ 90-day unplanned readmission rate*[50] ▪ 90-day emergency department use rate* ▪ Institutional number of days* ▪ Percent discharged to Institutional PAC* ▪ Percent discharged to Home Health* ▪ Percent discharged to None (patient went home, received no PAC)* ▪ SNF number of days* • Baseline case-mix: ▪ Average HCC Score ▪ Average age ▪ Percent eligible for Medicaid ▪ Percent disabled (excluding ESRD) ▪ Percent of patients with fracture (exclusively for MJRLE)
Shifts (from baseline to intervention period) in case-mix and payments outside the bundle period	• Average HCC Score • Average age • Percent eligible for Medicaid • Percent disabled (excluding ESRD) • Percent of patients with fracture (exclusively for MJRLE) • Total payments during days 1-90 of the Post Bundle Period*
Shifts (from baseline to intervention period) in quality, utilization, and payment outcomes	• Total Medicare standardized allowed payment for inpatient stay plus 90 days post-discharge period* • 90-day unplanned readmission rate* • 90-day emergency department use rate* • Institutional number of days* • Percent discharged to Institutional PAC* • Percent discharged to Home Health* • Percent discharged to None (patient went home, received no PAC)* • SNF number of days*
Gainsharing Activities	Participation in the gainsharing waiver

[50] The risk adjustment methodology was analogous to the one used for the BPCI Awardee Feedback Reports.

Dimension	Characteristics
Model, hospital and market characteristics	• Model characteristics: ▪ Percent under Awardee Convener (AC)/Designated AC Agreement ▪ Percent EIs with a 30-day length of episode (LOE) ▪ Average number of selected clinical episodes • Hospital characteristics: ▪ Size (bed count) ▪ Type of Ownership: Percent non-profit and percent for-profit ▪ Resident to bed ratio ▪ Disproportionate Share ▪ Number of discharges within clinical episode, 2011 ▪ Percent with prior pay-for-performance or bundled payment experience • Market characteristics: ▪ Percent urban ▪ Number of SNF beds/10,000 residents ▪ Percent located in market with an IRF
Shifts (from baseline to intervention period) in volume (MJRLE only)	• Average number of discharges

*Risk adjusted outcome

We displayed results for selected characteristics in "heat maps". These reflect how, in a given clinical episode and with respect to a variable of interest, the top and bottom performance groups deviate from a "benchmark" (i.e., the average performance of all BPCI participants in the group). We did not summarize all characteristics with "heat maps", but instead focused only on those that presented meaningful correlations with our performance measure.

To assess whether there were meaningful correlations, we used Spearman's Rank Order Correlation. This nonparametric test measures the strength and direction of the association between two ranked variables, in this case, our performance measure and the characteristic of interest. Spearman's coefficient range from -1 to +1. Values closer to +/- 1 indicate a stronger positive/negative association between two variables; a value of 1, for example, means a perfect positive association of ranks, and values closer to zero signify a weak association between variables. We created "heat maps" for those characteristics that had statistically significant (p value <0.10) Spearman's Rank Order Correlation coefficients with our performance measure.

5. Market dynamics analysis

We conducted a descriptive analysis to understand whether BPCI EIs are capturing a greater share of the patients in their markets, in the episodes in which they are participating. We call these "BPCI eligible" admissions and hypothesize that BPCI EIs will strive to increase their market share, attracting patients that would otherwise go to competitors not participating in BPCI. If so, we would observe an increasing volume of BPCI-eligible admissions among EIs, relative to competitors providing the same types of services in a market. We examined changes in market shares for Model 2 ACH EIs and Model 3 SNF and HHA EIs.

We also wanted to understand whether BPCI Model 2 ACH EIs are changing the PAC providers to which they refer BPCI patients, preferring to send patients to the higher quality PAC providers

in their markets. We examined the concentration of patients discharged from Model 2 ACH EIs who subsequently received PAC care, and the star rating of the SNFs and HHAs to which they were admitted. We expect that Model 2 ACH EIs enacted strategies to gain market share, and to form and strengthen relationships with preferred PAC providers, over time. Likewise, we expect Model 3 PAC EIs to enact strategies to gain market share.

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

1. Market share of BPCI EIs, for BPCI-eligible admissions in which they participate
2. Number of PAC providers accounting for 75% of Model 2 ACH EI PAC discharges
3. Percent of patients discharged from Model 2 EIs to PAC providers with high CMS Star ratings

We used core-based statistical areas (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. Where CBSA boundaries are not a good definition of the competitive market, we will have limited ability to detect meaningful shifts in market share. To improve our ability to detect meaningful changes, we identified markets where the CBSA was too large to accurately define the local health care market and removed these CBSAs from the market share analyses.

All analyses focus on three BPCI clinical episodes: MJRLE, CHF, and sepsis. These clinical episodes are high-volume conditions in which most BPCI EIs are engaged, span both surgical (MJRLE) and non-surgical (CHF and sepsis) conditions, and include one chronic (CHF) and one acute (sepsis) medical condition. Market share and referral patterns may be considerably different for different clinical conditions. For example, the preferred PAC for an elective joint replacement rehabilitation patient may be different than the preferred PAC for a frail CHF patient. Or SNF EIs may have less interest in, or ability to attract, patients with sepsis than patients undergoing elective joint replacement.

We segmented the EIs according to the quarter in which they first joined BPCI and started participating in each of the three clinical episodes included in this analysis, and we analyzed each cohort separately. The sections below offer more details about our analytic approach.

a. Outcome definitions

Market share of BPCI EIs for BPCI-eligible admissions in which they participate

An EI's market share is defined as the number of BPCI-eligible admissions to the EI, divided by the total number of the same type of admissions across similar providers (i.e., ACHs, SNFs, HHAs) in the market. We calculated market share separately for BPCI Model and EI type: Model 2 ACHs, Model 3 SNFs and Model 3 HHAs. For Model 3 EIs, we include admissions within 30 days after a qualifying inpatient stay.

Number of PAC providers accounting for 75% of Model 2 ACH EI PAC discharges

This measure reflects the size of the PAC referral network used by Model 2 ACH EIs, and the number of PAC providers that account for 75% or more of an ACH's discharges – the concentration of their PAC discharges among a set of preferred PAC providers. We calculated the

share of each Model 2 ACH EI's patients discharged to each PAC provider in the market. We counted the fewest number of PAC providers having a cumulative share of 75% of the EI's discharged patients. This measure is calculated and analyzed separately for discharges to SNFs and to HHAs. To identify discharges from hospitals to PAC providers, only a patient's first SNF or HHA after hospital discharge was counted, excluding SNF admissions more than five days after hospital discharge and HHA admissions more than 14 days after hospital discharge.[51]

Percent of patients discharged from Model 2 EIs to PAC providers with high CMS star ratings

Model 2 EIs may have financial incentives to work with higher quality PAC providers, to reduce the total cost of episodes resulting from fewer complications and readmissions to hospitals. This measure examines whether a larger share of patients discharged from BPCI Model 2 ACH EIs are admitted to high-quality PACs, as defined by CMS Compare databases for nursing homes/SNFs and HHAs, after joining BPCI.

Nursing Home Compare and Home Health Compare are consumer-oriented websites that provide information on the quality of care provided in nursing homes (including SNFs) and HHAs. Both feature a five-star rating system of composite scores based on a variety of setting-specific quality metrics. Nursing Home Compare was first published in November of 2002 and Home Health Compare was first published in July 2015. Both update their five-star ratings quarterly.

We used publicly available, archived CMS Compare data to construct this measure. We defined high quality PAC providers differently for SNFs and HHAs because the two five-star rating systems have different measures and national distributions. Nationally, approximately one-quarter of SNFs attain a five-star rating, and one-quarter of HHAs achieve a four-star rating or higher; we therefore used these thresholds. For each Model 2 ACH EI that discharged patients to SNFs (within 5 days), we calculated the percent of all SNF discharges that were to SNFs having a five-star rating as of March 2014 (the rating system release closest to the start of BPCI participation). For patients discharged to HHAs (within 14 days), we calculated the percent of all such discharges that were HHAs having a four-, four and a half-, or five-star rating as of July 2015.

We anchored the PAC providers' star rating to a particular point in time and used this constant measure of a provider's relative quality across all pre-BPCI and BPCI periods; we did not update their star rating. This is to address two limitations of using the Nursing Home and Home Health Compare star ratings to measure providers' quality. First, Home Health Compare data was not published until Q3 2015, so we do not have a quality rating for HHAs prior to Q3 2015. Second, CMS periodically updates the set of setting-specific quality metrics included in, or the specific weights and thresholds used in, the methodology for calculating the composite star ratings. For instance, in February 2015 Nursing Home Compare underwent a "rebasing" of its thresholds for computing the five-star distribution. Thus, while the overall distribution of providers' star-ratings might shift in future periods, some part of these shifts do not reflect changes in providers' levels of quality, but rather are due to changes in the measurement systems.

Some new SNFs did not have star ratings as of March 2014 due to lack of historical data. For these SNFs, ratings were assigned based on the earliest archived data set through January 2015

[51] These five and fourteen day restrictions for SNF and HHA referrals are consistent with the definitions in the analyses of claims- and patient assessment-based outcomes for the BPCI evaluation.

with non-missing data. Patients admitted to SNFs that still did not have a star rating reported were excluded from both the denominator and numerator of the measure. Likewise, if HHAs did not have star ratings in July 2015, patients discharged to these HHAs were excluded from both the denominator and numerator of the measure.

The second and third measures above have an important limitation: They examine whether BPCI has influenced Model 2 ACH EIs' discharge referral patterns, in favor of a set of preferred SNF providers in the local area. However, we cannot tell whether the discharges from a hospital to a PAC provider reflect a deliberate referral decision made by the hospital, or reflect independent beneficiary preferences and decisions since beneficiaries are free to choose among all PAC providers. For example, some beneficiaries may arrange for their PAC stay prior to a surgical procedure based on recommendations from their physician and loved ones or promotional activities by the PAC providers.

b. Market definition and selection

A market includes the geographic area from which the BPCI provider draws patients and competes to provide services. To define BPCI markets, we used a geopolitical boundary defined by CBSAs. By Q3 2015, there were a total of 105 CBSAs that contained at least one BPCI Model 2 ACH EI. There were 114 CBSAs that contained at least one Model 3 SNF or HHA EI.

CBSAs are socially and economically interdependent regions that are geographically circumscribed by commuting times to the core geographic areas. While CBSA is an appropriate boundary for many markets, CBSA may not accurately define the local health care markets in large urban areas, such as New York City or Chicago. The very small market shares of EIs in these markets suggests that the CBSA is too large to define the local health care market, making it difficult to detect meaningful changes in market share over time.

We assessed the appropriateness of the CBSA boundary in two ways and removed five markets where CBSA was too large to accurately define a local health care market. First, we examined the market share of BPCI EIs when CBSA was used to define the market. Extremely small market shares of BPCI EIs in a market could be a sign that the CBSA is too large. Second, we calculated two market metrics: patient outflow rate and patient inflow rate, to identify markets where CBSA is too large a boundary. The outflow rate is defined as the percent of residents in a CBSA seeking any type of inpatient treatment at any hospital located outside the CBSA. The inflow rate is defined as the percent of discharges from all hospitals located inside a CBSA for patients who reside outside the CBSA. Unusually low inflow and outflow rates (i.e.,. under 10%) also serve as indicator that the market boundary may be too broad. For example, the largest possible boundaries (i.e., the entire U.S.) would have extremely low inflows and outflows. Through these two means, we identified five large markets where CBSA was a poor market definition and removed them from the subsequent analysis of market shares. These five CBSAs were as follows:

1. Chicago-Joliet-Naperville, IL-IN-WI (16980)
2. New York-Northern New Jersey-Long Island, NY-NJ-PA (35620)
3. Detroit-Warren-Livonia, MI (19820)
4. Cincinnati-Middletown, OH-KY-IN (17140)
5. Los Angeles-Long Beach-Anaheim, CA (31080)

Additionally, we excluded CBSAs where the BPCI EI had a 100% market share for a clinical episode type in the baseline period (Q4 2011 through Q3 2012). Markets with only one hospital, which is participating in BPCI, are unusual and unlike other BPCI markets because there is no competition: the usual market dynamics may not be influencing that EI's decisions about capturing market share.

CBSAs were only excluded from the analyses of the change in Model 2 and Model 3 EIs' market shares of BPCI-eligible admissions. The other two measures focus on discharges from Model 2 EIs to PAC providers without consideration of the PAC providers' specific geographic location. The removal of a relatively large number of EIs from the market share analyses due to their location in large CBSAs limits the generalizability of our findings to all BPCI EIs, particularly EIs located in highly competitive markets. However, removing these EIs is necessary to improve our ability to draw meaningful inferences from our results. Since the EIs located in the five large markets have small baseline market shares, due to the large size of their CBSAs, this limits our ability to detect meaningful change in their market shares over time. Including these EIs in the analyses of market share would suppress the average change in market shares for all EIs, thereby obfuscating the change observed among EIs for where the CBSA more accurately defines their local health care market.

The number of EIs removed from the analysis due to these two restrictions are discussed below with our analytic approach.

c. Data

We used 100% of Medicare Part A claims between Q4 2011 and Q3 2015 to calculate market shares for hospitals, SNFs, and HHAs located in CBSAs where one or more BPCI EIs were located. These claims spanned the pre-BPCI and BPCI periods. We linked data from Medicare claims to CBSA designations found in the POS Files. To identify PAC referrals we created dyads, or provider pairs, representing any discharging inpatient hospital and the first admitting PAC provider after a patient was discharged. This file contained one row for each patient transition from an acute care hospital to a SNF or HHA, starting with the first inpatient claim for a BPCI-eligible admission in the pre-BPCI period. Data on star ratings for SNFs and HHAs were acquired from Data.Medicare.gov, and linked to this file using CMS Certification Numbers (i.e., CCN) for SNFs and HHAs.

d. Analytical approach

As mentioned above, we focused on three BPCI clinical episodes: MJRLE, CHF, and sepsis.

We conducted pre-post comparisons of the three market-related measures for each of the three clinical episodes separately at the BPCI EI level, tracking changes in measure rates from the baseline to intervention period to infer the impact of the BPCI initiative. Specifically, we first calculated the changes in market share and the two measures of PAC concentration from the baseline period to the intervention period at the individual EI level. We then calculated the mean, standard deviation, minimum, maximum, and quartiles of the calculated changes in each measure across all EIs.

BPCI allows rolling enrollment and disenrollment of participants, and new EIs have joined the initiative in every quarter up until Q3 2015 since the initiative was first implemented in Q4 2013. Furthermore, existing EIs can elect to participate in specific clinical episodes in any given quarter. To avoid obfuscating the results of pre-post comparisons due to participants' differing lengths of exposure to BPCI, we stratified all analyses according to the quarter in which the BPCI EIs first enrolled to participate in the clinical episode being examined. However, stratifying EIs by the quarter in which they joined reduces the number of episodes in the analysis. To address this concern, we only analyzed the quarters with the largest influx of BPCI participant (i.e., Q1 2014, Q2 2015, and Q3 2015) since these large volume quarters offer the best opportunity to detect meaningful changes in our measures. However, those who joined in Q3 2015 have had only three months of BPCI participation by the end of our follow-up period, Q3 2015. Thus, we limited our analyses to those that joined in Q1 2014 and Q2 2015 among Model 2 ACH EIs and Model 3 SNF EIs across all three clinical episodes. Substantial influxes of BPCI participating HHAs occurred in Q1 2014, Q2 2014, Q2 2015 and Q3 2015, but only to participate in the CHF episode. Thus, we analyzed Model 3 HHA EIs that started the CHF episode in Q1 2014, Q2 2014 and Q2 2015. We define each sample as a "cohort".

We defined the baseline period as Q4 2011 through Q3 2012, which is one year prior to the beginning of Phase 1 of the BPCI initiative. We segmented each EI's baseline and post-BPCI periods into six month intervals, and then calculated the change in each measure from the baseline period to each six-month intervention period. For example, if a Model 2 ACH EI joined BPCI to participate in MJRLE in Q1 2014, we calculated the change in market share of MJRLE episodes for this EI between the baseline period and the Q1 2014-Q2 2014, Q3 2014-Q4 2014, and Q1 2015–Q2 2015 periods, respectively. If an EI did not join BPCI until Q2 2015, then we calculated its change in market share once, between the Q2 2015-Q3 2015 period and the baseline period.

EIs removed from the analysis

Exhibit 20 presents the number of BPCI EIs in the Q1 2014 and Q2 2015 cohorts that were removed from the analyses of EIs' market shares for BPCI-eligible admissions in which they participate, for each Model overall and for each clinical episode. The majority of Model 2 ACH EIs removed were located in the five large CBSAs. Out of a total of 78 Model 2 ACH EIs in the Q1 2014 cohort participating in MJRLE, CHF, or sepsis episodes, 23 were located in the five large CBSAs, and three more were removed because they had 100% of market share of the clinical episode type in the baseline period. Some EIs that were removed from the analysis started participating in more than one of the three clinical episodes that we examined. Out of 115 ACH EIs in the Q2 2015 cohort, 18 were located in the five CBSAs removed from the analysis, and seven more were removed due to having 100% of market share during the baseline period.

These restrictions were applied in the same manner to Model 3 SNF and HHA EIs, and these numbers are also reported in Exhibit 20. Out of 55 SNF EIs and 17 HHA EIs in the Q1 2014 cohort, there were 22 SNF EIs and one HHA EI located in the five large CBSAs. Two more HHA EIs in this cohort were removed because they were not located in a CBSA (not shown in the table). Out of 154 SNF EIs and seven HHA EIs in the Q2 2015 cohort, there were 18 SNF EIs located in the five large CBSAs, and one SNF was removed from the analysis of CHF episodes because it had 100% of market share during the baseline period. Four more SNF EIs

were removed from this cohort because they were not located in a CBSA (not shown in table). No Model 3 HHA EIs were removed from the Q2 2014 cohort.

Exhibit 20: Numbers of EIs removed from analysis due to restrictions

Clinical Episode Type	Q1 2014 starters				Q2 2015 starters			
	Total Number of EIs Before Removal	Removed due to CBSA Definition*	Removed due to 100% Market Share**	Total Number removed	Total Number of EIs Before Removal	Removed due to CBSA Definition*	Removed due to 100% Market Share**	Total Number removed
Total Number of Model 2 EIs	78	23	3	26	115	18	7	25
MJRLE	66	18	2	20	89	15	5	20
CHF	31	10	1	11	55	5	5	10
Sepsis	11	5	0	5	50	6	3	9
Total Number of Model 3 SNF EIs	55	22	0	22	154	18	0	18
MJRLE	12	0	0	0	102	18	0	18
CHF	54	22	0	22	69	7	1	8
Sepsis	45	22	0	22	54	8	0	8
Total Number of Model 3 HHA EIs	17	1	0	1	7	0	0	0
CHF	17	1	0	1	7	0	0	0

* EIs were removed from analysis due to being located in the five CBSAs where CBSA was too big to define the local market.

** EIs were removed from analysis due to having 100% of market share of clinical episode type during the baseline period.

6. 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 participant experiences in the BPCI initiative.

For site visits conducted from October 2015 through December 2015, interviewers developed a single set of notes for the two-day interview session. These notes were added to the ATLAS.ti (version 7.0.91; Scientific Software Development GmbH, Berlin, Germany) qualitative database and coded for qualitative themes. For site visits from January 2016 through June 2016, coders used the site visit narrative report as source data (rather than the notes document) to improve coding efficiency and standardization. As with the notes, the site visit reports were imported to the ATLAS.ti database and coded for major qualitative findings. Data from quarterly interviews were gathered by means of note-taking in a standardized template during each call. These standardized notes for each quarterly interview were 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. The first step of coding identified key themes and the second step highlighted recurring and sub-themes. Each data document (i.e., notes or site visit reports) underwent independent review by two researchers with subsequent discussion to establish a common understanding about major themes and alignment of these findings with established codes. We created definitions for each code and adopted common conventions for the coding process. Each team member who participated in qualitative coding received training in using ATLAS.ti and was familiar with the BPCI initiative through model documents, IPs, the evaluation and monitoring plan, and participated in data gathering during site visits or quarterly interviews.

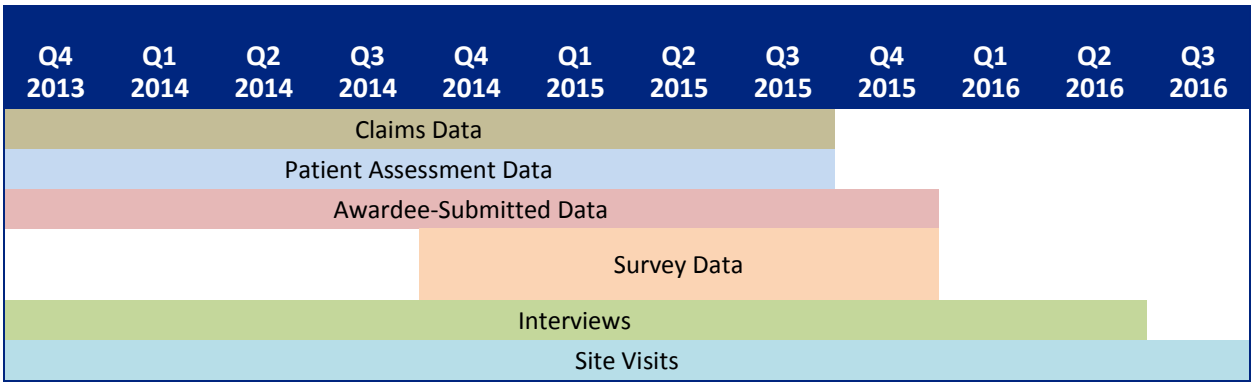
Analysis of the qualitative themes from case studies and interviews was guided by Research Questions A and C and the constant comparative method, a systematic data coding and analysis process during which specific quotes were categorized into themes with codes developed iteratively to reflect the data.[52] We developed the codes in steps, drafting a preliminary code structure after independent review by senior researchers.[53] The site visit code set from the most recent year of data collection was refined from the prior year's code set, with codes added to reflect new topic areas and new themes for this year. Codes were developed independently for each set of quarterly interviews. After coding the notes and reports, 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. We added new codes to capture new concepts as needed. Discrepancies in coding were discussed until consensus was reached regarding the most appropriate code.

[52] Glaser, B.G. & Strauss, A.L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New Brunswick, New Jersey: Aldine Transaction.
[53] Crabtree, B.F. & Miller, W.L. (1999). *Doing Qualitative Research*. Thousand Oaks, California: Sage Publications, Inc.

III. Model 2 Results and Discussion

This section presents information about the experience of Model 2 BPCI participants and their episodes of care. Exhibit 21 below shows the time period for each of the data sources used in this section. The quantitative outcomes are risk adjusted as described in Section II.D.2.

Exhibit 21: BPCI Quantitative and Qualitative Data Sources Used for Model 2 Results



Note: The risk-bearing phase (Phase 2) of BPCI began Q4 2013. Awardee-submitted data for gainsharing analysis covers Q4 2013 through Q1 2016.

A. Characteristics of the initiative and participants

This section summarizes the characteristics of the BPCI participants during the first two years of the initiative. Where relevant, we summarize the qualitative results from the last BPCI evaluation annual report in a call-out box in the beginning of a section. The narrative that follows the call-out box complements those results with additional insights gathered through the last year's qualitative data collection and analysis.

1. Key takeaways on Model 2 characteristics of the initiative and participants

- 215 Awardees that represented 422 acute care hospital (ACH) episode initiators (EIs) and 277 physician group practice (PGP) EIs joined the risk-bearing phase in Model 2 of BPCI. These participants initiated over 242,000 episodes during the first two years of the initiative.
- BPCI-participating hospital EIs were more likely to be larger, non-profit hospitals and set in urban locations compared to non-participating hospitals. BPCI-participating hospital EIs had higher standardized Part A payments for patients discharged with BPCI MS-DRGs in 2011 than non-participating hospitals (7% higher on average).
- BPCI-participating hospitals are located in larger and more competitive markets, with fewer skilled nursing facility (SNF) beds per 10,000 residents than markets with no Model 2 or 4 BPCI hospitals.
- Model 2 participants engaged in gainsharing and made use of program rule waivers to a limited extent. Between Q4 2013 and Q1 2016, 132 Awardees indicated their intent to enter into gainsharing agreements with partner providers. Over the same period, 24 Awardees distributed approximately $13.5 million in net payment reconciliation amounts (NPRA) and 11 Awardees distributed $3.1 million in internal cost savings (ICS). Among the Model 2 EIs under Awardees that indicated their intent to gainshare

in their IPs that were accepted by CMS, 30% used the three-day hospital stay waiver and 18% provided beneficiary incentives. Participants are not required to indicate their intent to use the telehealth or post-discharge home visit waivers and 5% or fewer EIs used them.

- Interviewees stated that obtaining initial physician buy-in to the BPCI initiative and having continued support and engagement is important for success in the initiative.
- Interviewees reported developing collaborative relationships with post-acute care (PAC) providers to manage patient care in the post-discharge period and reduce readmissions.
- Interviewees stated that relationships with Conveners or external contractors are important for supporting data management and analysis, as well as managing BPCI administratively.
- Interviewees reported setting patient expectations for discharge destination through early patient education, including consistent messaging from a multidisciplinary team, in an effort to change PAC use. Additionally, Awardees reported assessing patients for risks and available home supports to ensure safe transitions.
- Interviewees noted that sharing real-time data with PAC providers was often difficult because electronic health records (EHRs) differ between PAC providers and hospitals and willingness to share data across systems varies.
- Interviewees stated that it was difficult to influence PAC providers to change their care practices in less competitive markets or in facilities that were not owned by the EI.
- During the first two years of the BPCI initiative, 27 (6%) Model 2 hospital EIs and one Model 2 PGP EI withdrew from BPCI.

2. Participants

a. Entry decisions

Interview Findings Reported Previously

- Interviewees joined BPCI for the opportunity to learn and gain experience with bundled payment models or after being approached by a Convener
- Interviewees reported selecting Model 2 (vs. Model 4) due to the inclusion of the PAC stay and believed that this model offers the greatest opportunity to achieve savings.

By the third quarter of 2015, 215 Awardees that represented 422 ACH EIs and 277 PGP EIs were active in Model 2. Model 2 episodes accounted for over 85% of the more than 242,000 episodes initiated across Models 2, 3, and 4 during the first two years of the initiative.

Consistent with what we have reported previously, Model 2 interviewees indicated that they joined BPCI for the opportunity to learn. Interviewees sought opportunities to learn about population health management, best practices, and care improvement strategies. They also sought opportunities to learn about bundled payments, which they noted are the future of payment reform. Another reason interviewees cited for joining BPCI was the desire to be considered a leader in health care and payment reform. Interviewees noted that the experience gained from participating in BPCI would enhance their ability to negotiate with payers in the future.

Interviewees reported that the selection of Model 2 was frequently influenced by data analyses and recommendations from a Convener or outside consultants. Perceived financial opportunity was the most common rationale for selecting Model 2, particularly because of the opportunity to achieve savings by reducing PAC utilization. As in previous years, respondents reported selecting Model 2 because it provided an opportunity to learn about managing the entire continuum of care. Among PGP respondents, a common reason for selecting Model 2 over Model 3 was to have more control over the inpatient portion of the clinical episode.

b. Participant characteristics

During the first two years, a number of participants stopped participating in some or all of their BPCI clinical episodes. Sixty hospitals (14% of hospital EIs) and 39 PGPs (14% of PGP EIs) stopped participating in at least one clinical episode, and 27 hospitals (6%) and one PGP (0.4%) withdrew completely.

As of Q3 2015, the average length of participation in BPCI was three quarters for hospital EIs and one quarter for PGPs, in part because a majority of EIs (64% of hospitals and 70% of PGPs) joined BPCI in the last two quarters of enrollment (Q2 and Q3 2015).

Exhibits 22a & 22b compare characteristics of BPCI-participating hospital EIs with non-participating hospitals. Characteristics of Model 2 hospitals have changed during the initiative as additional hospitals enrolled in BPCI each quarter. In general, BPCI hospitals have become more similar to non-participating hospitals, although key differences remain between participants and non-participants.

Compared with non-participating hospitals, BPCI-participating hospital EIs were more likely to be non-profit (77% vs 57%), set in urban locations (92% vs 69%), and have a higher bed count (311 vs. 175). BPCI-participating hospitals had larger teaching programs, as indicated by a higher resident-to-bed ratio (0.12 vs. 0.05). They had nearly twice as many discharges for BPCI episode MS-DRGs during 2011 (3,004 vs. 1,598) compared with non-participating hospitals. For some metrics, participants and non-participants were similar. The disproportionate share percent was similar between BPCI participants (27%) and non-participants (29%), which indicates that both groups treated about the same proportion of Medicare Supplemental Security Income, Medicaid, or other low-income populations. Additionally, the proportion of BPCI-participating hospital EIs that were part of a chain was comparable to that of non-participating hospitals (52% vs. 53%) and they had a similar share of total inpatient days due to Medicare beneficiaries (39% vs. 42%).

One notable difference was that BPCI-participating hospital EIs had higher standardized Part A payments during the inpatient stay plus the 90-day post-discharge period (PDP) for patients discharged with BPCI MS-DRGs in 2011 than non-participating hospitals (Exhibit 22c).[1] Across all clinical episodes, average 2011 standardized Part A payments were $1,366 (7%) higher among BPCI-participating hospitals than among non-participating hospitals. The difference in standardized payments varied by clinical episode; the greatest difference in payment occurred in major bowel procedure episodes, where the mean payment was $3,477 (12%) higher in BPCI-participating hospital EIs than non-participating hospitals. The smallest difference was for major

[1] The results for BPCI hospital EIs shown in Exhibit 22c only include EIs participating in that episode as of Q3 2015.

joint replacement of the lower extremity (MJRLE), in which payments for BPCI-participating hospitals were $352 (1%) higher than for non-participating hospitals.

Exhibits 22a & 22b: Baseline Characteristics of BPCI-participating Hospital EIs and Non-participating Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospital EIs (N=419)		Non-participating Hospitals (N=2,774)	
		N	%	N	%
Ownership	For Profit	66	16%	638	23%
	Government	32	8%	542	20%
	Non-Profit	321	77%	1,594	57%
Urban/Rural	Rural	32	8%	872	31%
	Urban	387	92%	1,902	69%
Part of Chain	Yes	216	52%	1,469	53%

Characteristic	BPCI Hospital EIs (N=419)	Non-participating Hospitals (N=2,774)
	Mean	Mean
Bed Count	311	175
Number of Discharges for BPCI Episode MS-DRGs, 2011	3,004	1,598
Medicare Days Percent	39%	42%
Resident-to-bed ratio	0.12	0.05
Disproportionate Share Percent	27%	29%
Hospital Market Share	21%	27%

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals not participating in any BPCI initiative that reported values for all measures listed above and are not from Maryland. Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

Exhibit 22c: 2011 Standardized Allowed Part A Payments for Inpatient Stay Plus 90-day PDP, BPCI-participating Hospital EIs and Non-participating Hospitals, by Clinical Episode[1], Model 2, Q4 2013 - Q3 2015

Clinical Episode	BPCI Hospital EIs			Non-participating Hospitals		
	Number of Hospital EIs	Number of Discharges	Mean Payment Across Hospitals	Number of Hospitals	Number of Discharges	Mean Payment Across Hospitals
Major joint replacement of the lower extremity	289	58,890	$23,819	2,322	289,882	$23,467
Congestive heart failure	136	27,476	$19,645	2,590	302,335	$18,501
Simple pneumonia and respiratory infections	110	22,922	$19,128	2,616	361,079	$17,617
Chronic obstructive pulmonary disease, bronchitis, asthma	110	20,819	$15,330	2,594	327,638	$13,780
Sepsis	102	17,014	$28,714	2,519	257,621	$25,286
Urinary tract infection	64	8,174	$18,731	2,549	186,194	$17,069
Stroke	63	7,527	$26,500	2,396	150,164	$24,065
Renal failure	55	6,557	$20,671	2,430	164,471	$18,852
Medical non-infectious orthopedic	70	5,563	$23,413	2,393	114,691	$21,200
Cardiac arrhythmia	57	5,250	$13,430	2,497	184,508	$12,840
Acute myocardial infarction	76	5,074	$22,075	2,294	92,965	$21,638
Other respiratory	55	4,779	$27,701	2,437	149,623	$25,708
Esophagitis, gastroenteritis and other digestive disorders	43	4,485	$13,026	2,546	173,283	$11,795
Hip & femur procedures except major joint	72	4,387	$38,207	2,150	91,373	$36,586
Percutaneous coronary intervention	36	4,369	$19,331	1,227	124,125	$17,575
Cellulitis	55	4,280	$16,124	2,515	108,646	$14,404
Gastrointestinal hemorrhage	42	4,054	$15,865	2,457	145,446	$15,096
Nutritional and metabolic disorders	34	2,721	$17,188	2,544	122,596	$15,192
Coronary artery bypass graft	38	2,306	$39,003	832	39,030	$36,975
Cardiac valve	22	1,922	$52,588	736	29,918	$50,648
Spinal fusion (non-cervical)	26	1,633	$31,983	1,221	45,184	$30,299
Major bowel procedure	27	1,632	$31,926	2,132	76,020	$28,449
Revision of the hip or knee	28	1,100	$32,690	1,240	25,279	$29,812

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. Standardized allowed payments are based on 2011 claims and trended to 2015 dollars. BPCI participating hospitals for a given clinical episode are defined as Model 2 EIs, Q4 2013 – Q3 2015 who had at least 5 discharges in 2011 in the episode of relevance. Non-participating hospitals are all other hospitals not participating in any BPCI initiative that reported values for all measures listed in Exhibits 22a & 22b, had at least 5 discharges in 2011 in the episode of relevance, and are not in Maryland. Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

[1] The clinical episodes included in this table are those that had enough sample size to include in the risk adjusted difference-in-differences analysis presented in the Results section, Model 2 Impact of BPCI.

c. Participant Readiness

The majority of BPCI-participating hospitals reported prior experience with care redesign and payment incentives and had systems in place that prepared them for the BPCI initiative. BPCI-participating PGPs, on the other hand, were less likely to have had experience with care redesign and payment incentives prior to joining BPCI.

Prior Experience

Awardee-submitted data indicated that the redesign of care pathways was the most widely reported redesign experience among EIs prior to BPCI (see Exhibit 23). Prior experience related to the redesign of care pathways (59.1%), care coordination (58.4%), and enhancements in care delivery (57.9%) were reported most often among Model 2 ACH EIs. Prior experience with care redesign initiatives was less common among Model 2 PGPs.

Over half (54.4%) of Model 2 ACH EIs had prior experience with pay for performance. Model 2 PGP EIs reported little experience with payment incentives prior to their participation in BPCI.

Exhibit 23: Care Redesign and Payment Incentives Experience, Model 2, Prior to BPCI Participation*

		Model 2 Hospital EIs (N=401)		Model 2 PGP EIs (N=281)	
		N	%	N	%
Prior experience in care redesign initiatives	Redesign of Care Pathways	237	59.1%	85	30.2%
	Enhancements in Care Delivery	232	57.9%	18	6.4%
	Patient Activation, Engagement, & Risk Management	206	51.4%	18	6.4%
	Care Coordination	234	58.4%	19	6.8%
	System Changes to Support Care	198	49.4%	19	6.8%
	Other Redesign Activities	27	6.7%	9	3.2%
Prior experience in payment incentives	Bundled Payments	55	13.7%	10	3.6%
	Pay for Performance	218	54.4%	20	7.1%
	Shared Savings	167	41.6%	15	5.3%
	Other Payment Incentives	41	10.2%	18	6.4%

Source: Lewin analysis of Awardee-submitted data collected February 2015 through February 2016 for Model 2 episode initiators participating in BPCI between Q4 2013-Q4 2015.

*Prior experience in care redesign and payment incentive initiatives was reported when episode initiators first became active in Phase 2 of the initiative. Therefore, episode initiators that changed models are only included in the sample for the first model they participated in. This table only includes information about episode initiators that submitted data. Episode initiators that did not submit complete data are excluded from the counts in the table.

Electronic Health Record (EHR) Use and other Information Exchange

As shown in Exhibit 24, almost all (98.5%) Model 2 ACH EIs use an EHR. Of the 399 ACH EIs in Model 2 that submitted data, 393 reported using an EHR system. EIs that reported using an EHR system were also asked to report the meaningful-use functionalities of their systems. Computerized physician order entry (99.2%) and discharge instructions and care summary documents (99.2%) were the most common functionalities reported by Model 2 EIs. Medication

management and clinical decision support functionalities were also reported by 95.7% and 90.8% of Model 2 EIs, respectively.

Exhibit 24: Electronic Health Record (EHR) Use, Model 2 ACH, Baseline*

		Model 2 Hospital EIs (N=399)	
		N	%
	EIs with an EHR	393	98.5%
Meaningful-use functionalities (among EIs with EHRs)	Automated Quality Reporting	306	77.9%
	Discharge Instructions and Care Summary Documents	390	99.2%
	Medication Management	376	95.7%
	e-Prescribing	288	73.3%
	Computerized Physician Order Entry	390	99.2%
	Clinical Decision Support	357	90.8%

Source: Lewin analysis of Awardee-submitted data collected February 2015 and February 2016 for Model 2 ACH episode initiators participating in BPCI between Q4 2013-Q4 2015

*EHR measures were analyzed using data from the first reporting period episode initiators were required to report annual measures. This table only includes information about episode initiators that submitted data. Episode initiators that did not submit complete data are excluded from the counts in the table.

d. Episode and length selection

The count of EIs participating in each of the 48 clinical episodes during the first eight quarters of the initiative is shown in Exhibit 25. The average Model 2 EI participated in six clinical episodes, and one Model 2 EI participated in all 48 clinical episodes. The most popular clinical episode among Model 2 participants was MJRLE, in which 60% of EIs participated. Congestive heart failure was the second most common clinical episode and was chosen by 27% of EIs. During the first eight quarters of the initiative, EI participation was less than 10% in 28 clinical episodes. The vast majority of episodes chosen by Model 2 ACHs and Model 2 PGPs were 90 days in length (97% and 99%, respectively).

Compared to hospital EIs associated with an awardee convener or a designated awardee convener (N=261), single awardee and designated awardee hospital EIs (N=161) chose similar episode lengths (both chose 90-day episodes over 90% of the time), participated in fewer clinical episodes on average (2.9 vs 6.5), and were more likely to participate in only one clinical episode (59% vs 37%).

Exhibit 25: Participation of Episode Initiators by Clinical Episode, Model 2, Q4 2013 - Q3 2015

Clinical Episode	Episode Initiators by Participant Type (N=699)		
	Hospital (N=422)	PGP (N=277)	% of EIs
Major joint replacement of the lower extremity	294	126	60%
Congestive heart failure	136	56	27%
Chronic obstructive pulmonary disease, bronchitis, asthma	110	76	27%
Simple pneumonia and respiratory infections	111	67	25%
Sepsis	103	67	24%
Hip & femur procedures except major joint	74	70	21%
Acute myocardial infarction	78	53	19%
Urinary tract infection	65	60	18%
Medical non-infectious orthopedic	70	54	18%
Cellulitis	56	58	16%
Renal failure	55	57	16%
Cardiac arrhythmia	58	47	15%
Stroke	63	40	15%
Other respiratory	56	44	14%
Esophagitis, gastroenteritis and other digestive disorders	44	54	14%
Nutritional and metabolic disorders	35	55	13%
Gastrointestinal hemorrhage	43	45	13%
Fractures of the femur and hip or pelvis	37	49	12%
Percutaneous coronary intervention	37	45	12%
Red blood cell disorders	26	48	11%
Diabetes	27	37	9%
Gastrointestinal obstruction	26	36	9%
Transient ischemia	26	36	9%
Lower extremity and humerus procedure except hip, foot, femur	33	28	9%
Syncope & collapse	16	44	9%
Chest pain	21	38	8%
Spinal fusion (non-cervical)	29	28	8%
Medical peripheral vascular disorders	22	35	8%
Major bowel procedure	28	26	8%
Amputation	20	33	8%
AICD generator or lead	5	48	8%
Major joint replacement of the upper extremity	26	26	7%
Coronary artery bypass graft	41	10	7%
Revision of the hip or knee	36	15	7%
Other vascular surgery	30	20	7%

Clinical Episode	Episode Initiators by Participant Type (N=699)		
	Hospital (N=422)	PGP (N=277)	% of EIs
Atherosclerosis	18	32	7%
Double joint replacement of the lower extremity	29	19	7%
Pacemaker	19	24	6%
Cardiac valve	28	13	6%
Cervical spinal fusion	24	16	6%
Combined anterior posterior spinal fusion	15	17	5%
Cardiac defibrillator	12	18	4%
Removal of orthopedic devices	17	11	4%
Major cardiovascular procedure	16	12	4%
Pacemaker device replacement or revision	10	15	4%
Back & neck except spinal fusion	15	9	3%
Complex non-cervical spinal fusion	11	12	3%
Other knee procedures	6	17	3%

Source: Lewin Analysis of CMS' BPCI database, June 2016.

Note: The sum of the total EIs participating in each of the clinical episodes exceeds the total number of EIs because EIs can participate in more than one clinical episode.

Interview Findings Reported Previously

- Interviewees selected episodes based upon existing expertise, physician engagement, learning opportunities, savings opportunities, and patient volume.
- Interviewees that chose the 30-day episode length, did so because it was perceived as less risky (i.e. less chance of unrelated complications or readmissions) than the 90-day episode length.
- Interviewees selected the 90-day episode length for the following reasons: 1) desire to control a longer continuum of care, 2) opportunities for cost reduction in post-acute care, and 3) a lower discount rate from CMS.

Similar to previous years, respondents identified patient volumes, financial expectations, physician engagement, and alignment with existing organizational strengths as reasons for selecting clinical episodes. Interviewees reiterated previous findings related to selecting the length of episode.

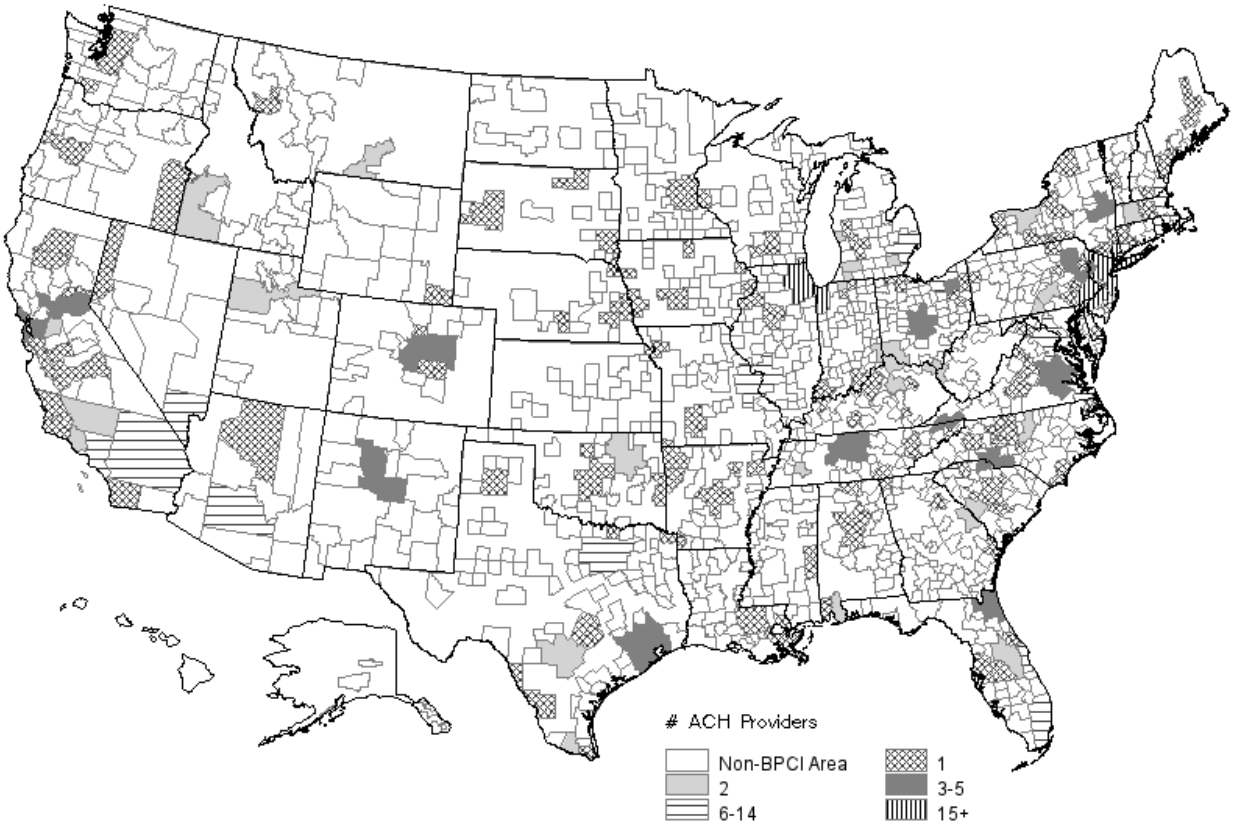
3. Market characteristics

The 419 Model 2 hospitals, and 277 Model 2 PGPs, and 23 Model 4 hospitals were located throughout the country, as pictured in Exhibits 26 and 27. In this section, we present the market characteristics[2] of the hospitals that participated in Models 2 and 4 (BPCI markets) and the

[2] 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).

characteristics of markets with no BPCI-participating hospitals (non-BPCI markets) through Q3 2015.[3]

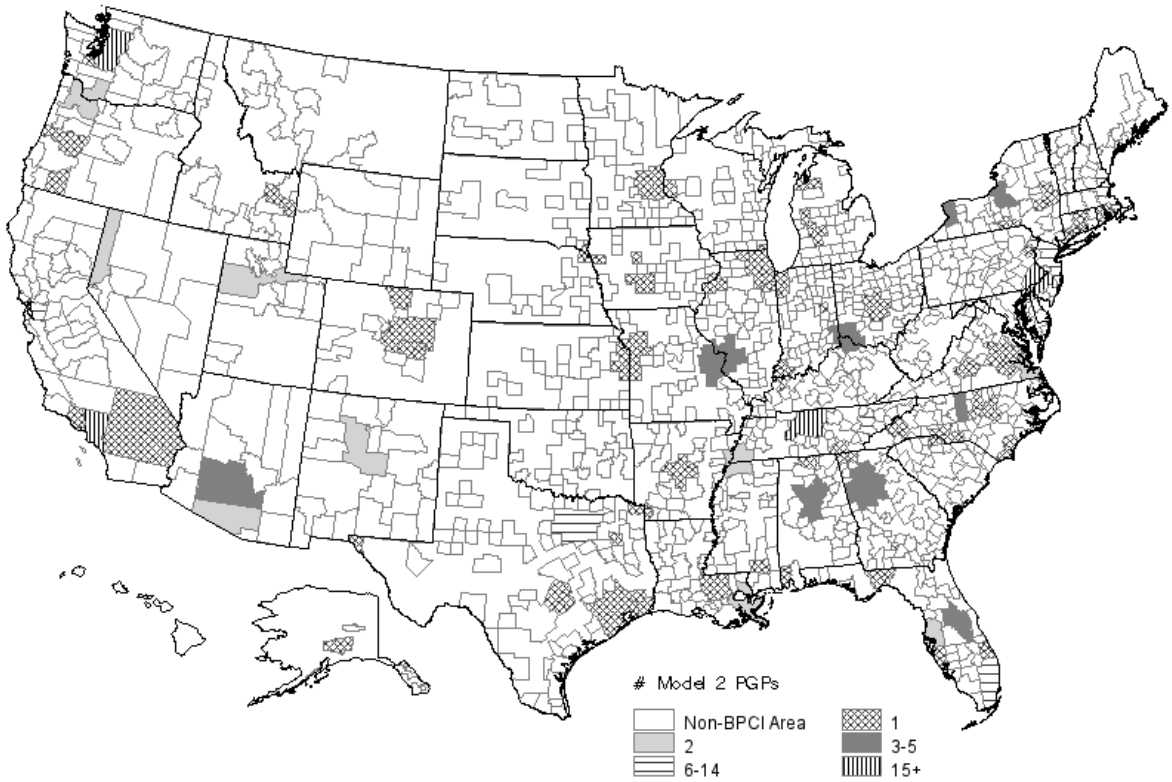
Exhibit 26: Number of BPCI-Participating Hospitals by CBSA, Model 2 and Model 4, Q4 2013 – Q3 2015



Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating hospital EIs.

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

Exhibit 27: Number of BPCI-Participating PGPs by CBSA, Model 2, Q4 2013 – Q3 2015



Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating PGP EIs.

Markets with BPCI-participating hospitals differed from markets without BPCI hospitals. As shown in Exhibit 28, BPCI markets tended to have larger populations (average 1.1 million residents), whereas non-BPCI markets were smaller (average 126,000 residents). On average, median household income was higher in BPCI markets ($49,119) than in non-BPCI markets ($43,095). BPCI markets also had, on average, higher Medicare Advantage (MA) penetration than non-BPCI markets (23.7% vs. 17.0%).

Consistent with location in larger areas, there was a higher concentration of primary care physicians (PCPs) per 10,000 residents in the BPCI markets (7.6 vs. 6.2), as well as a higher concentration of physician assistants/nurse practitioners (PA/NPs) per 10,000 residents (7.4 vs. 5.8). The difference between BPCI and non-BPCI markets was more pronounced with respect to specialists per 10,000 residents (9.3 vs. 4.7). BPCI markets had fewer SNF beds (56.4 vs. 73.8) per 10,000 residents than non-BPCI markets. The proportion of Medicare-aged residents was similar for BPCI and non-BPCI markets.

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 indicate lower competition and greater market power for local providers. The relatively low mean Herfindahl index for hospitals in BPCI markets (0.40) suggests competition among hospitals, with none of them dominating the market. In contrast, the mean Herfindahl index in non-BPCI markets was higher (0.73), suggesting that these markets were relatively less competitive and probably dominated by fewer hospitals with more market share.

Exhibit 28: Characteristics of BPCI Markets and Non-BPCI Markets, Model 2 and Model 4 Hospitals, Q4 2013 – Q3 2015

Market Characteristics	BPCI Markets N=174; 18.5% of Markets				Non-BPCI Markets N=768; 81.5% of Markets			
	Mean	Median	25th	75th	Mean	Median	25th	75th
Hospital Penetration	52.6%	47.0%	28.0%	73.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index – hospital	0.40	0.32	0.17	0.54	0.73	1.00	0.50	1.00
Herfindahl Index – SNF	0.12	0.07	0.03	0.17	0.36	0.30	0.18	0.45
Herfindahl Index – HHA	0.29	0.21	0.09	0.39	0.57	0.56	0.28	1.00
Herfindahl Index – IRF	0.33	0.00	0.00	1.00	0.08	0.00	0.00	0.00
Medicare Advantage Penetration	23.7%	22.3%	13.8%	33.7%	17.0%	14.2%	7.8%	22.9%
Population	1,123,402	428,242	162,697	925,899	125,900	59,418	37,068	122,253
Median Household Income	$49,119	$47,086	$42,800	$53,839	$43,095	$42,023	$37,894	$47,589
% Age 65+	14%	13%	12%	15%	15%	15%	13%	17%
PCPs Per 10,000	7.6	7.5	6.3	8.5	6.2	5.9	4.6	7.4
Specialists Per 10,000	9.3	8.1	5.7	11.2	4.7	3.9	2.3	5.9
PA/NPs Per 10,000	7.4	6.7	5.0	9.0	5.8	5.4	3.6	7.3
SNF Beds Per 10,000	56.4	51.5	37.1	70.2	73.8	68.0	45.3	93.9
LTCH Beds Per 10,000	0.9	0.5	0.0	1.5	0.4	0.0	0.0	0.0
IRF Beds Per 10,000	0.7	0.0	0.0	1.0	0.3	0.0	0.0	0.0
CAH Beds Per 10,000	0.6	0.0	0.0	0.4	2.0	0.0	0.0	1.9

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

Notes: BPCI markets are CBSAs that contain one or more Model 2 and 4 BPCI-participating hospitals. Providers not located in a CBSA were assigned to the largest CBSA within their Hospital Referral Region (HRR). Non-BPCI markets are CBSAs that do not have a Model 2 or 4 BPCI hospital. Areas of the country that are not in a CBSA are therefore not included in these non-BPCI markets. 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. HHA = home health agency. Variable definitions are in **Appendix L**.

4. Model incentive structure characteristics

a. Conveners in BPCI

Interview Findings Reported Previously

- Conveners reported selecting the facilitator convener (FC) role as opposed to the Awardee role because they wanted to gain expertise with alternative payment models but did not want to assume financial risk.
- FCs provided a variety of services for their EIs, including general education about BPCI, program administration, and data analytics. FCs also made recommendations about episode and waiver selection and served as intermediaries between their Awardees/EIs and CMS.
- AC and DAC roles varied; some were actively engaged in the administration and implementation of BPCI (e.g., helping to implement care redesign, assisting with episode selection), while others adopted a more passive approach by primarily providing resources and guidance (e.g., facilitating the exchange of best practices among BPCI-participating providers).

Interviewees indicated they joined under a Convener because this structure offered certain advantages. As noted in prior evaluation reports, hospital interviewees reported that their FCs provided a range of services, including data analysis (such as packaging CMS data and reviewing reconciliation data) and facilitating the sharing of information and best practices across BPCI participants. In addition to providing these services, interviewees noted that ACs assisted with care redesign activities, such as developing and instituting care protocols, assisting with case management and patient tracking, and facilitating relationships with PAC partners. Other benefits of joining under a Convener reported by interviewees included having administrative support for the initiative and sharing risk.

b. Partners

Interview Findings Reported Previously

- Interviewees mentioned forming partnerships and improving coordination with independent entities, such as physicians, physician group practices, and non-affiliated PAC providers. Interviewees also noted that partnerships enabled them to focus on population health and tracking patient outcomes.
- According to interviewees, success of the acute care/PAC partnerships depended 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 interviewees used gainsharing, while others offered higher patient referral volumes.
- Interviewees hired consultants to perform data analysis and administrative functions. They noted that these contractors played significant roles in managing data systems, performing internal data analyses, facilitating data sharing across other partners, and tracking patients.
- Interviewees described physician partnerships that facilitated information sharing and care standardization. Positive relationships and increased engagement with physicians were critical to implementing care redesign.

BPCI participants partnered with multiple types of organizations to support the initiative. We conducted quarterly interviews about partnering decisions and asked about these decisions in the Model 2 site visits. Interviewees noted collaborating with a variety of partners, including PAC providers, external contractors, and local community organizations.

Representatives of hospitals and PGPs described several characteristics that were considered when selecting preferred PAC providers as partners: star ratings or other quality metrics (i.e., readmission rate and length of stay), geographic location, patient volume, and their willingness to collaborate on care redesign. Similar to findings from previous years, interviewees indicated that successful partnerships were dependent on PAC providers' willingness to collaborate, communicate, and change practice patterns.

Quarterly interviewees and site visit interviewees described differing approaches to guiding beneficiaries on which PAC provider to select. Many, but not all, interviewees reported providing beneficiaries with a preferred provider list. Others focused on educating beneficiaries about PAC options, but did not try to influence provider selection. Interviewees anecdotally described better outcomes for beneficiaries who selected preferred providers due to enhanced care coordination. Conversely, interviewees mentioned difficulty in managing beneficiaries who chose non-preferred PAC providers, who were less responsive to improving communication or changing protocols.

Interviewees also described third-party entities as key partners. Partnerships with external contractors and consultants fulfilled needs that included data analytics, care redesign, and program administration. As in previous years, interviewees noted the significant role these partners played in managing claims data, generating quality outcome reports for internal use, and sharing data with partners. PGP interviewees also noted that partnerships with third-parties were beneficial to the development of care protocols, coordinating and managing care, identifying BPCI patients, tracking patient volume, and facilitating relationships with PAC partners. In addition to providing data analytics, interviewees described administrative roles external partners played. Administrative consultants helped interviewees with project management, payment reconciliation, implementation protocols, and CMS relations. Most PGP interviewees and some hospital interviewees noted that their Conveners helped with data analytics and administrative support.

> "You need to get the data analytic capacity somehow. I don't think we'll ever bring it in house."
>
> – Model 2 Hospital

Arrangements with hospitals were highly variable among PGP interviewees. Most interviewees reported having informal agreements with the hospitals at which their BPCI episodes were initiated. While many interviewees mentioned having good relationships with hospital staff, they also reported that more formal arrangements, including gainsharing agreements, often failed due to a lack of interest from hospitals. Conversely, other interviewees described success in obtaining formal gainsharing or management agreements with hospital partners.

Several interviewees described collaborative partnerships with organizations in their communities. One interviewee suggested that BPCI participants should examine how community organizations can enhance services provided to patients to improve outcomes. For example, an interviewee noted a partnership with a local fire department that would make home visits to beneficiaries with certain diagnoses within a specified geographical area. The hour-long home visits would occur weekly for five weeks and include discussions about medications and wound care. Other interviewees discussed partnerships with palliative care providers, physician groups in their communities, and organizations such as Meals on Wheels.

c. Waiver use

The BPCI initiative allows participants to use waivers from certain Medicare program requirements to facilitate the implementation of care redesign interventions. The waiver from the three-day inpatient hospital stay requirement for SNF coverage can be used if the Awardee describes its use in the Implementation Protocol (IP) and the IP is accepted by CMS. An EI may elect to use or not use a waiver indicated by its Awardee in their IP. All participants have access to waivers regarding telehealth services and home visit services. Most Model 2 EIs (62%) were eligible to use the three-day hospital stay waiver, but less than a third of those eligible used the waiver (Exhibit 29). Very few EIs used the telehealth (5%) or post-discharge home visit (0%) waivers.

In order for an Awardee to provide beneficiary incentives or to engage in gainsharing agreements under BPCI that may be protected under the BPCI fraud and abuse law waivers, the Awardee must include its plans for those activities in the IP. An EI may or may not elect to provide beneficiary incentives or enter into gainsharing agreements indicated by its Awardee in their IP. Nearly half of Model 2 EIs could provide beneficiary incentives, but less than a third of those eligible to provide incentives actually did so (Exhibit 29).

Exhibit 29: Participation of Episode Initiators in Various Waivers, Model 2, Q4 2013 - Q3 2015*

	Model 2 EIs (N=627)			
	Model 2 EIs Allowed to Use Waiver		Model 2 EIs that Used Waiver of those Allowed to Use Waiver	
Model 2 Waivers	N	%	N	%
Three-day hospital stay waiver	394	63%	117	30%
Beneficiary Incentives	309	49%	56	18%
Telehealth**	NA	NA	32	5%
Post-Discharge Home Visit**	NA	NA	2	0.3%

*The number of Episode Initiators includes only those who had at least one BPCI episode between Q4 2013 – Q3 2015. Data for gainsharing waiver are available at the Awardee level and presented in Exhibit 31.

**The Post-Discharge Home Visit and Telehealth waivers are available to all Model 2 EIs without specifying it in their Awardee's Implementation Protocol. Therefore, the denominator used to calculate % of M2 EIs that used these two waivers is the number of Model 2 EIs through Q3 2015.

Sources: Lewin analysis of Awardee Implementation Protocols for Q3 2015 BPCI participants, Medicare claims data for episodes initiated Q4 2013 – Q3 2015, and Awardee-submitted data Q4 2013 – Q3 2015.

Note: The 627 EIs that participated in Q4 2013 - Q3 2015 are distributed among 214 Model 2 Awardees. EIs include both hospitals and PGPs.

Three-day hospital stay waiver

Interview Findings Reported Previously

- Use of the three-day hospital stay waiver was low in part due to concerns that discharging a beneficiary to a SNF after less than a three-day hospital stay could result in financial liability for the SNF and beneficiary if used incorrectly.

Medicare rules require that beneficiaries have an inpatient hospital stay of at least three days to qualify for Medicare coverage of SNF care. The three-day hospital stay waiver allows Medicare beneficiaries treated by a Model 2 EI to qualify for Medicare SNF coverage (as long as other

criteria are met) even if their hospital stay was less than three days. Although 394 EIs could use this waiver, Medicare claims data indicates that only 117 EIs used it through Q3 2015 for 1,846 episodes (1% of all Model 2 episodes).

Similar to results from earlier reports, most interviewees stated that they have not used the waiver or that they used it infrequently, oftentimes due to their inability in real time to determine which patients were eligible for BPCI. These interviewees were concerned about the lack of Medicare coverage of the SNF stay if the patient was later found to not be in a BPCI episode. For example, interviewees described cases where the 'working' MS-DRG, included within a BPCI clinical episode, was different from the discharge MS-DRG, which was not included in BPCI. As a result, Medicare would not cover the SNF stay after a short hospitalization. In addition, Model 2 hospitals were more likely to discharge patients home rather than to a SNF. This reason was especially apparent among interviewees participating in orthopedic surgery episodes. Finally, a handful of participants believed that it was not clinically appropriate to discharge some patients (e.g., those with co-morbidities or non-surgical episodes) in less than three days, thus limiting the use of the waiver.

Interviewees identified two reasons for requesting the three-day hospital stay waiver: reducing the inpatient length of stay and hastening the patient's recovery. Interviewees also reported that they wanted to reserve the option to use the waiver, even if they did not initially have plans for using it. Most interviewees favored a Medicare policy change that would allow Medicare SNF coverage following a hospitalization of less than three days as a means to reduce costs.

Waiving the three-day hospital stay requirement for Medicare SNF coverage may contribute to costs being shifted from Medicaid to Medicare for Full Benefit Medicare-Medicaid Enrollees (FBMMEs) who are residents of a nursing facility. FBMME nursing facility residents who receive care in an acute care hospital may trigger Medicare SNF coverage when they return to the nursing facility. By reducing the hospital length of stay requirement, more FBMMEs could switch from Medicaid nursing facility to Medicare SNF coverage, which would typically result in higher payments to the nursing facility because Medicare SNF payments are often higher than Medicaid rates. We explored the extent to which this may occur under BPCI. **Appendix M** summarizes a descriptive analyses on the use of the three-day hospital stay waiver among this subpopulation.

Beneficiary incentives

The waiver of certain fraud and abuse laws allows EIs to provide beneficiaries with an in-kind, episode-related item or service that would not typically be allowed under existing law and regulation. As displayed in Exhibit 29 above, through Q3 2015, 309 EIs (49%) were allowed to provide incentives because their Awardee included plans for the beneficiary incentives in their IP. We analyzed Awardee-submitted data detailing beneficiary incentives that were distributed between Q4 2013 and Q3 2015. Fifty-six Model 2 EIs reported that they distributed 3,170 incentives to BPCI beneficiaries during this period. One or more beneficiary incentives were distributed in a total of 2,724 episodes (1% of the total episodes initiated under Model 2 EIs). These incentives ranged in value from $2.00 to $5,648 per item or service during Q4 2013 to Q3 2015. A detailed description of the distribution of beneficiary incentives by category provided through Q3 2015 is shown in Exhibit 30. Medication management tools continued to be the most common beneficiary incentive distributed, followed by equipment.

Exhibit 30: Beneficiary Incentives Distributed by EIs to Beneficiaries, Model 2, Q4 2013 - Q3 2015

Incentive Description	Awardees allowed to use incentives	EIs allowed to use incentives and had at least one BPCI episode	Awardees that provided one or more incentives	Episodes receiving one or more incentives	Average cost per incentive provided
Transportation	18	52	7	614	$71.77
Equipment	16	17	7	757	$31.82
Home care/home visits	16	22	5	211	$387.90
Living arrangement services	7	8	4	30	$535.00
Telehealth/technology	14	63	3	23	$85.60
Wellness program/resources	12	237	1	1	$30.00
Medication management tools	11	13	2	1,157	$11.70

Source: Lewin Program Adherence Report based on analysis of Awardee-submitted data regarding disseminated beneficiary incentives through Q3 2015.

Interview Findings Reported Previously

- Interviewees noted that the ability to provide beneficiary incentives could be used to address issues in care delivery (e.g., meals for patients who could not otherwise cook, animal boarding for patients recovering from surgery whose pets presented a fall risk).
- Some interviewees found providing beneficiary incentives to be administratively challenging because they are required to report each incentive and beneficiary name to CMS.

Interviews provided specific examples of beneficiary incentives, such as fitness center programs and general support at home. While quantitative data indicate that medication management tools was the most frequently provided beneficiary incentive, Model 2 interviewees most commonly described providing incentives to help beneficiaries with transportation to outpatient therapy or rehabilitation. Most interviewees were aware that beneficiary incentives were permitted and protected under the fraud and abuse law waivers and asked that they be maintained as an option, despite low utilization.

Gainsharing

Interview Findings Reported Previously

- Interviewees viewed gainsharing as a tool to change practice patterns by incentivizing physicians and other partners to buy in to the care redesign effort. Although some interviewees cited increased provider engagement as a reason for participating in gainsharing, other interviewees expressed doubt as to whether gainsharing actually increased provider engagement.

Of the 112 Awardees in BPCI from Q4 2013 through Q2 2015, most (66%) entered into gainsharing agreements. Twenty-four Awardees distributed approximately $13.5 million in NPRA and 11 Awardees distributed approximately $3.1 million in ICS between Q4 2013 through Q1 2016. Distribution of NPRA is dependent on several factors, such as whether they had positive NPRA, the specifics of the organization's gainsharing methodology, frequency of distribution, delayed reconciliation, and dispute of reconciliation results.

Exhibit 31 describes gainsharing distributions between Q4 2013 and Q1 2016 received by type of gainsharing partner. Among gainsharing partners, physicians were the most common to receive a distribution. Hospitals and PGPs were also common gainsharing partners.

Exhibit 31: Gainsharing Distributions Received by Partner Type, Model 2, Q4 2013 - Q1 2016

Gainsharing Partners	Awardees that reported eligible partners of this type	Number of Partners Receiving a Distribution	Number of Partners receiving NPRA	Number of Partners receiving ICS
Physicians	101	324	242	139
PGPs	51	12	10	3
Hospitals	57	20	19	4
Institutional PAC	18	11	11	-
HHAs	13	2	2	1
Other	12	10	10	-

Source: Lewin analysis of Awardee-submitted data collected February 2016 and May 2016 for Model 2 EIs participating in BPCI between Q4 2013 – Q1 2016. HHAs = home health agencies.

Based on our interviews, the most common reason for using gainsharing was to increase engagement and improve communication with partners. A couple of interviewees explained that they wanted to motivate specific changes in behavior through gainsharing, such as reducing readmissions or influencing physicians' decisions regarding discharge location. Similar to results from previous years, there was no consensus among interviewees on the ability of gainsharing to change behaviors. Some interviewees did not think that gainsharing would change provider behavior, whereas others believed gainsharing played a strong role in the adoption of new behaviors and processes (e.g., using patient risk stratification and new protocols).

A minority of the Model 2 interviewees stated that their organizations did not participate in gainsharing. There were diverse reasons given for this decision. One of the more often cited reasons was that interviewees were waiting to decide if gainsharing would be beneficial. These interviewees wanted additional data or more experience in BPCI before moving forward with gainsharing. Another reason for not gainsharing specifically with PAC providers was the belief that providing patient volume was a much bigger incentive to cooperate with EIs than gainsharing. Some of the Model 2 interviewees chose not to set up gainsharing agreements because their physicians were employed by the hospital and motivated to implement care redesign changes without gainsharing.

Telehealth and Post-Discharge Home Visit waivers

Only 32 out of the 699 Model 2 EIs used the telehealth waiver. As a result, we did not get much information about the waiver use in site visits or quarterly interviews. A couple of interviewees explained that the waiver could help them connect with rural patients, but they did not currently use the waiver.

Similarly, the post-discharge home visit waiver was rarely mentioned in interviews. Only two Model 2 EIs used the post-discharge home visit waiver, likely due to the challenges associated with using the waiver. For example, one of the interviewees noted that their patients required more services than what would be allowed by the waiver. Another interviewee felt that the payment amounts were not adequate given the restrictions on which providers were able to conduct the home visit. We also heard that one home visit every 30 days, on average, is not sufficient to meet the needs of patients who may not meet the definition of "homebound," yet would benefit from home visits.

5. Care redesign and cost savings

Interview Findings Reported Previously

- Standardization of protocols was frequently cited as an effective care redesign strategy, which included standardization of medications, development of pain management protocols, algorithms for selecting surgical implants, and protocols for discharge planning and follow-up care.
- Interviewees reported increasing efforts to discharge patients directly home (with or without home health services) instead of to an institutional PAC setting, when appropriate.
- Interviewees created methods to assess patient risk that included considering medical history, home safety, and current medications. Interviewees adopted varying approaches to care for patients with different risk scores.
- Interviewees stated that care navigators or care coordinators were a key component of their BPCI programs. These staff only provided services to BPCI patients and their positions were created as a result of BPCI. Generally, these staff met with patients and provided education prior to surgery, discussed post-discharge options, and tracked patients throughout the entirety of the episode.
- Interviewees noted an increase in the number of interdisciplinary team meetings to discuss BPCI patients' needs and progress.
- Interviewees indicated three general cost savings strategies: 1) standardize devices; 2) reduce PAC utilization; and 3) reduce readmissions.
- To reduce PAC utilization, interviewees reported trying to reduce SNF length of stay and substituting home health care or outpatient services for SNF when appropriate.
- Reducing readmission rates was another key strategy to achieve cost savings. Common methods used to reduce readmissions included patient risk stratification to target services and coordination with care partners to encourage patients to call the participant's office if they had a concern, rather than going straight to the emergency room.

Care redesign and care coordination were discussed during site visits and interviews this past year. Model 2 interviewees discussed tactics such as patient education and communication, discharge planning, patient follow-up, patient risk stratification, standardization of care pathways, and hiring new staff. Most Model 2 interviewees indicated that these changes were applied to all patients with a given condition, although some aspects of care redesign were described as being only for BPCI patients.

Model 2 interviewees stressed the importance of patient education and consistent messaging across all interactions with patients. A frequently mentioned theme was that BPCI resulted in more patient education, which was emphasized and delivered earlier. Although a few interviewees indicated that their patient education classes were mandatory for elective admissions, most interviewees reported that they were not required. Specific educational tools mentioned by interviewees included pre-hospitalization in-person classes, printed materials,

videos, and webinars. Educational topics included exercises for patients to perform before and after surgery, patient safety, symptom recognition, and appropriate places for follow-up care.

> "Once our physicians became involved in the discharge process and in setting expectations, our patients became a lot more agreeable."
> – Model 2 Hospital Representative

Model 2 interviewees indicated that their educational efforts were critical in managing patient and family expectations, particularly with post-discharge planning. Interviewees noted that BPCI led to an increased focus on educating patients about what happens after discharge. One interviewee described this as a cultural shift and commented that with their MJRLE patients, "This perspective has changed what we do during the joint class. Now we've added details about reasons to call your PCP, to call your surgeon, and to return to the ER." When patient education was discussed, the majority of interviewees stated that they discuss PAC options with patients, often sharing a provider list with them. Some interviewees reviewed PAC quality and preferred partners with their patients.

Interviewees indicated that many of their care redesign strategies focused on discharging patients earlier and controlling costs during the post-discharge period (PDP). In general, interviewees reported that they began planning for discharge earlier in the inpatient stay; some even began the process prior to the patient's surgery. Several Model 2 interviewees indicated that they conducted home evaluations - considering the layout of the residence as well as the availability of family support to ensure the patient could be discharged home safely. While numerous interviewees described their discharge planning process as a multidisciplinary team effort, some noted that their physicians' involvement in the process was essential to ease patients' concerns about being discharged directly home.

Model 2 interviewees indicated that after entering BPCI, staff spent more time considering the most appropriate discharge destination for their patients. One interviewee noted that they were willing to keep patients for an extra day in the hospital if it allowed them to then be discharged home rather than to a SNF. Another interviewee reported it may be more cost-effective to discharge patients to an inpatient rehabilitation facility (IRF) for a few days rather than to discharge the patient home and risk a readmission. A few PGP interviewees noted that their Convener provides them with the Care at the Right Location (CARL) tool to assist them in determining the appropriate discharge destination; however they reported that this tool was not frequently used. Interviewees noted that increased attention on identifying the most appropriate PAC setting based on patients' functional status and care needs was a result of their participation in BPCI.

> "We have very little influence on discharge from the SNF. That's the culture that we have to change."
> – Model 2 PGP

Model 2 interviewees indicated that they implemented post-discharge strategies to reduce hospital readmissions. Similar to what we heard in previous years, interviewees helped their patients schedule follow-up appointments with their PCP, and some maintained contact with their patients to ensure they attended the appointment. Additionally, interviewees noted that they monitored PAC providers' activities once their patients were discharged from the hospital. In particular, these interviewees tracked their patients' PAC LOS and monitored patients for readmissions. Two PGP interviewees also noted that they worked to ensure that the SNFs did not keep patients longer than medically necessary. Finally, Model 2 interviewees described various ways in which they coordinated with

their PAC partners, including notifying the PAC provider when they would receive a BPCI patient, sharing care protocols and discharge summaries with the PAC facility, and discussing any patient readmissions with their PAC partners. Some PGP interviewees also reported that they had staff members who would round on patients in the SNF setting.

Additionally, Model 2 interviewees often used risk stratification as a tool to create tailored approaches to their BPCI patients' care and to match patients' needs with available resources. Risk stratification was mentioned most commonly as a tool used to determine the frequency and intensity of follow-up after discharge, particularly post-discharge telephone calls. Multiple interviewees indicated that they conducted more follow-up calls with patients who had higher risk levels. For example, one PGP EI representative noted that patients were tracked throughout their hospital stay and assigned a red, yellow, or green risk score. This score was based upon clinical judgment, presence of comorbidities, health care utilization, and cognitive and functional status. The PGP EI then used the risk scores to prioritize patients for follow-up.

Other uses of risk stratification included predicting readmissions, identifying patients who needed additional inpatient care, determining the most appropriate discharge destination, and pre-operative medical optimization, that is, screening individuals to determine if they should delay surgery to achieve another medical goal such as weight loss or stabilization of diabetes.

Interviewees reported using various tools and methods for assessing patient risk. Some focused on patients' comorbidities alone, some used a variety of patient metrics beyond comorbidities, and others used validated tools. These tools included the LACE index (length of stay, acuity, co-morbidities, and emergency department visits), functional independence measure (FIM) scores, and CURB-65 (which is specific to pneumonia and includes confusion, blood urea nitrogen, respiratory rate, blood pressure and age 65 or older). A couple of interviewees mentioned using risk management tools that were created by their Conveners.

Similar to earlier reports, Model 2 interviewees focused on standardization of care protocols as one of their primary care redesign efforts. For example, a few interviewees described standardizing their pre-operative procedures to optimize patients for surgery, while some focused on developing standard discharge protocols to prepare patients to be discharged home. Interviewees also described various approaches to standardizing their pain management protocols. Some interviewees reported that they engaged their pharmacists in managing their BPCI patients' medications, while a few interviewees described testing various pilot programs for medication adherence. Conversely, one Model 2 PGP interviewee reported that they did not make any effort to standardize their care as a result of their participation in BPCI.

In addition to the standardization of care protocols, a few site visit interviewees discussed device standardization and consolidation of vendors as a strategy to achieve cost savings. Two of these interviewees reported that they restricted the number of their vendors or relied on a small set of preferred vendors to manage costs. One Model 2 hospital interviewee noted that they could apply significant market pressure on the price of implanted devices due to the size of their hospital system.

As previously reported, interviewees said they needed to hire additional staff for BPCI. Several reported hiring clinical staff, such as additional physical therapists or nurses. Some hired non-

clinical staff to assist with the additional administrative requirements of BPCI, such as data entry and coordination with PAC facilities. Multiple Model 2 interviewees created new roles, including case managers to identify BPCI patients in the ED, BPCI medical directors, and BPCI financial directors. Many interviewees highlighted the new position of care coordinator, which was typically responsible for discharge planning, patient follow-up throughout the episode, and communication with PAC providers. One PGP interviewee described the care coordinator role as the "glue" that holds their BPCI program together.

6. Implementation challenges

Interview Findings Reported Previously

- MJRLE episodes included patients who are receiving a joint replacement due to a fracture and those having an elective joint replacement. Patients with fracture are more likely to require post-acute care, have longer lengths of stay, and generally cannot benefit from interventions like patient education or pre-operative planning because there is limited time before the surgery.[4]
- Episodes may include services (such as mental health or substance abuse services) that episode initiators have no control over.
- Interviewees described situations where BPCI patients had clinically appropriate procedures unrelated to their BPCI episode that were scheduled to occur in the 30, 60, or 90-day period. Although these were unrelated to the BPCI episode, interviewees were still penalized for the readmission.
- Beneficiaries may transfer to Medicare Advantage, in which case they are no longer eligible for BPCI. This change in patient eligibility can impact the NPRA.
- Interviewees expressed dissatisfaction with their inability to direct beneficiaries to preferred PAC facilities. If PACs do not receive a significant patient volume, they are less likely to partner in redesign efforts. This can make it difficult for participants to track patients after hospital discharge.
- Other challenges to partnerships that interviewees cited included motivating PAC providers to reduce length of stay when appropriate and coordinating data sharing across systems that may not be compatible.
- Interviewees described difficulty in changing the culture of care for patients, families, and physicians, who may be hesitant to adopt new protocols and change the way they've practiced throughout their careers.
- Interviewees described challenges they faced in achieving cost savings, including the difficulty managing patients with chronic diseases, as they often require more extensive intervention, and the need to hire additional full-time employees to meet demands of participation, such as data analysis and patient tracking.
- Interviewees cited the need for focused outreach and communication to ensure buy-in and participation among gainsharing entities.
- Interviewees noted that taking a more localized approach to managing gainsharing may be more successful, as Conveners are often outside entities with few ties to EIs and their PAC providers.
- Interviewees described greater success gainsharing with physicians than with PAC partners due to challenges such as low patient volume for PAC providers and gainsharing not outweighing daily per diem revenues.

Interviewees noted that they have also faced challenges in redesigning PAC use. Interviewees stated that market characteristics often shape their ability to influence care redesign among PAC providers. In competitive PAC markets where there are many SNFs, interviewees reported that SNFs are more open to adopting care redesign to ensure continued collaboration with EIs.

[4] In response to concerns from Awardees about the difference in cost for MJRLE episodes, in October 2016, participants were allowed to have different target prices for MJRLE due to fracture and MJRLE not due to fracture.

Conversely, in less competitive markets with fewer SNFs, they are less likely to collaborate with EIs. Other interviewees noted challenges changing the culture among providers with regards to PAC use. One interviewee stated that physicians tend to focus on acute care only, so this participant has attempted to shift the physician's focus to the entire episode of care.

SNFs that were not affiliated with the EIs were less likely to be influenced by the hospital because of different financial incentives. Because of this, interviewees have had to balance their approach to maintain a positive relationship with SNFs while encouraging them to adopt care redesign interventions. Further, non-affiliated SNFs may have different data collection methods and electronic medical record systems. This has also been a challenge for interviewees as they attempt to transfer files and follow up on patients.

A few interviewees said that managing patient expectations regarding PAC was a challenge. Managing expectations was especially difficult when patients or their friends and family had prior experience in a PAC setting. Interviewees said that patients often wished to stay in the PAC facility for the maximum time period. Some patients and their families specifically had expectations for longer stays at SNFs. Reasons for this expectation included that the patients wanted to get the maximum-covered benefit, the SNF was convenient, or the SNF was resort-like and in high demand. One interviewee noted that providing education early in the episode was a key strategy for setting patient expectations.

Finally, medication reconciliation, pain management for patients after surgery, and information transfer from one provider to another were all identified as challenges. Site visit interviewees indicated that patients experienced challenges filling prescriptions and that any delays could result in readmissions. One interviewee noted that their IT system did not allow their medical team to transfer medication reconciliation information to other providers; therefore they had to manually transfer medication information. Several interviewees noted difficulties in sharing patient data between their BPCI team and PAC partners. Some interviewees noted that PAC partners were unwilling to share data or that a PAC facility was unable to share information because it used a different EHR system.

a. Model 2 PGP challenges

Model 2 PGP interviewees have a few unique challenges given their structure. One interviewee noted that hospital administrators were often less accommodating with PGP staff members compared to those in employed physician groups. Another PGP interviewee mentioned that decisions regarding patient care (e.g., discharge planning) were often made by hospital staff members (such as case managers) rather than by the physicians. Discrepancies between hospital codes (used in hospital billing) and current procedure terminology (CPT) codes (used in physician billing) were also reported as a challenge to the PGP structure in BPCI. As one interviewee explained, the BPCI program seems to be designed with hospitals in mind and that PGPs feel "like a square peg in a round hole." One interviewee mentioned that hospital partners have no motivation to follow the PGP processes or utilize their tools, which resulted in a strained partnership.

7. Participants that terminated or withdrew from BPCI

Interview Findings Reported Previously

- Among the Model 2 Awardees, participants that terminated their BPCI participation as of June 2015, the most frequently cited reason was the additional administrative burden and costs associated with BPCI.

During the past year we completed interviews with Awardees that terminated their BPCI participation. Each of the Model 2 Awardees we spoke with cited financial reasons for terminating their participation in BPCI. One respondent determined that additional resources, such as new analysts and greater utilization of physicians' time, were necessary for them to be successful in BPCI; however, they concluded that these additional costs combined with the anticipated risk for their clinical bundles were too great a financial burden for them to remain in the model. Other respondents noted that they faced large initial losses that were not sustainable, and they indicated that their losses came from events that they believed should have been excluded from their clinical bundles. The reintroduction of downside risk (i.e., being required to repay CMS for any spending over their target amount, which had been suspended in the previous year) was another reason that some respondents terminated their BPCI participation.

By the third quarter of 2015, 27 BPCI hospitals (6% of all Model 2 ACH EIs) had withdrawn from the initiative. Exhibits 32 and 33 compare the market, provider, and characteristics of hospitals that withdrew from BPCI within the first two years (Q4 2013 through Q3 2015) to hospital EIs that remain in the BPCI initiative.[5] On average, hospitals that withdrew had a longer average tenure in the BPCI initiative and were more likely to operate under a Convener than participating BPCI hospitals that had not withdrawn (Exhibit 32). Further, Exhibit 32 indicates that withdrawn hospitals initiated a similar number of episodes per clinical episode per quarter of participation compared to BPCI hospitals that have not withdrawn. Not surprisingly, withdrawn hospitals were also less likely to have positive NPRA during their tenure in the initiative.

Exhibit 32: BPCI characteristics, BPCI-participating Hospital EIs and Hospitals that have withdrawn from BPCI, Model 2, Q4 2013 – Q3 2015

Characteristic	BPCI Hospital EIs that have not Withdrawn (N=395)	Withdrawn BPCI Hospitals (N=27)
	Mean	Mean
Average Tenure in BPCI (months)	7.7	13.0*
Average number of BPCI episodes initiated per quarter within a given clinical episode	39.1	37.7
Percentage of Hospitals operating under a Convener	61%	70%
Percentage of EIs with Positive NPRA+	54%	33%

Source: Awardee-submitted data, claims data, and CMS' BPCI database.

*Indicates statistical significance at 5% level

+This measure takes into account EIs that initiated at least one episode between Q4 2013 and Q3 2015. 15 EIs were excluded because they initiated zero BPCI episodes over the course of the intervention period (not withdrawn N=380, withdrawn N=27).

[5] *Withdrawn Hospitals* include all BPCI hospitals that dropped all of their clinical episodes at some point during the first eight quarters of the initiative. *BPCI Hospital EIs that have not withdrawn* include all hospitals that have continued participation in at least one clinical episode during the first eight quarters of the initiative.

Exhibits 33a, 33b, & 33c compare BPCI-participating and withdrawn hospitals on a number of provider and market characteristics. Hospitals that withdrew from BPCI were similar to BPCI-participating hospitals that have not withdrawn from the initiative, with one exception. Compared to BPCI hospitals that remained in the initiative, withdrawn BPCI hospitals were more likely to be located in more populous regions (p=0.03) with an IRF.

Exhibits 33a, 33b & 33c: Provider and Market Characteristics, BPCI-participating Hospital EIs and Hospitals that have withdrawn from BPCI, Model 2, Q4 2013 – Q3 2015

Hospital Characteristics		BPCI Hospital EIs that have not Withdrawn (N=392)		Withdrawn BPCI Hospitals (N=27)	
		N	%	N	%
Ownership	Non-Profit	300	77%	21	78%
	Government	29	7%	3	11%
	For-Profit	63	16%	3	11%
Urban/Rural	Urban	361	92%	26	96%
Part of Chain	Yes	202	52%	14	52%

Characteristic	BPCI Hospital EIs that have not Withdrawn (N=392)	Withdrawn BPCI Hospitals (N=27)
	Mean	Mean
Bed Count	312	301
Number of Admissions for BPCI Episode MS-DRGs, 2011	3,029	2,645
Medicare Days Percent	39%	36%
Resident-to-bed ratio	0.12	0.11
Disproportionate Share Percent	27%	30%

Market Characteristics	BPCI Hospital EIs that have not Withdrawn (N=392)	Withdrawn BPCI Hospitals (N=27)
	Mean	Mean
Hospital-Market Share	21%	18%
Herfindahl Index	0.22	0.18
Medicare Advantage Penetration	25%	27%
PCPs per 10,000	7.9	7.9
SNF beds per 10,000	53.9	47.6
IRF in Market	67%	85%
Population	3,561,376	5,707,471*
Median Household Income	$53,275	$53,844
% Age 65+	13%	13%

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2015.

Note: Withdrawn hospitals are compared to all BPCI Hospital EIs that have not dropped all of their episodes from BPCI. Exhibits 33a, 33b, & 33c exclude three BPCI-participating hospitals that received Medicare certification after 2011.

*Indicates statistical significance at 5% level.

B. Model 2 Impact of BPCI

This section presents the impact estimates of Model 2 BPCI episodes on payments, utilization, and quality of care based on episodes that were initiated by acute care hospitals during the first eight quarters of the BPCI initiative (Q4 2013 through Q3 2015). Outcomes were analyzed across clinical episodes and in some cases across groups of clinical episodes. These non-mutually exclusive groupings include chronic condition episodes, planned episodes, episodes with greater than 40% of total baseline episode payments driven by the anchor inpatient stay, and the 10 episodes with the highest proportion of total baseline episode payments due to post-acute care (PAC) use. Please see **Appendix N** for more details on groupings of clinical episodes.

1. Payment, utilization, quality

We present results for the key outcomes across 23 Model 2 clinical episodes in this section. The exhibits present the estimated differential change in risk adjusted outcomes for patients receiving care from hospitals participating in BPCI between the baseline and the intervention period relative to the same change for the patients receiving care from hospitals in a comparison group (obtained from a difference-in-differences framework). See Section II.D for additional details on the statistical approach. All claim and assessment-based results for the 23 Model 2 clinical episodes are in **Appendix O**.

Across all Model 2 episodes, only MJRLE had a significant reduction in total payments for the inpatient stay plus 90-days post discharge. The lack of any other statistically significant results may be due to the fact that the average length of participation across all clinical episodes is only three quarters, a result of the large influx of participants in Q2 2015 and Q3 2015. Although not yet resulting in statistically significant reductions in total payments, there appears to be a change from the baseline to intervention period across clinical episodes in reducing SNF and IRF utilization and increasing home health agency (HHA) utilization. Episodes with reduced utilization of institutional PAC and increased utilization of HHAs tended to also have reductions in total payments. In general, quality of care was maintained across all episodes.

a. Sample Characteristics

Before discussing the impact of BPCI on payments, utilization, and quality, we present some basic statistics by clinical episode in Exhibit 34 to better understand the BPCI sample used for the impact analysis. By clinical episode, the number of matched EIs ranged from 18 to 276; EIs initiated between 597 and 40,974 episodes over the first eight quarters of the initiative. Because providers were allowed to join BPCI over an extended period, these data represent an average of three quarters of participation. In all but one clinical episode analyzed, over 50% of participating EIs joined BPCI in Q7 or Q8. As many as 21 (28%) EIs discontinued participation in a single episode over the first eight quarters.

Exhibit 34: Characteristics of the Matched BPCI Providers included in the DiD Estimates, Model 2, Q4 2013 – Q3 2015

Clinical Episodes	Matched EIs (#)	Matched Intervention Period Episodes (#)	Average length of participation (quarters)	EIs that joined BPCI in Q8 (%)	EIs that joined BPCI in Q7 or Q8 (%)	EIs that terminated participation in BPCI (#)	Intervention Period Episodes from EIs that exited (%)
Acute myocardial infarction	74	2,345	3	24%	72%	10	27%
Cardiac arrhythmia	52	2,167	3	13%	75%	2	6%
Cardiac valve	18	2,135	4	33%	44%	2	9%
Cellulitis	48	2,190	3	21%	69%	1	4%
Congestive heart failure	131	13,552	3	26%	63%	21	28%
COPD, bronchitis, asthma	108	8,286	3	23%	69%	13	25%
Coronary artery bypass graft	33	1,681	4	36%	52%	2	0%
Esophagitis, gastroenteritis and other digestive disorders	39	1,379	2	21%	82%	3	16%
Gastrointestinal hemorrhage	38	2,274	3	11%	58%	6	22%
Hip & femur procedures except major joint	66	2,823	3	23%	56%	5	7%
Major bowel procedure	25	1,086	4	28%	40%	4	11%
Major joint replacement of lower extremity	276	40,974	3	36%	66%	15	8%
Medical non-infectious orthopedic	65	2,591	3	23%	66%	4	8%
Nutritional and metabolic disorders	26	894	3	31%	65%	1	11%

Clinical Episodes	Matched EIs (#)	Matched Intervention Period Episodes (#)	Average length of participation (quarters)	EIs that joined BPCI in Q8 (%)	EIs that joined BPCI in Q7 or Q8 (%)	EIs that terminated participation in BPCI (#)	Intervention Period Episodes from EIs that exited (%)
Other respiratory	46	1,736	2	17%	80%	2	3%
Percutaneous coronary intervention	32	1,629	3	34%	75%	4	24%
Renal failure	51	2,927	2	24%	73%	4	26%
Revision of hip or knee	22	597	4	18%	32%	6	37%
Sepsis	97	10,484	2	32%	79%	6	18%
Simple pneumonia and respiratory infections	103	9,523	3	28%	71%	11	20%
Spinal fusion (non-cervical)	26	1,242	3	27%	42%	6	34%
Stroke	57	3,802	3	32%	63%	7	18%
Urinary tract infection	58	2,394	2	26%	81%	2	1%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2015 for BPCI providers. This table is limited to the matched BPCI providers used to calculate the difference-in-differences (DiD) results in the reminder of this section.

b. Key payment, utilization, and quality of care outcomes

How have the average standardized allowed amounts (Medicare payments and coinsurance/ copayments combined) changed under BPCI?

There were no clear patterns across all clinical episodes in the relative change from baseline to intervention in the total allowed payment amount (2015$) and in the total amount included in the bundle definition. Also, only a few of the differences between BPCI and comparison episodes were statistically significant (Exhibit 35). One reason we may not observe strong patterns or statistically significant results in these two payment outcomes is that the average tenure under BPCI was three quarters, as only relatively few participants joined BPCI at the outset of the initiative. A much larger number of EIs joined in the last two quarters (Q2 2015 and Q3 2015). Given the care redesign necessary to change payment and quality outcomes, it may be too soon to observe changes.

However, patterns in specific payment measures are apparent across clinical episodes. In BPCI episodes, the HHA standardized allowed amount increased relative to the change in the comparison group in 19 of the 23 clinical episodes (83%) and the change was statistically significant in six clinical episodes ($p<0.10$). As HHA payments increased, SNF payments tended to decrease. Among the 19 clinical episodes in which HHA payments increased relative to the change in the comparison group, 12 clinical episodes (63%) had relative declines in SNF payments (Exhibit 36). However, only one of these clinical episodes showed a statistically significant reduction in SNF payments ($p<0.05$). Among the 12 clinical episodes with sufficient sample size to calculate IRF payments, IRF payments decreased relative to the change in the comparison group for two thirds of the episodes ($p<0.05$ for two clinical episodes). These patterns suggest that BPCI participants attempted to reduce episode payments by reducing institutional PAC use and increasing use of HHAs, a strategy numerous BPCI participants have indicated they employ.

This pattern of reduced payments for PAC can also be observed in the 10 clinical episodes with the highest proportion of total costs during the baseline period for PAC. All 10 clinical episodes exhibited increased HHA payments (3 statistically significant, $p<0.10$) and 9 of 10 exhibited decreased SNF payments (1 statistically significant, $p<0.01$) (see **Appendix N** for more details on groupings of clinical episodes). In addition, total allowed payments for the inpatient stay and the 90-day PDP (2015$) declined relative to the change in the comparison group in all but one clinical episode. However, this reduction was only statistically significant for MJRLE ($p<0.01$).

In most clinical episodes, standardized allowed amounts for readmissions increased relative to the change in the comparison group, although none of these differences were statistically significant. It is important to remember, however, that hospital readmission rates have been declining since the initiation of Medicare's Hospital Readmissions Reduction Program in 2012. It may be particularly difficult for BPCI participants to continue to reduce readmission rates as they focus on reducing episode payments.

Exhibit 35: Impact of BPCI on Medicare Allowed Payment (2015$) Outcomes, by Clinical Episode, Model 2 ACH, Baseline to Intervention, Q4 2013-Q3 2015

Clinical Episode	Number of Episodes Q4 2013-Q3 2015	Total amount included in bundle definition[1]	Total allowed payment amount, IP through 90-day PDP	Readmissions Standardized Allowed Amount, 90-day PDP	SNF Standardized Allowed Amount, 90-day PDP[2]	HHA Standardized Allowed Amount, 90-day PDP[2]	IRF Standardized Allowed Amount, 90-day PDP[2]
Acute myocardial infarction	2,345	$205	$101	$104	$469	-$32	
Cardiac arrhythmia	2,167	$509	$483	-$87	$21	$72	
Cardiac valve	2,135	$680	$919	$398	$501	$161	-$371
Cellulitis	2,190	$6	$9*	$303	-$357	$188	
Congestive heart failure	13,552	-$315	-$253	-$49	$118	$70	-$143
COPD, bronchitis, asthma	8,286	-$109	-$59	$272	$33	$71	$29
Coronary artery bypass graft	1,681	$1,059	$1,238	$227	$274	$33	-$113
Esophagitis, gastroenteritis & other digestive disorders	1,379	$946	$925	$175	$522	-$35	
Gastrointestinal hemorrhage	2,274	-$754	-$367	$179	-$379	$91	
Hip & femur procedures except major joint	2,823	-$54	-$164	$177	-$940	$377	$291
Major bowel procedure	1,086	-$176	-$617	$189	-$494	$94	
Major joint replacement of the lower extremity	40,974	-$1,260	-$1,273	-$71	-$711	$92	-$435
Medical non-infectious orthopedic	2,591	-$824	-$1,032	-$90	-$616	$107	-$42
Nutritional and metabolic disorders	894	$44	$294	-$108	-$714	$23	
Other respiratory	1,736	$1,252	$1,305	$254	$299	$64	
Percutaneous coronary intervention	1,629	-$647	-$658*	$278	-$88	-$110	
Renal failure	2,927	-$238	-$32	$253	$26	$103	-$129
Revision of the hip or knee	597	$318	$482	$278	-$454	$115	
Sepsis	10,484	$20	-$114	$282	-$82	$97	-$127

Clinical Episode	Number of Episodes Q4 2013-Q3 2015	Total amount included in bundle definition[1]	Total allowed payment amount, IP through 90-day PDP	Readmissions Standardized Allowed Amount, 90-day PDP	SNF Standardized Allowed Amount, 90-day PDP[2]	HHA Standardized Allowed Amount, 90-day PDP[2]	IRF Standardized Allowed Amount, 90-day PDP[2]
Simple pneumonia & respiratory infections	9,523	-$499	-$224	$97	-$235	$62	$68
Spinal fusion (non-cervical)	1,242	$1,089	$1,013	$313	$198	-$137	-$466
Stroke	3,802	-$182	-$142	-$195	-$24	$39	$11
Urinary tract infection	2,394	$113	$135*	$375	-$523	$147	

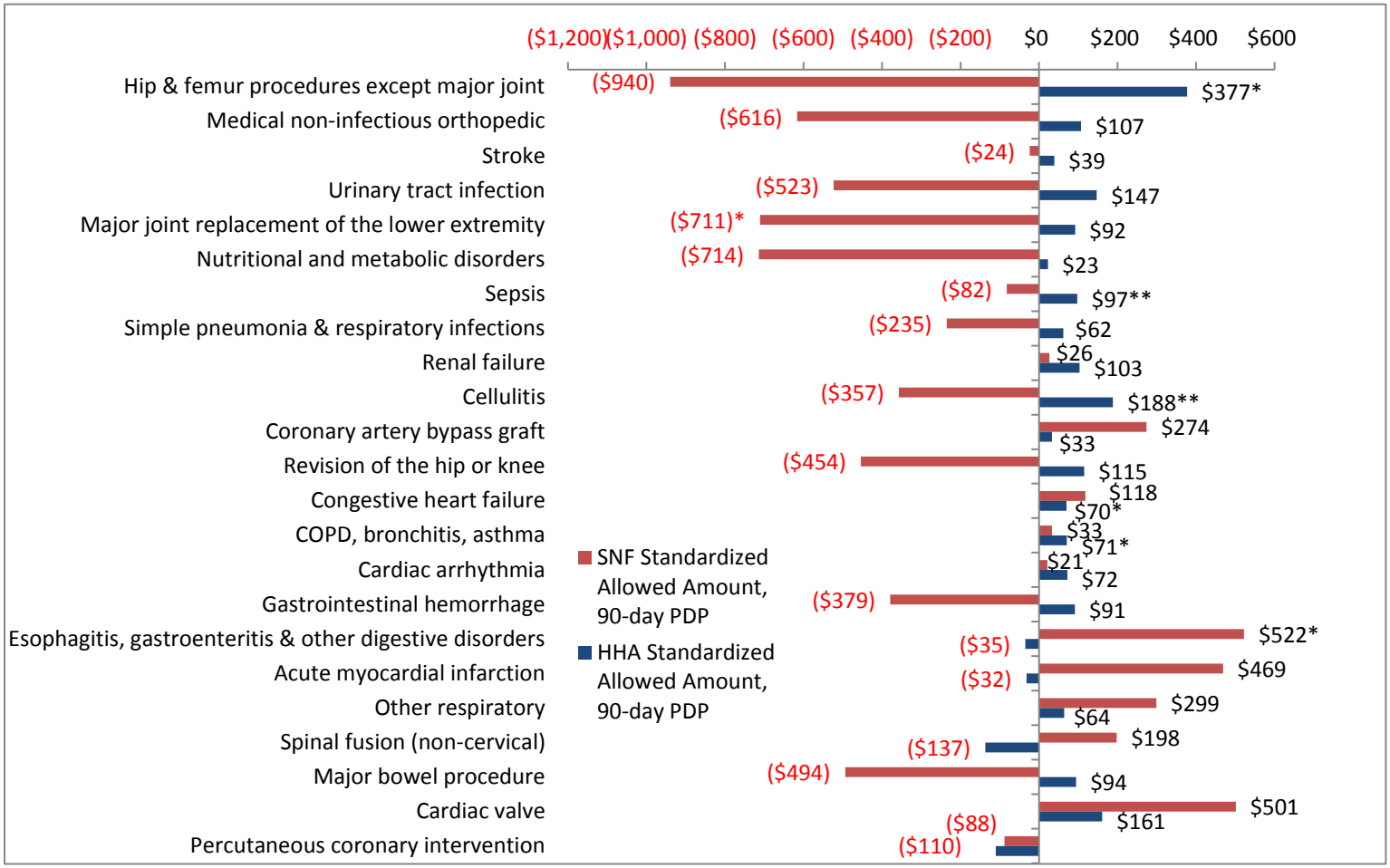
Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. A blank cell indicates that the outcome cannot be presented due to insufficient sample size. Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. PDP = post-discharge period. IP = inpatient.

[1]The total amount included in the bundle definition is based on only the 90-day episodes.

[2]These payment measures are not conditional upon use of the service.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.
* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

Exhibit 36: Impact of BPCI on HHA and SNF Payments, 90-day PDP, Sorted by Descending Proportion of Costs in Baseline for PAC, Baseline to Intervention, Model 2 ACH, Q4 2013 - Q3 2015



Note: These difference-in-differences (DiD) results are sorted by the proportion of costs during the baseline period for PAC and range from 52% (hip & femur procedures except major joint) to 8% (percutaneous coronary intervention). The payment measures are not conditional upon service use. Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. PDP = post-discharge period. IP = inpatient. SNF = skilled nursing facility. HHA = home health agency.
**Indicates DiD estimates are statistically significant at the 5% level
*Indicates DiD estimates are statistically significant at the 10% level

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How have the services changed under BPCI?

As shown in Exhibit 37 below, for most clinical episodes we detect no statistically significant change in the proportion of patients discharged to PAC. However, among those who received any PAC, there was a reduction in the percentage discharged to institutional PAC in about two-thirds of clinical episodes, three of which were statistically significant ($p<0.10$). Consistent with the PAC provider payment changes from the baseline to intervention period, there appeared to be an overall decline in the use of SNF care and an increase in HHA care across the clinical episodes (Exhibit 37). The number of SNF days for BPCI patients who used SNF care declined relative to patients in the comparison group in 16 of the 23 clinical episodes (although only two declines were statistically significant, $p<0.10$). The number of HH visits among BPCI patients who had at least one visit increased from the baseline to the intervention period relative to the change in the comparison group in 20 of the 23 clinical episodes (two statistically significant, $p<0.05$).

This change from baseline to intervention period is more notable across the 10 clinical episodes with the highest proportion of PAC provider costs in the baseline period, as shown in Exhibit 38. In all of these episodes the number of SNF days decreased relative to the change in the comparison group (statistically significant for two episodes, $p<0.05$). In 9 of these clinical episodes, the number of HH visits increased (with only one statistically significant, $p<0.01$). These increases and decreases in HHA and SNF utilization are consistent with the respective increases and decreases in HHA and SNF payments observed in these clinical episodes.

Across all clinical episodes, there was no systematic change in the inpatient length of stay (LOS). Hospitals have had financial incentives to reduce LOS since the inpatient prospective payment system was implemented in 1983. Therefore, it is not surprising to see minimal change in this utilization outcome. Even among the 7 clinical episodes with over 40% of their baseline costs attributable to the anchor stay, there was no clear direction in terms of changes in LOS as a response to BPCI. Additionally, most of these episodes had a high use of PAC in the baseline period, which could have led participants to focus on PAC rather than inpatient LOS.

Exhibit 37: Impact of BPCI on Utilization Outcomes, by Clinical Episode, Model 2 ACH, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of Episodes Q4 2013-Q3 2015	Percent discharged to PAC	Percent discharged to an institution out of those who received any PAC	Anchor Inpatient Stay LOS	Number of HH Visits, 90-day PDP[1]	Number of SNF Days, 90-day PDP[2]	Number of Institutional Days, 90-day PDP[3]
Acute myocardial infarction	2,345	0.9 pp	1.3 pp	-0.1	0.3	1.9	1.5
Cardiac arrhythmia	2,167	1.2 pp	1.1 pp	0.0	0.6	-1.2	-1.3
Cardiac valve	2,135	6.5 pp	-10.9 pp	0.2	0.2	2.6	3.9
Cellulitis	2,190	-0.6 pp	-3.9 pp	-0.1	0.4	-0.2	-1.1
Congestive heart failure	13,552	0.4 pp	1.6 pp	0.0	0.8	0.1	0.1
COPD, bronchitis, asthma	8,286	0.8 pp	1.2 pp	0.0	0.6	-1.5	-1.5

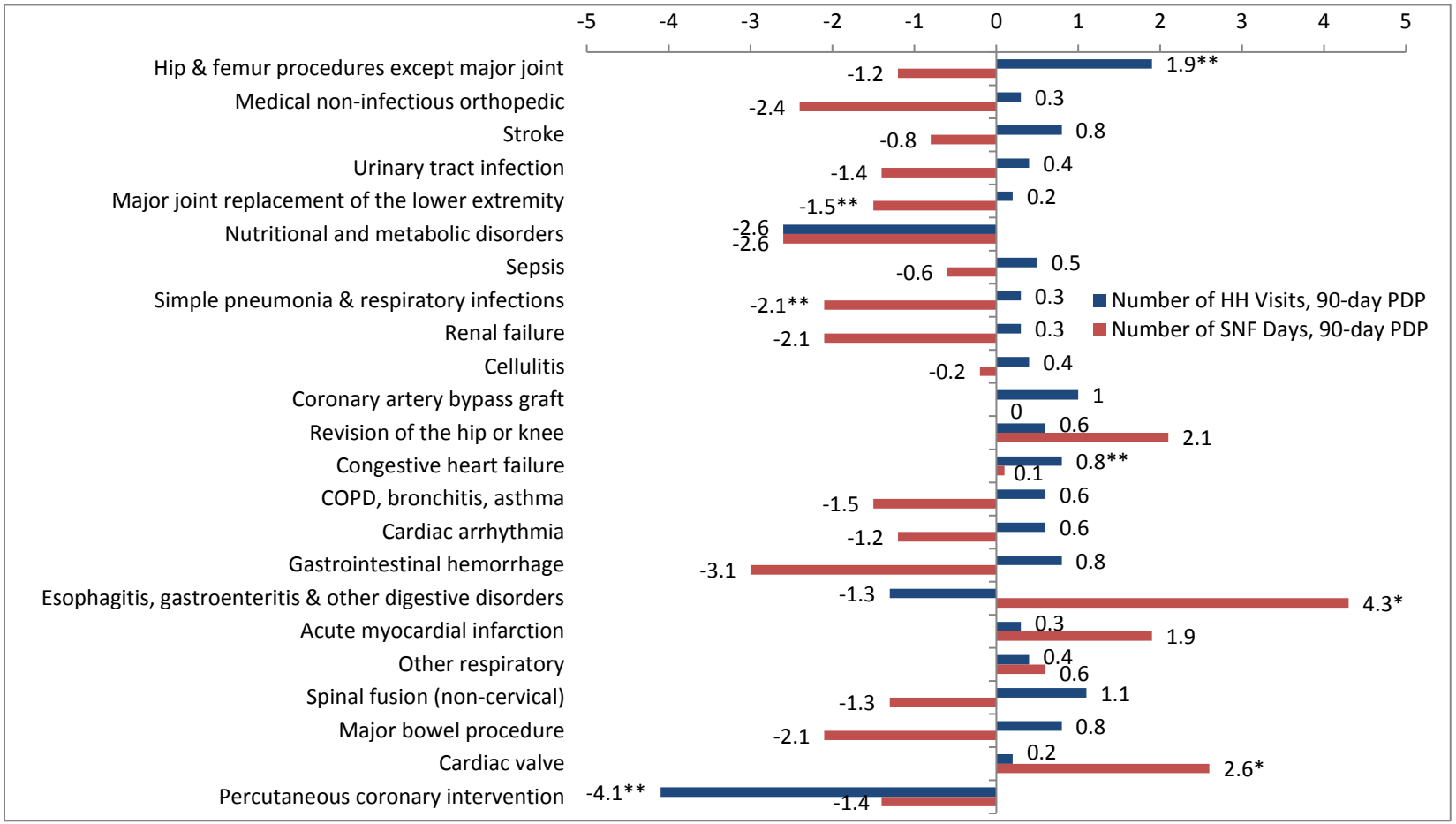
Clinical Episode	Number of Episodes Q4 2013-Q3 2015	Percent discharged to PAC	Percent discharged to an institution out of those who received any PAC	Anchor Inpatient Stay LOS	Number of HH Visits, 90-day PDP[1]	Number of SNF Days, 90-day PDP[2]	Number of Institutional Days, 90-day PDP[3]
Coronary artery bypass graft	1,681	-1.6 pp	-0.2 pp	0.2	1.0	0.0	1.7
Esophagitis, gastroenteritis & other digestive disorders	1,379	-2.8 pp	-1.3 pp	0.0	-1.3	4.3	3.4
Gastrointestinal hemorrhage	2,274	-1.1 pp	-0.5 pp	-0.1	0.8	-3.1	-3.8
Hip & femur procedures except major joint	2,823	0.2 pp	0.1 pp	0.0	1.9	-1.2	-1.7
Major bowel procedure	1,086	-2.2 pp	-1.6 pp	0.1	0.8	-2.1	-1.2
Major joint replacement of the lower extremity	40,974	-1.7 pp	-6.0 pp	-0.1	0.2	-1.5	-1.0
Medical non-infectious orthopedic	2,591	-1.7 pp	1.6 pp	0.0	0.3	-2.4	-2.5
Nutritional and metabolic disorders	894	-0.3 pp	-3.5 pp	-0.1	-2.6	-2.6	-2.4
Other respiratory	1,736	3.3 pp	-4.9 pp	0.2	0.4	0.6	0.5
Percutaneous coronary intervention	1,629	-2.1 pp	2.0 pp	-0.1	-4.1	-1.4	-2.3
Renal failure	2,927	-0.6 pp	-0.5 pp	-0.1	0.3	-2.1	-1.2
Revision of the hip or knee	597	3.7 pp	-2.6 pp	0.2	0.6	2.1	0.5
Sepsis	10,484	1.3 pp	-1.6 pp	0.2	0.5	-0.6	-0.7
Simple pneumonia & respiratory infections	9,523	1.7 pp	-1.7 pp	0.1	0.3	-2.1	-2.2
Spinal fusion (non-cervical)	1,242	-3.4 pp	4.7 pp	0.2	1.1	-1.3	1.1
Stroke	3,802	-0.1 pp	1.8 pp	0.0	0.8	-0.8	-1.7
Urinary tract infection	2,394	0.9 pp	-1.4 pp	0.0	0.4	-1.4	-1.8

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. PAC = post-acute care. PDP = post-discharge period. LOS = length of stay. HH = home health. SNF = skilled nursing facility.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

[1]Beneficiaries must have spent a minimum of one day in a HH setting during the 90-day PDP.
[2]Beneficiaries must have spent a minimum of one day in a SNF setting during the 90-day PDP.
[3]Beneficiaries must have spent a minimum of one day in a SNF, IRF, or LTCH setting during the 90-day PDP.

Exhibit 38: Impact of BPCI on HHA and SNF Utilization, 90-day PDP, Sorted by Proportion of Costs in Baseline for PAC, Baseline to Intervention, Q4 2013 - Q3 2015, Model 2



Are participants shifting services outside of the episode period or increasing services not included in the bundle, which may reduce overall savings to Medicare?

BPCI participants may attempt to reduce episode costs by changing the timing of services that would be otherwise included in the bundle so that they are furnished before the anchor hospitalization or after the end of the bundle period. Although these tactics could reduce episode payments, they would not necessarily achieve Medicare savings.

Among the 9 clinical episodes that are generally due to planned or elective hospitalizations, there may be an opportunity to time the anchor hospitalization such that unrelated outpatient services are not provided concurrently with the BPCI episode of care, which could reduce within-bundle costs (Exhibit 39). However, there was no indication that this occurred in planned episodes, based on both the lack of statistically significant results and the lack of any patterns in the changes in non-significant results in the 30-day pre-bundle Part B payment outcome. However, for 8 out of the 10 episodes for chronic conditions, there were greater increases in payments in the 30 days after the end of the episode, relative to the comparison group (though only one clinical episode had a statistically significant difference, $p<0.01$). This may indicate that BPCI participants had patients postpone certain services that would be included in the bundle until after the end of the episode. Alternatively, participants may have provided services to patients with chronic condition episodes to prevent a relapse or maintain their function past the end of the episode. The emergent pattern of increased post-episode spending was not evident for any of the other categories of clinical episodes.

Episode payments could also be lower if fewer services within the bundle definition are provided. However, if this results in higher payments for services excluded from the bundle, overall Medicare payments may not be lower. Certain readmissions and services for specific conditions are excluded from the bundle. These excluded readmissions and services vary by clinical episode.[6] Payments for services not included in the bundle increased in three of the 23 episodes (COPD, renal failure, and gastrointestinal hemorrhage were statistically significant at the 5% level, Exhibit 39). The payment increases were predominately due to excluded readmissions.

For COPD, renal failure, and gastrointestinal hemorrhage, the proportion of BPCI episodes that had any excluded readmissions increased relative to the comparison episodes (see Exhibit 40). Both excluded and included readmissions increased in COPD episodes, which may indicate an overall quality problem that contributes to readmissions. For renal failure and gastrointestinal hemorrhage episodes, excluded readmissions, but not included readmissions, increased.

[6] Centers for Medicare and Medicaid Services. (2016). Bundled Payments for Care Improvement Learning & Resources Area. Retrieved from https://innovation.cms.gov/initiatives/Bundled-Payments/learning-area.html

Exhibit 39: The Impact of BPCI on Allowed Payment Outcomes outside of the bundle, by Clinical Episode, Model 2 ACH, Baseline to Intervention, Q4 2013-Q3 2015

Clinical Episode	Number of Episodes Q4 2013-Q3 2015	Total amount not included in bundle definition[1]	Part B, 30-day pre-bundle period	Total Part A and B 1-30 post-bundle
Acute myocardial infarction	2,345	-$277	$62	-$99
Cardiac arrhythmia	2,167	$87	-$20	$386
Cardiac valve	2,135	-$29	$32	$479
Cellulitis	2,190	$121	$197	-$260
Congestive heart failure	13,552	-$52	-$7	$38
COPD, bronchitis, asthma	8,286	$154	$8	$15
Coronary artery bypass graft	1,681	-$14	$152	$365
Esophagitis, gastroenteritis & other digestive disorders	1,379	$121	$184	$67
Gastrointestinal hemorrhage	2,274	$331	-$54	-$484
Hip & femur procedures except major joint	2,823	-$138	-$12	$238
Major bowel procedure	1,086	-$38	-$178	$70
Major joint replacement of the lower extremity	40,974	-$21	-$21	-$84
Medical non-infectious orthopedic	2,591	$35	-$119	-$206
Nutritional and metabolic disorders	894	$4	$11	$221
Other respiratory	1,736	-$14	-$35	$1,235
Percutaneous coronary intervention	1,629	$70	$306	-$139
Renal failure	2,927	$271	-$20	-$381
Revision of the hip or knee	597	$20	$77	-$651
Sepsis	10,484	-$65	-$13	$192
Simple pneumonia & respiratory infections	9,523	$93	$33	-$105
Spinal fusion (non-cervical)	1,242	$91	$143	$327
Stroke	3,802	$98	-$112	-$224
Urinary tract infection	2,394	-$35	-$96	$61

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark light orange shaded cells, respectively. 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.

[1]The total amount not included in bundle definition values include 90-day episodes only.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers

Exhibit 40: Impact of BPCI on Any Readmission Excluded from the Bundle, Model 2 ACHs, Q4 2013-Q3 2015

Clinical episode	Number of Intervention Episodes		BPCI		Comparison		DiD Estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P value
COPD	7,858	7,840	2.9%	3.1%	3.2%	2.6%	0.8pp	0.1pp	1.5pp	0.03
Gastrointestinal hemorrhage	2,248	2,255	4.4%	5.3%	4.1%	3.4%	1.6pp	0.0pp	3.2pp	0.06
Renal failure	2,768	2,766	5.1%	5.6%	5.8%	4.9%	1.5pp	0.0pp	2.9pp	0.05

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. DiD = Difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; pp = percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

How has quality of care changed under BPCI?

Claim- and assessment-based outcomes suggest that across all Model 2 episodes, quality of care was generally maintained during the first two years of the BPCI initiative. Among patients discharged to a SNF, IRF, or HHA there were no patterns of reductions in the proportion of patients who had improved ADLs. These results were based on the approximately three-quarters of patients who received care long enough to have two assessments and may not be representative of all patients discharged to PAC. As shown in Exhibit 41, only 22 (10%) of the 211 ADL outcomes were statistically significant. For eight of these measures there was an increase in the proportion of BPCI patients that had improvement in ADLs, while the remaining 14 significant outcomes were a decrease in the proportion of BPCI patients that had an improvement relative to the comparison group. In addition, as found in the utilization results summarized above and as described by Awardees during our interviews, BPCI participants attempted to send fewer patients to PAC, particularly institutional PAC, and reduced institutional PAC LOS as a strategy to reduce Medicare FFS payments and increase NPRA. This, in turn, may have resulted in a sicker PAC patient population with less time to recover. Although our improvement in ADL outcomes control for beneficiary health status at PAC admission, we might expect to see reductions in the proportion of patients with improved ADLs during their PAC period due to the shorter SNF stays. However, there was not a strong relationship between SNF days and improvement in ADLs during the initial SNF stay. SNF days were reduced in 16 of the 23 clinical episodes, yet only 5 had a statistically significant reduction in at least one of the three SNF ADL outcomes ($p<0.10$).

Overall, claim-based results also suggest that the quality of care was maintained under BPCI Model 2. BPCI-participating providers had a relative decline in mortality that was statistically significant in three clinical episodes ($p<0.10$). There was also a statistically significant reduction in emergency department use for one clinical episode ($p=0.07$). While more than half of the clinical episodes had positive point estimates for the change in emergency department use and readmission rates, the magnitudes tended to be small and only two were statistically significant at the 10% level. Hip and femur procedures except major joint had an increase in unplanned readmissions of 2.7 percentage points ($p=0.06$) and spinal fusion (non-cervical) had an increase in emergency department use of 4.5 percentage points ($p=0.01$) relative to the change in the comparison group.

Exhibit 41: Impact of BPCI on Claim-based and Assessment-Based Quality Outcomes, by Clinical Episode, Model 2 ACH, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of Episodes Q42013-Q3 2015	All-cause mortality rate, 90-day PDP	ED Use, 90-day PDP	Unplanned readmission rate, 90-day PDP	ADL, HHA, improved ambulation	ADL, HHA, improved bathing	ADL, HHA, improved bed transferring	ADL, HHA, improved lower body dressing	ADL, HHA, improved upper body dressing	ADL, SNF, improved mobility function	ADL, SNF, improved overall function	ADL, SNF, improved self-care function	ADL, IRF, change in self-care score	ADL, IRF, change in mobility score
Acute myocardial infarction	2,345	-2.5 pp	-1.5 pp	1.9 pp	-4.5 pp	-0.6 pp	-0.1 pp	1.4 pp	-3.0 pp	2.0 pp	0.8 pp	2.3 pp	1.4	0.0
Cardiac arrhythmia	2,167	0.3 pp*	1.5 pp	0.0 pp*	-0.2 pp	-1.3 pp	0.6 pp	4.9 pp	10.8 pp	1.6 pp	2.2 pp	2.7 pp		
Cardiac valve	2,135	0.5 pp	0.9 pp	1.6 pp*	-1.2 pp	-3.3 pp	-0.1 pp	-3.8 pp	-2.7 pp	-2.1 pp	3.5 pp	-5.3 pp	0.0	-0.6
Cellulitis	2,190	1.1 pp	-2.2 pp	0.9 pp*	-1.1 pp	0.3 pp	3.6 pp	6.7 pp	7.1 pp	8.8 pp	2.1 pp	-0.8 pp	-0.6	0.2
Congestive heart failure	13,552	-0.4 pp*	-0.6 pp	-0.2 pp	2.5 pp	-2.1 pp	0.6 pp	-1.0 pp	-1.7 pp	2.2 pp	1.8 pp	-0.5 pp	-1.2	-0.2
COPD, bronchitis, asthma	8,286	-0.3 pp	1.0 pp	1.3 pp	1.6 pp	3.4 pp	3.1 pp	3.0 pp	1.0 pp	-1.5 pp	2.1 pp	0.4 pp	-0.2	-0.7
Coronary artery bypass graft	1,681	1.0 pp	2.4 pp	0.0 pp	1.2 pp	3.1 pp	0.7 pp	0.1 pp	1.0 pp	-1.8 pp	-3.0 pp	-0.2 pp	-1.3	-1.3
Esophagitis, gastroenteritis & other digestive disorders	1,379	-0.6 pp	2.7 pp	0.3 pp*	10.8 pp	8.2 pp	6.6 pp	3.22 pp	7.5 pp	-0.7 pp	-0.3 pp	-5.0 pp		
Gastrointestinal hemorrhage	2,274	-1.0 pp*	-0.2 pp	-2.2 pp	-5.5 pp	-3.7 pp	-2.2 pp	-3.3 pp	-2.0 pp	-5.2 pp	-1.3 pp	-4.3 pp		
Hip & femur procedures except major joint	2,823	-0.2 pp	0.3 pp	2.7 pp	6.1 pp	4.6 pp	5.5 pp	-3.1 pp	3.8 pp	-2.3 pp	-2.0 pp	-2.2 pp	-0.5	-0.4
Major bowel procedure	1,086	0.3 pp	0.9 pp	-0.7 pp	-1.2 pp	-3.4 pp	-2.5 pp	-4.0 pp	2.9 pp	3.9 pp	8.4 pp	5.6 pp	2.6	1.8
Major joint replacement of the lower extremity	40,974	0.1 pp	0.4 pp	-0.3 pp	-1.0 pp	-0.2 pp	-1.2 pp	-0.7 pp	0.0 pp	-2.9 pp	-2.2 pp	-3.9 pp	-0.3	-0.1

Clinical Episode	Number of Episodes Q42013-Q3 2015	All-cause mortality rate, 90-day PDP	ED Use, 90-day PDP	Unplanned readmission rate, 90-day PDP	ADL, HHA, improved ambulation	ADL, HHA, improved bathing	ADL, HHA, improved bed transferring	ADL, HHA, improved lower body dressing	ADL, HHA, improved upper body dressing	ADL, SNF, improved mobility function	ADL, SNF, improved overall function	ADL, SNF, improved self-care function	ADL, IRF, change in self-care score	ADL, IRF, change in mobility score
Medical non-infectious orthopedic	2,591	0.3 pp	1.2 pp	0.2 pp	0.8 pp	4.2 pp	-4.4 pp	-1.1 pp	-0.6 pp	1.5 pp	-1.6 pp	-3.8 pp	-0.7	-0.7
Nutritional and metabolic disorders	894	-3.6 pp	3.7 pp	2.7 pp	-9.0 pp	7.5 pp	-5.7 pp	6.8 pp	7.8 pp	13.0 pp	1.1 pp	-3.3 pp		
Other respiratory	1,736	-1.9 pp	1.7 pp	2.5 pp	3.0 pp	-0.7 pp	-3.9 pp	2.1 pp	0.3 pp	4.4 pp	2.4 pp	-3.5 pp	-0.9	-0.4
Percutaneous coronary intervention	1,629	0.7 pp	-0.7 pp	0.1 pp	3.1 pp	4.8 pp	2.6 pp	4.9 pp	7.6 pp	-12.2 pp	-4.0 pp	8.0 pp		
Renal failure	2,927	-2.4 pp	-1.2 pp	0.0 pp	4.1 pp	-2.3 pp	6.8 pp	4.8 pp	1.8 pp	-0.4 pp	4.9 pp	-4.4 pp	0.4	-0.4
Revision of the hip or knee	597	0.7 pp	3.1 pp	0.1 pp	-15.9 pp	-3.8 pp	-12.1 pp	-5.9 pp	-4.1 pp	-8.9 pp	-8.0 pp	-14.1 pp	2.0	2.1
Sepsis	10,484	-0.1 pp	-0.1 pp	0.2 pp	1.3 pp	3.9 pp	-0.6 pp	1.4 pp	2.0 pp	-1.0 pp	-1.9 pp	-3.8 pp	-0.1	0.4
Simple pneumonia & respiratory infections	9,523	0.6 pp*	-0.2 pp	-0.3 pp	-2.4 pp	4.6 pp	2.2 pp	1.0 pp	1.8 pp	-3.0 pp	-1.8 pp	2.1 pp	0.5	0.6
Spinal fusion (non-cervical)	1,242	0.0 pp	4.5 pp	0.6 pp	-2.6 pp	2.8 pp	-3.1 pp	1.5 pp	-4.3 pp	-2.1 pp	-8.6 pp	-13.4 pp	-2.4	-1.6
Stroke	3,802	0.4 pp	-2.3 pp	-0.1 pp	0.5 pp	-5.0 pp	-5.0 pp	-4.1 pp	-2.3 pp	-2.8 pp	-5.7 pp	-7.8 pp	0.0	0.6
Urinary tract infection	2,394	-1.5 pp	1.2 pp	0.6 pp	3.0 pp	-1.0 pp	1.9 pp	0.5 pp	-1.4 pp	-3.3 pp	-1.8 pp	0.2 pp	0.1	-0.2

Note: The estimates in this table are the results of a difference-in-differences model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. A blank cell indicates that the outcome cannot be presented due to insufficient sample size. ED=Emergency Department. PDP = post-discharge period. HHA = home health agency. ADL = activity of daily living. SNF = skilled nursing facility. pp = percentage points.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

2. Patient functional status, health status, and health care experience

Do patients treated under BPCI report differences in changes of functional status, health status, and health care experience compared with similar patients who were not in BPCI?

We have completed five waves of beneficiary surveys (Exhibit 42). In each survey wave, we asked respondents to report their functional status as they recalled it from before their clinical episode and also months later when filling out the survey. We asked about their overall mental and physical health and their health care experiences. This section of the report presents risk adjusted results for each survey wave, reflecting differences in outcomes between respondents receiving care from a BPCI EI and those in a matched comparison group. For Model 2, we include only results from surveys of beneficiaries with episodes initiated by acute care hospitals. We present analyses separately for each wave, rather than pooled across all waves, because EIs were able to opt in and out of clinical episodes over time, and also to report changes over time. Estimates for the BPCI and the comparison groups that include confidence intervals for the estimated differences are located in **Appendix P**.

Exhibit 42: Model 2 Clinical Episodes Analyzed for Beneficiary Survey, Wave 2-5

Clinical Episode	Wave 2	Wave 3	Wave 4	Wave 5
ACH – Cardiac arrhythmia				✓
ACH – Congestive heart failure (CHF)			✓	✓
ACH – Chronic obstructive pulmonary disease, bronchitis, asthma (COPD)			✓	✓
ACH – Major joint replacement of the lower extremity (MJRLE)	✓	✓	✓	✓
ACH – Simple pneumonia and respiratory infection (Pneumonia)			✓	✓
ACH – Sepsis			✓	✓
ACH – Stroke				✓
ACH – Urinary tract infection (UTI)				✓

Notes: Wave 1 results are excluded from this section because Wave 1 did not include episode-specific results. Episodes (respondents) in each survey wave were sampled from hospital discharges occurring in the following periods: Wave 2, October to November 2014; Wave 3, February to March 2015; Wave 4, May to June 2015; Wave 5, October to November 2015.

Exhibits 43-45 present point estimates of the differences in survey-based outcomes between respondents treated at BPCI EIs and respondents treated at matched comparison hospitals. Measures of interest are listed in the first row. Changes in functional status (Exhibit 43) are further classified as rates of improvement and rates of decline, from before to after the respondents' clinical episodes. Results across all measures are summarized below.

For most Model 2 episodes in our survey, BPCI and comparison respondents reported similar changes in functional status from before their hospitalization to after their episode, with a few exceptions. Specifically, BPCI respondents with a MJRLE or cardiac arrhythmia episode reported better changes in functional status relative to comparison respondents. In contrast, BPCI respondents with sepsis reported several worse changes in functional status, although the results were not consistent across waves. Model 2 BPCI and comparison respondents reported similar levels of physical and mental health status, with a couple of exceptions. Model 2 BPCI respondents with a sepsis episode were significantly more likely to report symptoms of depression, relative to comparison respondents; while Model 2 BPCI respondents with a MJRLE episode were significantly less likely to report symptoms of depression. Model 2 BPCI and comparison

respondents also generally reported similar health care experiences for most clinical episodes. In the few cases where significant differences occurred (i.e. MJRLE, pneumonia and sepsis episodes), such differences did not persist in multiple waves. There were also no statistically significant differences between BPCI and comparison respondents in their overall satisfaction with recovery, for any type of clinical episode.

a. Improvement and Decline in Functional Status

For most clinical episodes analyzed in Waves 2 through 5, BPCI and comparison respondents reported similar changes in functional status from before their hospitalization to after their episode, although there were some exceptions to this change (Exhibit 43). In three of the eight clinical episodes analyzed from Model 2, multiple measures were statistically significantly different between BPCI and comparison respondents in a consistent direction (i.e., all positive or all negative). Specifically, BPCI respondents with a MJRLE episode consistently reported better changes in functional status relative to the comparison group across multiple waves. BPCI respondents with sepsis indicated several worse changes in functional status in Wave 4, but we did not observe these differences in Wave 5. BPCI respondents with cardiac arrhythmia reported several superior changes in functional status in the single wave in which they were analyzed.

Although results were not statistically significant in all waves, Model 2 BPCI respondents with a MJRLE episode reported consistently higher rates of improvement relative to comparison respondents for two mobility-related measures: walking without rest (range: 3.3 to 8.2 percentage points) and using stairs (range: 1.1 to 7.9 percentage points). BPCI respondents with MJRLE episodes also consistently reported improvement in the degree to which pain limited their regular activities (over a range between 2.8 to 6.6 percentage points).[7]

Model 2 BPCI respondents who were hospitalized for sepsis tended to report worse functional status changes relative to the matched comparison group. In Wave 4, BPCI respondents indicated much smaller improvements, or even declines, for 4 of 7 measures (bathing, dressing, using the toilet, or eating, $p<0.01$; use of mobility device, $p=0.09$; walking without rest, $p=0.05$; and using stairs, $p=0.01$). However, none of these differences remained large or statistically significant in Wave 5, and only one measure (improvement in pain) was significantly worse among BPCI respondents with sepsis in Wave 5 ($p=0.08$).

For Model 2 cardiac arrhythmia episodes, which were only assessed in Wave 5, we found statistically significantly higher rates of improvement for 4 of 7 measures of functional status (planning regular tasks, $p=0.3$; use of mobility device, $p=0.04$; physical/emotional problems,

[7] Subsequent results from the Wave 6 (Spring 2016) and Wave 7 (Summer 2017) surveys indicated a reversal of the positive changes in functional status among Model 2 BPCI respondents with MJRLE episodes. In Wave 6, BPCI respondents did not report any significantly higher rates of improvement or lower rates of decline on any measure relative to comparison respondents. Rather they indicated significantly lower rates of improvement or higher rates of decline for 4 of 7 functional status measures, including regular activities limited by pain ($p<0.01$), and use of a mobility device ($p=0.03$) planning regular tasks, use of a mobility device, using stairs, and the degree to which pain limited their regular activities. In Wave 7, BPCI respondents did report significantly higher rates of improvement in pain limiting everyday activities ($p=0.03$), but still reported lower rates of improvement or higher rates of decline for 3 functional status measures, including planning regular tasks, walking without rest ($p=0.07$) and using stairs ($p=0.03$).

p=0.03; and pain, p=0.03), and significantly lower rates of decline for 3 of 7 measures (use of mobility device, p<0.01; physical/emotional problems, p<0.01; and pain, p=0.03).

b. Patient Reported Health Status

For five of the eight clinical episodes analyzed, BPCI and comparison respondents reported similar levels on a composite depression indicator, a self-reported measure of physical health, and a self-reported measure of mental health (Exhibit 44). However, Model 2 BPCI respondents with a MJRLE episode reported significantly fewer symptoms of depression relative to matched comparison respondents in Wave 4 (-5.0 percentage points, p=0.07) and Wave 5 (-4.0 percentage points, p=0.07). Additionally, BPCI respondents with a MJRLE episode reported better overall physical and mental health than the comparison respondents in Wave 4 (+6.4 percentage points for physical health, p=0.03; +3.8 percentage points for mental health, p=0.08). We also found that Model 2 BPCI respondents with a sepsis episode reported significantly more symptoms of depression relative to matched comparison respondents in Waves 4 and 5 (Wave 4: +5.8 percentage points, p=0.09; Wave 5: +6.3 percentage points, p=0.05). Although BPCI respondents with simple pneumonia reported more symptoms of depression than comparison respondents in Wave 4 (5.50 percentage points, p=0.10), BPCI and comparison respondents reported similar symptoms of depression in Wave 5. It is important to note that we were not able to measure changes in health status (depression, physical health, or mental health) for respondents over time. The survey asked about health status only after the episode and it did not ask respondents to recall their status before the episode. Although the results were case-mix adjusted using survey responses and data available on Medicare claims, it is possible that results reflect underlying unobservable differences in health status between BPCI and comparison respondents that existed before their BPCI episode, rather than the effect of the BPCI program.

c. Health Care Experience

Perceptions of care experience were generally similar between BPCI and comparison respondents (Exhibit 45). For two clinical episodes (cardiac arrhythmia and COPD) there were no statistically significant differences between BPCI and comparison respondents. In three additional clinical episodes (CHF, stroke, UTI) there was only one statistically significant difference.

Although we found few statistically significant results overall, we observed three clinical episodes where BPCI respondents reported significantly worse care experiences for multiple measures in a single wave (p<0.1). In Wave 2, BPCI respondents with a MJRLE episode were less likely than comparison respondents to report that they were discharged at the right time (-3.4 percentage points, p=0.07) and less likely to agree that medical staff clearly explained what follow-up appointments or treatments would be needed before they were discharged (-2.8 percentage points, p<0.01). BPCI respondents with a pneumonia episode in Wave 4 were also less likely than comparison respondents to agree that their preferences for post-discharge services were considered (-4.6 percentage points, p=0.06), and less likely to agree that they understood how to take care of themselves after discharge (-4.2 percentage points, p=0.02). In addition, BPCI respondents with a sepsis episode in Wave 4 reported worse experiences than the comparison respondents on four health care experience measures, including that they never received conflicting advice from the medical staff (-8.0 percentage points, p=0.03), always received the appropriate level of care (-8.4 percentage points, p=0.03), that medical staff always spoke to them in their preferred language (-4.1 percentage points, p=0.04) and that patient

preferences were taken into account in deciding what health care services were needed after leaving the hospital (-3.8 percentage points, p=0.10).

We also found one instance where BPCI respondents reported significantly better care experiences for multiple measures in a single wave. In Wave 4, BPCI respondents with a MJRLE episode were more likely than comparison respondents to report they never received conflicting advice from the medical staff (+6.5 percentage points, p=0.04), and medical staff always spoke to them in their preferred language (+1.9 percentage points, p=0.09).

In only one instance did any significant difference persist beyond one survey wave. Specifically, BPCI respondents with a MJRLE episode were more likely to report that staff always spoke to them in their preferred language in Waves 4 and 5 (Wave 4: +1.9 percentage points, p=0.09; Wave 5: +2.4 percentage points, p=0.09). There was no difference between BPCI and comparison respondents in overall satisfaction with recovery for any clinical episode that was significant at the 5% or 10% level.

Exhibit 43: Improvement and Decline in Functional Status, Difference between BPCI and Comparison Survey Respondents, Model 2 Acute Care Hospitals, 2014-2015

Clinical episode	Wave	N BPCI responses *	Bathing, dressing, using the toilet, or eating		Planning regular tasks		Use of mobility device		Walking without rest		Using stairs		Physical/ emotional problems limit social activities		Pain limiting regular activities	
			Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.
Cardiac arrhythmia	5	303	-0.26	-3.75	7.26	-4.64	6.40	-9.56	1.70	-3.03	-0.53	-3.62	8.31	-13.98	7.99	-7.73
Congestive heart failure	4	371	0.23	-0.69	4.82	-0.59	3.87	-1.46	1.39	0.46	2.82	-3.04	-3.09	2.20	-2.35	6.35
	5	349	1.41	1.20	2.01	-2.01	-5.84	9.00	-3.33	-0.74	1.41	-1.09	3.06	-2.16	0.46	-5.21
COPD, bronchitis, asthma	4	291	3.50	0.38	-1.03	0.31	0.46	-0.22	2.30	-2.91	2.51	-0.28	5.53	3.83	-3.24	0.05
	5	365	-4.97	7.99	0.21	2.89	0.92	-3.36	4.05	-1.39	-1.09	-2.98	4.09	-0.82	8.04	-6.21
Major joint replacement of the lower extremity	2	364	0.48	-0.13	-3.97	4.22	-2.71	2.52	8.16	-1.34	7.48	-1.31	3.81	-2.31	4.02	-1.45
	3	365	2.65	-2.44	2.30	-2.03	0.10	0.33	3.30	-1.49	4.04	-2.61	-0.02	-0.47	3.73	-3.44
	4	369	-1.32	-1.01	2.90	-0.99	1.93	-4.33	4.17	-1.15	7.92	-3.00	7.06	-1.31	6.57	-6.25
	5	372	8.42	-5.84	4.06	-4.43	-2.04	0.66	5.33	-2.90	1.10	1.17	6.54	0.01	2.78	2.15
Simple pneumonia and respiratory infections	4	367	-3.87	6.05	-0.88	1.19	0.09	0.49	2.26	1.06	3.26	-0.94	-2.55	3.44	5.88	-4.97
	5	333	3.21	-3.90	6.52	-3.90	-0.19	1.33	0.37	-1.27	-1.40	3.45	-0.85	1.82	-1.83	7.11
Sepsis	4	344	-9.01	5.32	-2.73	1.73	-2.95	4.99	-4.96	6.68	-7.78	3.81	-2.47	2.24	-1.41	0.21
	5	408	3.35	-4.56	-1.11	0.09	0.33	-0.82	-0.46	1.21	-0.55	0.62	-1.43	0.57	-6.24	-0.19
Stroke	5	396	-4.86	-1.31	0.02	-2.54	1.76	-2.31	2.23	-5.15	0.15	-0.82	-1.33	-0.18	4.63	-1.84
Urinary tract infection	5	298	-4.29	2.05	1.37	-0.26	0.28	-0.44	1.79	-0.66	6.72	-3.28	5.98	-3.26	3.44	-1.26

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P**.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. Abbreviations: COPD, Chronic obstructive pulmonary disease; Dec., decline in functional outcomes or stay at the worst status; Imp., improvement in functional outcomes or stay at the best status.

Source: Lewin analysis of BPCI and comparison Model 2 beneficiary surveys collected during 2014 and 2015.

Exhibit 44: Patient Reported Health Outcomes, Difference between BPCI and Comparison Survey Respondents, Model 2 Acute Care Hospitals, 2014-2015

Clinical episode	Wave	N BPCI responses*	Composite depression binary indicator†	Self-reported physical health binary indicator‡	Self-reported mental health binary indicator‡
Cardiac arrhythmia	5	303	-3.83	4.46	1.93
Congestive heart failure	4	371	1.21	-2.47	3.81
	5	349	2.40	0.58	-0.56
COPD, bronchitis, asthma	4	291	0.40	3.50	-0.32
	5	365	-1.34	-3.82	4.86
Major joint replacement of the lower extremity	2	364	1.66	-0.15	0.65
	3	365	-2.80	1.44	0.71
	4	369	-5.04	6.39	3.81
	5	372	-4.00	3.72	0.43
Simple pneumonia and respiratory infections	4	367	5.50	-2.89	-3.11
	5	333	-1.57	-0.75	3.71
Sepsis	4	344	5.75	-0.96	-0.30
	5	408	6.32	-1.81	-4.12
Stroke	5	396	-3.73	1.89	-0.24
Urinary tract infection	5	298	-1.47	3.63	4.35

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P.**

†The composite depression indicator is a binary measure equal to one when respondents reported a score of 3 or more on the Patient Health Questionnaire-2 (PHQ-2) and otherwise equals zero.

‡ The self-reported physical and mental health are binary measures equal to one when respondents reported that their health was excellent, very good or good, and equal to zero when respondents reported fair or poor health.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* COPD, Chronic obstructive pulmonary disease.

Source: Lewin analysis of BPCI and comparison Model 2 beneficiary surveys collected during 2014 and 2015.

Exhibit 45: Health Care Experience, Difference between BPCI and Comparison Survey Respondents, Model 2 Acute Care Hospitals, 2014-2015

Clinical episode	Wave	N BPCI responses*	Never received conflicting advice	Always received appropriate care	Staff always used preferred language	Discharged at right time	Preferences considered for services after discharge†	Understand care of self†	Medication instructions clearly explained†	Follow-up explained†	Able to manage health†	Satis-faction with recovery ‡
Cardiac arrhythmia	5	303	-0.43	-1.56	0.16	-2.21	2.37	-1.05	-0.86	0.35	0.90	-4.51
Congestive Heart Failure	4	371	-2.09	3.46	2.75	0.46	0.42	-1.63	0.09	-1.11	0.38	1.18
	5	349	6.93	0.68	0.96	-1.32	-1.84	-2.45	-2.67	1.05	0.60	1.01
COPD, bronchitis, asthma	4	291	5.00	-1.39	-1.74	-3.01	-0.46	2.60	-2.13	1.87	0.09	-4.10
	5	365	0.37	-2.94	1.78	-0.24	1.69	1.01	2.13	2.71	1.04	5.79
Major joint replacement of the lower extremity	2	364	-1.91	-2.25	-2.32	-3.41	0.84	-0.11	-0.95	-2.75	1.03	-0.16
	3	365	1.28	2.72	-0.39	-1.70	0.95	-0.39	-2.19	-0.86	0.68	0.89
	4	369	6.48	-3.37	1.90	-1.20	1.30	-0.15	0.49	0.21	1.47	1.68
	5	372	1.45	2.14	2.36	-1.47	-0.52	-0.40	1.22	-2.23	0.53	-1.74
Simple Pneumonia	4	367	2.88	-3.94	-0.67	-3.75	-4.63	-4.24	-2.86	-0.58	1.52	-0.54
	5	333	-4.64	-1.18	-0.82	4.67	2.31	-0.55	-2.33	0.48	0.21	-1.59
Sepsis	4	344	-7.99	-8.42	-4.08	-0.93	-3.77	-1.33	-1.62	2.68	-0.84	-4.30
	5	408	2.54	-2.51	-3.54	-4.18	0.41	-0.80	-2.42	0.49	-2.44	-4.15
Stroke	5	396	-5.48	-2.96	-2.46	-7.02	2.22	-2.53	1.24	-1.55	-0.37	-1.42
Urinary tract infection	5	298	-2.20	3.54	1.98	-1.21	7.48	0.27	1.29	3.70	2.00	1.92

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P.**

†Measure reflects that respondents either agreed or strongly agreed with this statement.

‡ Measure reflects that respondents were either quite a bit satisfied or extremely satisfied with their recovery.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* COPD, chronic obstructive pulmonary disease.

Source: Lewin analysis of BPCI and comparison Model 2 beneficiary surveys collected during 2014 and 2015.

3. Change in patient mix

BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce episode payments below their target and increase their NPRA. However, with the exception of the planned non-fracture MJRLE episodes where patients are identified prior to their hospitalization for the procedure, it is unclear that hospitals could directly influence any change in patient mix. Regardless of their ability to select healthier patients after joining BPCI, changes in patient mix could potentially impact the participant's NPRA. The impact analysis on payment, utilization, and quality presented in Section B.1 above controls for changes in these patient characteristics, but it does not directly examine any changes in patient mix. In this section we present results for the patient characteristics across Model 2 ACH episodes to determine whether there are any differences in the changes in patient mix between BPCI and comparison providers that may affect Medicare payments. We compared the change between the baseline and intervention period for BPCI patients to the change in the matched comparison group of patients in basic demographic characteristics, prior health status, and the utilization of care in the six months prior to the index hospitalization. We identified strata where BPCI participants had a change in patient mix that differed from the comparison group based on the presence of at least one characteristic having a statistically significant change. See Section II.D for additional details on the DiD analysis.

Of the 25 Model 2 strata, eight strata had indications of a healthier BPCI patient mix in the intervention period relative to the change for the comparison group, while 10 strata had indications that the BPCI-participating hospitals had a less healthy patient mix in the intervention period relative to the baseline period (Exhibit 46). Non-fracture MJRLE had statistically significant results indicating a healthier BPCI patient mix in the intervention period relative to the change for the comparison group. Given that there may be opportunities for participants to influence the mix of patients receiving MJRLE not due to fracture and changes indicating a relative improvement in patient mix for these BPCI episodes, the results suggest potential patient selection. BPCI-participating hospitals' non-fracture MJRLE patients were less likely to have had prior health care use, including inpatient acute care hospital use, skilled nursing facility and institutional nursing facility use, and were more likely to have no prior institutional health care use in the intervention period than the baseline period, compared to the change for the comparison group. There were several indications that BPCI renal failure patients became healthier from before the initiative to the intervention period (patients eligible for Medicaid, patients who were disabled, HCC indicators, ED use, HH use, prior PAC) relative to the change for the comparison group. It is difficult, however, to understand how BPCI participants could select healthier renal failure patient given the characteristics of this condition.

For 10 out of 25 strata, there are some indications that the BPCI-participating hospitals had a less healthy patient mix in the intervention period than in the baseline period. Two of the 10 strata (cellulitis and sepsis) had at least four out of 14 variables that were statistically significant. For six of the 25 strata, changes in patient characteristics had a combination of results indicating both changes to a healthier and a less healthy patient mix from the baseline to the intervention period relative to the comparison group.[8] In the cardiac valve clinical episode, no outcomes were statistically significant.

[8] CABG is stratified by emergent/non-emergent episodes and MJRLE is stratified by non-fracture/fracture episodes.

Exhibit 46: Relative Changes in Claim-based Characteristics of BPCI and Matched Comparison Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 2, Q4 2011 – Q3 2015

Clinical Episode	Number of Episodes Q4 2013- Q3 2015	Demographics					Prior Health Status		Utilization Six Months Prior to Index Hospitalization									
		Age: 20-64 years	Age: 65- 79 years	Age: 80+ years	Female	Medicaid Eligibility	Disability, no ESRD	Count of HCC indicators	Inpatient Acute Care Hospital	Emergency Room Admission	Home Health	Inpatient Rehabilitation Facility	Skilled Nursing Facility	Psychiatric Hospital	Long-term Care Hospital	Institutional Nursing Facility	No Institutional Care	No Post-acute Care
Acute myocardial infarction	2,345	-1.8	-0.7	2.5	2.3	-1.5	-2.2	0.0	0.6	0.1	-1.9	0.1	0.5	0.2	-0.2	-0.1	0.6	0.5
Cardiac arrhythmia	2,167	-0.1	-2.8	2.9	-2.3	2.3	0.5	0.1	2.8	-0.5	2.8	0.3	-0.2	0.8	0.0	0.2	-3.0	-0.2
Cardiac valve	2,135	-0.7	-2.3	3.0	2.5	-0.4	-1.1	-0.1	-2.7	-1.4	-2.2	0.0	-1.1	0.0	-0.2	-1.1	3.3	3.4
Cellulitis	2,190	1.0	-1.3	0.3	1.0	0.3	0.8	0.3	6.7	-0.3	0.9	-0.1	3.7	0.2	0.1	2.3	-6.3	-2.4
Congestive heart failure	13,552	-0.5	0.7	-0.2	2.2	0.2	-0.3	0.1	0.4	-1.4	-0.1	-1.0	0.8	-0.1	-0.4	0.7	-0.5	0.5
COPD	8,286	0.8	0.6	-1.4	0.4	2.2	1.5	0.0	1.1	0.3	0.3	-0.4	0.8	0.3	-0.6	0.4	-1.2	-0.6
Coronary artery bypass graft (emergent)	868	-0.2	-2.1	2.3	3.6	3.4	0.3	0.1	-0.2	6.1	1.2	0.2	0.9	0.3	-0.1	1.1	-0.2	-4.7
Coronary artery bypass graft (non-emergent)	813	-2.2	0.8	1.4	3.2	-1.6	-3.8	0.2	3.6	1.4	1.0	-0.4	0.4	0.3	-0.4	0.0	-4.1	-0.4
Esophagitis, gastroenteritis and other digestive disorders	1,379	0.7	-0.7	0.0	2.0	1.2	0.8	-0.1	-0.4	0.0	1.1	0.1	0.7	-0.9	0.3	-0.6	-0.5	-3.8
Gastrointestinal hemorrhage	2,274	0.3	0.2	-0.4	-1.5	3.5	0.6	0.0	0.9	1.3	1.6	-0.6	-0.9	0.3	0.0	-2.1	-0.6	-3.0
Hip & femur procedures except major joint	2,823	1.3	-1.7	0.4	-1.2	5.7	0.8	0.1	0.5	-0.5	-0.3	-0.9	-0.1	-0.3	0.0	-0.6	0.6	2.8
Major bowel procedure	1,086	-1.3	0.4	0.9	-1.0	-3.9	-1.5	0.1	-0.6	-0.4	-0.3	0.1	1.5	0.6	-0.1	1.4	-0.2	1.6
Major joint replacement of lower extremity (fracture)	5,506	-0.4	0.5	-0.1	0.6	-1.1	-0.2	0.1	0.9	0.9	-0.3	1.0	1.1	0.4	0.0	1.5	-1.9	0.0
Major joint replacement of lower extremity (non-fracture)	35,468	-0.5	0.5	-0.1	-0.4	-0.5	-0.4	0.0	-0.6	-0.3	-0.4	0.0	-0.6	0.0	0.0	-0.4	0.6	0.5
Medical non-infectious orthopedic	2,591	0.7	-2.6	1.8	0.6	2.6	0.9	0.0	2.2	-0.2	0.5	-0.8	2.3	0.2	-0.2	2.7	-1.6	1.1
Nutritional and metabolic disorders	894	-0.7	-4.7	5.4	-2.3	-3.9	-0.4	-0.1	-0.6	-2.1	3.6	-0.2	0.0	-0.6	-0.3	-1.6	0.1	1.8
Other respiratory	1,736	-2.4	5.3	-2.9	-2.3	-7.3	-2.1	0.2	3.1	1.7	-0.1	0.4	-1.5	0.4	-0.3	-1.1	-2.7	1.6

Clinical Episode	Number of Episodes Q4 2013- Q3 2015	Demographics					Prior Health Status		Utilization Six Months Prior to Index Hospitalization									
		Age: 20-64 years	Age: 65- 79 years	Age: 80+ years	Female	Medicaid Eligibility	Disability, no ESRD	Count of HCC indicators	Inpatient Acute Care Hospital	Emergency Room Admission	Home Health	Inpatient Rehabilitation Facility	Skilled Nursing Facility	Psychiatric Hospital	Long-term Care Hospital	Institutional Nursing Facility	No Institutional Care	No Post-acute Care
Percutaneous coronary intervention	1,629	-0.9	1.6	-0.7	1.8	-1.3	-0.6	-0.1	-0.1	2.0	-3.7	-1.2	-1.0	-0.1	-0.4	-1.3	-0.4	-1.3
Renal failure	2,927	-2.8	-1.1	3.9	1.1	-2.9	-2.7	-0.3	-2.1	-4.3	-3.5	0.7	0.1	0.2	-0.4	-0.6	2.7	6.2
Revision of hip or knee	597	-0.1	6.7	-6.7	-5.2	-4.3	-0.5	0.1	5.1	1.0	0.2	1.7	3.7	0.2	-1.3	3.5	-3.7	-0.7
Sepsis	10,484	1.1	-1.6	0.5	0.4	1.2	1.2	0.1	0.2	1.1	-0.1	-0.1	1.4	0.3	-0.1	2.6	-0.5	-0.7
Simple pneumonia and respiratory infections	9,523	-0.4	0.8	-0.4	-0.1	-0.7	-0.3	0.1	0.5	0.1	-1.3	-0.3	0.0	0.2	-0.5	-0.1	-0.1	1.1
Spinal fusion (non-cervical)	1,242	-2.7	1.3	1.5	-0.3	-1.2	-3.0	0.2	-0.2	-0.8	-0.3	-0.2	1.6	0.1	0.0	1.1	0.2	1.1
Stroke	3,802	-0.4	-0.7	1.1	0.9	-0.7	-0.3	0.0	2.1	-2.5	-0.8	-1.0	0.2	0.2	0.1	0.3	-2.4	0.1
Urinary tract infection	2,394	0.1	-1.1	1.1	1.4	0.2	0.4	0.0	0.2	3.2	0.9	-0.8	1.3	0.6	0.3	-0.4	-1.4	-1.9

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

4. Market Dynamics

Participation in BPCI might result in a shift in provider market share for one or more clinical episodes. To measure whether this has occurred, we calculated the market share of Model 2 ACH EIs over time, to determine whether EIs captured a greater share of BPCI-eligible episodes after starting participation in BPCI. The market share of Model 2 ACH EIs was calculated as the number of BPCI-eligible episodes admitted to a BPCI EI, divided by the total number of the same type of episodes admitted to all hospitals in the market.

As discussed in the methods section, we defined a BPCI market as the CBSA in which an EI is located. We identified five markets where the CBSA was too large to accurately define a local health care market (Chicago, New York, Detroit, Cincinnati, and Los Angeles) and we excluded these markets and their EIs from the analysis of market share. We also excluded EIs that had a 100 percent market share for a clinical episode in the baseline period (i.e., there were no other hospitals in their CBSA with any Medicare patients in that clinical episode), and those that did not have at least one episode in a given clinical episode in each of the relevant time periods before and after BPCI began. The removal of the largest markets, and all their EIs, limits the generalizability of our findings, especially for large and highly competitive markets. However, removing these large markets and their EIs is necessary to draw meaningful inferences. Since the EIs located in the five largest markets each had small baseline market share (due to the many competitors in these large CBSAs) it is difficult to detect meaningful change in their market shares over time. Including these EIs in the market share analyses would obfuscate the changes occurring in more typical markets that are well defined by CBSA boundaries.

We also examined whether Model 2 ACH EIs tended to discharge their patients to a smaller set of PAC providers, possibly indicating deliberate channeling of patients to higher quality or more efficient PAC providers. To measure the concentration of BPCI-eligible episodes discharged from EI hospitals to PAC providers, we calculated the number of SNF providers to which 75 percent of an EI's patients who used SNF care were admitted, after being discharged from the hospital.[9] We did the same for patients who used HHA services after hospital discharge. To understand whether the PAC providers used by BPCI EIs' patients were of higher quality, we also examined the CMS quality ratings of SNFs and HHAs used by these patients. We calculated the share of patients who used SNF or HHA care after leaving the hospital, and used high-quality SNFs and HHAs, respectively. We defined high-quality SNFs as those with a five-star rating on CMS's Nursing Home Compare website, and high-quality HHAs as those with a four-star or higher rating on the Home Health Compare website. For both measures, we excluded EIs that did not have at least one episode in a given clinical episode that was discharged to a SNF (or HHA) in each of the relevant time periods before and after BPCI began.

We calculated these three measures (market share, PAC concentration, PAC quality ratings) separately for three clinical episode types: MJRLE, CHF, and sepsis. These are high-volume conditions in which most BPCI EIs are engaged, spanning both surgical (MJRLE) and non-surgical (CHF and sepsis) conditions, and one chronic (CHF) and one acute (sepsis) medical

[9] This measure reflects change in the size of the PAC referral network used by an EI over time. We explored how the number of PAC providers used by EIs varied across EIs based on cumulative shares of 50%, 75% and 90% of the EI's discharged patients. Our overall conclusion remained the same across the three thresholds, but we were best able to detect and present meaningful variation in the outcome across providers and over time with the 75% threshold.

condition. The three subsections below present changes in the three market measures for participating EIs, tracking the average change in rates from the baseline period (Q4 2011-Q3 2012) to the BPCI period across EIs. EIs were stratified by the calendar quarter they started participating in one of the three selected clinical episodes. We segmented the baseline period and each EI's BPCI participation period into six-month intervals using data through Q3 2015, and then calculated the change in each measure from the baseline period to each six-month BPCI participation period, to observe changes over time. The analysis focuses on the two largest EI cohorts for each clinical episode, for which we had at least six months of complete claims data in the BPCI participation period at the time of this analysis: Q1 2014 and Q2 2015 (i.e., the quarters when EIs started participation in a BPCI episode). Some cohorts had a small number of EIs, and our exclusion restrictions further reduce the number of EIs that are analyzed in each cohort. Therefore, caution is recommended when drawing conclusions from the results based on a small number of EIs.

Our analysis found that the magnitudes of the average changes in Model 2 ACH EIs' market shares of MJRLE, CHF, and sepsis episodes were small across EI cohorts. The changes we observed do not suggest that EIs increased their market share of BPCI-eligible episodes after starting participation in BPCI. Evidence for whether Model 2 ACH EIs channeled their patients to fewer SNFs or HHAs after starting participation in BPCI varied across the three clinical episodes. There is a consistent pattern, however, that BPCI Model 2 ACH EIs discharged greater shares of MJRLE, CHF, and sepsis patients to highly rated SNFs and HHAs after starting participation in BPCI.

a. Change in the Market Share of Model 2 ACH EIs

Has the provider market share of BPCI-eligible episodes changed under BPCI?

The changes in market share that we observed do not suggest that the market share for BPCI EIs participating in MJRLE, CHF, or sepsis episodes increased (Exhibit 47). In fact, we detected no clear pattern of change in market shares of MJRLE episodes over time. Also, the average market share of CHF and sepsis episodes declined slightly from baseline to the BPCI intervention period.

The magnitude of the average change in EIs' market share from baseline to the BPCI intervention period tended to be small across the clinical episodes and EI cohorts. For example, there were 45 Model 2 ACH EIs in the Q1 2014 cohort for MJRLE episodes, and their average baseline market share was 18.8%. On average, these EIs' market shares increased by 0.1 percentage points from baseline to the Q1 2014/Q2 2014 period, increased by 0.2 percentage points from baseline to the Q3 2014/Q4 2014 period, and decreased by 0.6 percentage points from baseline to the Q1 2015/Q2 2015 period. There were 20 Model 2 ACH EIs in the Q1 2014 cohort for CHF episodes. Their average market share was 19.8% in the baseline period. On average, these EIs' market shares declined by 1.5 to 1.8 percentage points from baseline to each of the three six-month BPCI intervention periods. The market shares for the 41 EIs in the Q2 2015 cohort for sepsis episodes declined by an average of about one percentage point in the Q2 2015/Q3 2015 period, from an average baseline market share of 16.1%.

Exhibit 47: Average Change in Market Share of MJRLE, CHF, and Sepsis Episodes for Model 2 ACH EIs

Clinical Episode	Statistic	Q1 2014 EI Cohort				Q2 2015 EI Cohort	
		Average Market Share in Baseline	Average Percentage Point Change from Baseline			Average Market Share in Baseline	Average Change from Baseline
			Q1 2014 / Q2 2014	Q3 2014 / Q4 2014	Q1 2015 / Q2 2015		Q2 2015 / Q3 2015
MJRLE*	N	45	45	45	45	69	69
	Mean	18.8	0.1	0.2	-0.6	15.9	-0.3
CHF**	N	20	20	20	20	45	45
	Mean	19.8	-1.5	-1.8	-1.7	16.3	-0.04
Sepsis†	N	6	6	6	6	41	41
	Mean	7.9	-0.3	-1.5	-0.8	16.1	-0.9

* *MJRLE*: 18 EIs in the Q1 2014 cohort and 15 EIs in the Q2 2015 cohort were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. Two EIs in the Q1 2014 cohort and five EIs in the Q2 2015 cohort had 100% market share at baseline and were excluded. One EI in the Q1 2014 cohort was excluded because they did not have at least one episode in all baseline and BPCI intervention periods.

** *CHF*: 10 EIs in the Q1 2014 cohort and five EIs in the Q2 2015 were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. One EI in the Q1 2014 cohort and five EIs in the Q2 2015 cohort had 100% market share at baseline and were excluded. One EI in the Q1 2014 cohort was excluded because it did not have at least one episode in all baseline and BPCI intervention periods.

† *Sepsis*: Five EIs in the Q1 2014 cohort and six EIs in the Q2 2015 were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. Three EIs in the Q2 2015 cohort had 100% market share at baseline and were excluded.

Source: Lewin analysis of Medicare claims data for discharges from Q4 2011 through Q2 2015 for BPCI and non-BPCI providers within CBSAs that have at least two BPCI providers and meet our market inclusion criteria.

b. Change in Number of PAC Providers Accounting for 75 Percent of Model 2 ACH EI PAC Discharges

Has the number of PAC providers receiving the Model 2 ACH EI PAC discharges changed under BPCI?

The evidence for whether Model 2 ACH EIs channeled their patients to fewer, selected SNFs and HHAs varied across the three types of clinical episodes, and between the types of PAC facilities the patients used (i.e., SNF vs. HHA) (Exhibit 48). Therefore there is no systematic pattern of increasing concentration in PAC providers used for BPCI patients. Compared to the baseline, EIs tended to discharge 75% of MJRLE patients to slightly more SNFs (i.e., lower concentration) but slightly fewer HHAs (i.e. higher concentration) in the BPCI participation periods. They tended to discharge 75% of CHF patients to slightly fewer SNFs and HHAs, and 75% of sepsis patients to more SNFs and HHAs.

For example, among 64 ACH EIs in the Q1 2014 cohort participating in MJRLE, 75% of MJRLE patients discharged to SNFs were concentrated in an average of 6.7 SNFs during the baseline period; this increased by an average of 0.3 to 0.7 SNFs from baseline to the BPCI intervention periods, indicating less concentration in the SNFs used by their patients. In contrast, for the same cohort of EIs whose patients used HHA care, 75% of MJRLE patients discharged to HHAs were concentrated in an average of 3.9 HHAs during the baseline period, and this

decreased by an average of 0.05 to 0.1 HHAs from baseline to the BPCI intervention periods, indicating more concentration in the HHAs used by their patients.

For 27 EIs in the Q1 2014 cohort participating in CHF whose patients used SNF care, the average number of SNFs to which 75% of patients were discharged was 4.9 in the baseline and this increased by an average of 0.4 SNFs from baseline to the Q1 2014/Q2 2014 period, decreased by 0.2 SNFs from baseline to the Q3 2014/Q4 2014 period and decreased by 0.04 SNFs from baseline to the Q1 2015/Q2 2015 period. For 30 EIs in the same cohort whose patients used HHA care, the average number of HHAs used for CHF patients was 3.4 in the baseline and this decreased by an average of 0.3 to 0.6 HHAs in the BPCI intervention periods.

For 50 EIs in the Q2 2015 cohort participating in sepsis whose patients used SNF care, the average number of SNFs to which 75% of patients were discharged was 5.3 in the baseline and this increased by an average of 0.9 SNFs in the Q2 2015/Q3 2015 period. Similarly, for 49 EIs in the same cohort whose patients used HHA care, the average number of HHAs used for sepsis patients was 2.9 in the baseline and this increased by an average of 0.8 HHAs in the Q2 2015/ Q3 2015 period.

Exhibit 48: Average Change in Number of PAC Providers Accounting for 75 Percent of MJRLE, CHF, and Sepsis Discharges to SNFs and HHAs from Model 2 ACH EIs

Clinical Episode	Statistic	Q1 2014 Cohort				Q2 2015 Cohort	
		Average Number of PAC Providers in Baseline	Average Change in Number of PAC Providers used from Baseline			Average Number of PAC Providers in Baseline	Average Change from Baseline
			Q1 2014 / Q2 2014	Q3 2014 / Q4 2014	Q1 2015 / Q2 2015		Q2 2015 / Q3 2015
MJRLE*	N (to SNF)	64	64	64	64	89	89
	Mean (to SNF)	6.7	0.7	0.3	0.7	5.8	0.3
	N (to HHA)	64	64	64	64	89	89
	Mean (to HHA)	3.9	-0.1	-0.05	-0.08	3.5	-0.3
CHF**	N (to SNF)	27	27	27	27	50	50
	Mean (to SNF)	4.9	0.4	-0.2	-0.04	3.6	-0.3
	N (to HHA)	30	30	30	30	52	52
	Mean (to HHA)	3.4	-0.6	-0.6	-0.3	3.2	-0.1
Sepsis†	N (to SNF)	11	11	11	11	50	50
	Mean (to SNF)	7.3	0.7	1.6	1.6	5.3	0.9
	N (to HHA)	11	11	11	11	49	49
	Mean (to HHA)	3.3	0.8	0.05	0.4	2.9	0.8

**MJRLE*: Two EIs in the Q1 2014 cohort were excluded because they did not have at least one episode discharged to a SNF or HHA in all baseline and BPCI intervention periods.

***CHF*: One EI in the Q1 2014 cohort and three EIs in the Q2 2015 cohort were excluded because they did not have at least one episode discharged to a HHA in all baseline and BPCI intervention periods.

† *Sepsis*: One EI in the Q2 2015 cohort was removed because it did not have at least one episode discharged to a HHA in all baseline and BPCI intervention periods.

Notes: N represents the number of EIs that discharged at least one patient to SNF or HHA.

Source: Lewin analysis of Medicare claims data for discharges from Q4 2011 through Q2 2015 for BPCI and non-BPCI providers within CBSAs that have at least two BPCI providers and meet our market inclusion criteria.

c. Change in Percent of Patients Discharged from Model 2 ACH EIs to PAC Providers with High CMS Star Ratings

Has the percent of patients discharged from Model 2 ACH EIs to high-quality PAC providers changed under BPCI?

Although there was no consistent change in the number of PAC providers to which Model 2 ACH EIs' patients were discharged, we did find evidence that EIs tended to discharge greater shares of MJRLE, CHF, and sepsis patients to highly rated SNFs and HHAs after they started participating in BPCI (Exhibit 49). Despite a few changes in opposite directions over the entire intervention period for EIs in the Q1 2014 cohort, the percentage of MJRLE, CHF, and sepsis patients discharged from these EIs to highly rated SNFs and HHAs generally increased after BPCI. For the Q2 2015 cohort of EIs, the share of patients discharged to highly rated SNFs and HHAs increased, on average, from baseline to the first six months of BPCI participation, for all three clinical episode types.

For example, among 64 ACH EIs in the Q1 2014 cohort participating in MJRLE episodes, the average share of MJRLE patients discharged to SNFs that used a five-star rated SNF was 35.9 percent in the baseline; this share increased by an average of 3.3 to 6.0 percentage points across the BPCI intervention periods. For the same EIs, the average share of patients discharged to HHAs that used a four-star or higher rated HHA was 31.3% in the baseline period. On average, this share declined slightly from baseline to Q1 2014/Q2 2014, but then increased by 0.3 percentage points from baseline to the Q3 2014/Q4 2014 period and increased by one percentage point in the Q1 2015 / Q2 2015 period.

Among 89 EIs that started participating in MJRLE episodes in Q2 2015, the average share of patients discharged to SNFs that used a five-star rated SNF in the baseline period was 33.2%, and this increased by an average of 1.9 percentage points from baseline to the Q2 2015/Q3 2015 period. For the same cohort of EIs, the average share of patients discharged to HHAs that used a four-star or higher rated HHA was 32.7% in the baseline period and this increased by an average of 1.4 percentage points from baseline to the Q2 2015/Q3 2015 period. Changes from the baseline to intervention period were similar for both cohorts of EIs participating in CHF and Sepsis episodes, and the magnitudes of change over time tended to be larger than the changes for MJRLE episodes.

Exhibit 49: Average Change in Share of Patients Discharged from Model 2 ACH EIs to SNFs and HHAs with High CMS Star Ratings: MJRLE, CHF, and Sepsis Episodes

Clinical Episode	Statistic	Q1 2014 Cohort				Q2 2015 Cohort	
		Average Share in Baseline	Average Percentage Point Change from Baseline			Average Share in Baseline	Average Change from Baseline
			Q1 2014 / Q2 2014	Q3 2014 / Q4 2014	Q1 2015 / Q2 2015		Q2 2015 / Q3 2015
MJRLE*	N (to SNF)	64	64	64	64	89	89
	Mean (to SNF)	35.9	6.0	2.9	3.3	33.2	1.9
	N (to HHA)	64	64	64	64	89	89
	Mean (to HHA)	31.3	-0.03	0.3	1.0	32.7	1.4
CHF**	N (to SNF)	27	27	27	27	50	50
	Mean (to SNF)	35.9	1.4	-3.0	7.0	32.7	3.6
	N (to HHA)	30	30	30	30	52	52
	Mean (to HHA)	21.0	-2.8	2.9	4.6	37.0	6.2
Sepsis†	N (to SNF)	11	11	11	11	50	50
	Mean (to SNF)	24.3	4.8	-0.2	1.9	24.4	1.6
	N (to HHA)	11	11	11	11	47	47
	Mean (to HHA)	7.9	5.0	17.3	8.6	38.6	4.1

**MJRLE*: Two EIs in the Q1 2014 cohort were excluded because they did not have at least one episode discharged to a SNF or HHA, in all baseline and BPCI intervention periods.

***CHF*: One EI in the Q1 2014 cohort and three EIs in the Q2 2015 cohort were excluded because they did not have at least one episode discharged to a HHA in all baseline and BPCI intervention periods.

† *Sepsis*: One EI in the Q2 2015 cohort was excluded because it did not have at least one episode discharged to a HHA in all baseline and BPCI intervention periods. Two more EIs were excluded from the Q2 2015 cohort because they did not have at least one episode discharged to a HHA with a calculated star rating in all baseline and BPCI intervention periods.

Notes: N represents the number of EIs that discharged at least one patient to SNF or HHA.

Source: Lewin analysis of Medicare claims data for discharges from Q4 2011 through Q2 2015 for BPCI and non-BPCI providers within CBSAs that have at least two BPCI providers and meet our market inclusion criteria.

C. Factors Contributing to the Variation in NPRA among BPCI Providers

This section examines the factors that may have contributed to the variation in the financial performance of BPCI-participating hospitals. Unlike the difference-in-differences (DiD) analysis, which quantifies the overall impact of BPCI by comparing changes in outcomes of BPCI providers to changes of a control group, this analysis assesses variation among BPCI hospitals. Furthermore, whereas the DiD analysis uses Medicare Part A and B payments to assess financial performance, this analysis focuses on providers' standardized net payment reconciliation amount (NPRA). NPRA measures the difference between the target price for services provided during the episode of care and the total dollar amount of Medicare fee-for-service expenditures for that episode.[10,11] We

[10] NPRA applies to Models 2 and 3. There is no NPRA for Model 4 because participants keep any difference between the prospectively determined amount and their episode costs.

[11] When a provider's episode payments are below the target price (i.e., positive NPRA), the provider receives this amount from CMS. When a provider's episode payments exceed the target price (i.e., negative NPRA) the provider may need to return the amount to CMS. The target price is set by discounting historical episode payments by 2 or 3 percent, depending on the Model and episode options selected.

examined NPRA across three clinical episodes in Model 2: chronic obstructive pulmonary disease (COPD), which includes bronchitis and asthma;[12] MJRLE; and congestive heart failure (CHF).[13]

Our analysis found that, on average, top performers (i.e., hospitals with the largest average NPRA per episode relative to their target price) across the three clinical episodes reduced their use of institutional post-acute care (PAC), increased the share of patients discharged home without any PAC, and reduced the rate of unplanned readmissions. Circumstances prior to the start of BPCI may have made it easier for hospitals with higher standardized NPRA to reduce their use of institutional PAC during the intervention. These hospitals historically had discharged a higher share of their MJRLE and CHF patients to PAC and had healthier CHF patients during the baseline period. While top performers made the largest shifts towards less expensive PAC services, top performers participating in MJRLE also experienced the most favorable shifts in patient mix, as evidenced by their younger and healthier patient population during the intervention period. At the same time, top MJRLE performers experienced an increase in their overall volume of discharges, which was driven by an increase in the average number of MJRLE discharges that were non-fracture relative to the baseline period among hospitals in this group. Across the three clinical episodes, MJRLE provided the largest positive NPRA. This might have created incentives for hospitals to increase the number of MJRLE procedures. Alternatively, it may be easier for hospitals located in markets where there is overall growth in MJRLE procedures to earn NPRA by capitalizing on efficiencies in providing care to a higher volume of patients.

1. Differences in Performance across COPD, MJRLE, and CHF Participants

Within the three clinical episodes we analyzed, we found variation in financial performance, as measured by NPRA among BPCI hospitals, as shown in Exhibit 50. For a given clinical episode, top performers were defined as hospitals with standardized NPRA above the 75th percentile of the standardized NPRA distribution. Average performers were defined as hospitals with standardized NPRA between the 25th and 75th percentiles of the standardized NPRA distribution. Bottom performers were those with standardized NPRA at or below the 25th percentile of the distribution.[14] COPD top performers had an average NPRA of $1,752 per episode, whereas bottom performers had an average NPRA of -$2,433. CHF top performers had an average NPRA of $2,002 per episode, whereas bottom performers had an average NPRA of -$2,361 per episode. Hospitals participating in MJRLE had larger gains relative to those observed among hospitals participating in COPD or CHF. MJRLE top performers had an average NPRA of $2,490 per episode, whereas bottom performers had an average NPRA of -$1,647 per episode.[15]

[12] Although this clinical episode consists of COPD, bronchitis and asthma, for the remainder of this section we will refer to all of them as COPD.

[13] BPCI providers were required to meet certain criteria to be included in this analysis such as having at least 50 COPD episodes during both the BPCI baseline and intervention periods. For more detail on all the inclusion and exclusion criteria, please see the Methods Section.

[14] To adjust for differences in baseline payments and thus target prices among BPCI providers, we standardize a hospital's average NPRA per case by dividing it by the hospital's average target price.

[15] The analyses presented in this section encompass the first seven quarters of the initiative (from October 2013 through June 2015) and a baseline period from October 2011 through September 2012.

Exhibit 50: Average NPRA by Performance Group, Q4 2013 – Q2 2015[16]

Episode	Performance Group	EIs (N)	Mean	Median	Min	Max
COPD	Top Performers	7	$1,752	$1,616	$1,386	$2,377
	Average Performers	16	-$95	$36	-$1,777	$1,266
	Bottom Performers	8	-$2,433	-$2,122	-$4,580	-$1,304
MJRLE	Top Performers	24	$2,490	$1,939	$1,423	$7,117
	Average Performers	48	$471	$329	-$377	$2,113
	Bottom Performers	25	-$1,647	-$1,222	-$6,053	-$412
CHF	Top Performers	12	$2,002	$1,692	$816	$5,172
	Average Performers	24	-$251	-$153	-$1,734	$940
	Bottom Performers	12	-$2,361	-$2,423	-$3,917	-$950

Source: Lewin analysis of net payment reconciliation amount (NPRA) data for episodes that began Q4 2013-Q2 2015.

The financial performance of hospitals for a clinical episode is consistent with the proportion of hospital EIs that stopped participating in the clinical episode. Through September 30, 2015, top performers had the lowest dropped episode rate across the three clinical episodes: 0% in COPD, 4% in MJRLE and 17% in CHF (see Exhibit 51). Except for MJRLE, bottom performers tended to have the highest dropped episode rates relative to average and top performers (63% in COPD and 67% in CHF). Hospitals participating in MJRLE had less incentive to stop participating in MJRLE relative to COPD or CHF participants given the higher likelihood of realizing positive NPRA from an MJRLE episode. Across all performance groups, COPD and CHF had the highest average dropped episode rates (39% for COPD and 29% for CHF), and MJRLE had the lowest (13%), as shown in Exhibit 51.

Exhibit 51: Hospital Episode Initiators that Stopped Participating in the Clinical Episode, by Performance Group, Q4 2013 – Q3 2015

Episode	Performance Group	EIs (N)	Number of EIs who stopped participating in the clinical episode	Percent of EIs who stopped participating in the clinical episode
COPD	Top Performers	7	0	0%
	Average Performers	16	7	44%
	Bottom Performers	8	5	63%
	Total	31	12	39%
MJRLE	Top Performers	24	1	4%
	Average Performers	48	8	17%
	Bottom Performers	25	4	16%
	Total	97	13	13%
CHF	Top Performers	12	2	17%
	Average Performers	24	4	17%
	Bottom Performers	12	8	67%
	Total	48	14	29%

Source: Lewin analysis of CMS's BPCI database and net payment reconciliation amount data for episodes that began Q4 2013-Q2 2015.

[16] While we used standardized NPRA to define performance groups, this table displays actual NPRA per episode summary statistics.

2. Strategies and Opportunities for Realizing Positive NPRA

Top performing hospitals used similar strategies to decrease Medicare payments across the three clinical episodes. A Spearman rank-order correlation[17] analysis found that hospitals with higher standardized NPRA decreased their use of institutional PAC across COPD, MJRLE, and CHF episodes (see Exhibit 52).[18] Instead of discharging patients to institutional PAC, hospitals with higher standardized NPRA participating in COPD and MJRLE discharged a greater share of patients to home health or home without home health (for hospitals participating in MJRLE, the correlation between higher standardized NPRA and the proportion of patients discharged to home health was statistically significant). Hospitals with higher standardized NPRA in CHF decreased both the share of patients discharged to institutional PAC and the share discharged to home health.[19] Decreases in the rate of unplanned readmissions were also associated with high performance across clinical episodes, although for hospitals participating in MJRLE episodes this correlation was not statistically significant.[20]

Circumstances at baseline may have provided more opportunity for hospitals with higher standardized NPRA to reduce payments during BPCI. We observed that prior to the start of BPCI, hospitals with higher standardized NPRA were associated with higher use of PAC, measured by a lower share of patients discharged home without any type of PAC, relative to hospitals with lower standardized NPRA (see Exhibit 53).[21] A younger and healthier patient mix (measured by age and overall HCC score) was also associated with higher performance, although this result was only statistically significant for hospitals participating in CHF. These initial circumstances might have facilitated reduced use of institutional PAC among hospitals with higher standardized NPRA.

High performance was also consistently significantly correlated with changes in patient mix among hospitals participating in MJRLE, more so than in other episodes. Relative to the baseline period, hospitals with higher standardized NPRA participating in MJRLE treated younger and less complex patients during the BPCI intervention period[22] (see Exhibit 53). MJRLE hospitals with higher standardized NPRA also had a decrease in their share of patients with fracture and at the same time had an increase in the average number of discharges each quarter. We explored changes in MJRLE volume in more depth and discuss our results below.

[17] A Spearman rank-order correlation assesses the strength and direction of an association between two ranked variables. The correlation is less sensitive to outliers as it calculates the correlation on the ranked values rather than on the values itself.

[18] The Spearman correlation coefficients for the change in percent discharged to institutional PAC were: COPD r_s=-0.40, MJRLE r_s =-0.52, and CHF r_s =-0.38. The coefficients for the change in percent discharged home without any type of PAC were: COPD r_s=0.15, MJRLE r_s =0.14, and CHF r_s =0.27. The coefficients for the change in the percent discharged to home health were: COPD r_s=0.20, MJRLE r_s =0.30, and CHF r_s =-0.01.

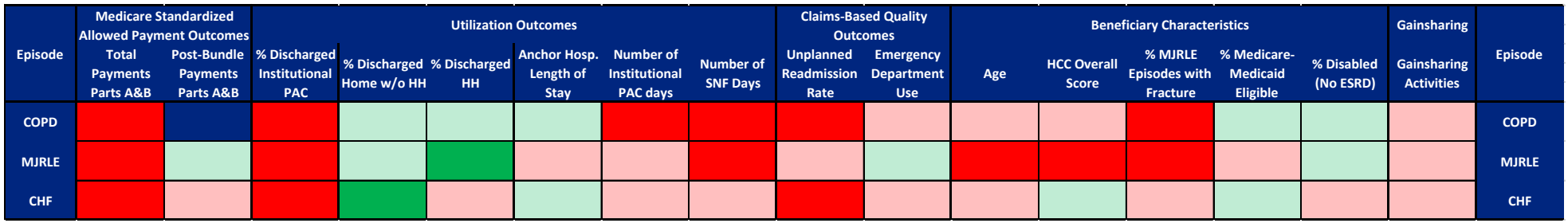
[19] For more detail on these numbers, please see the COPD, MJRLE and CHF issue brief chapters.

[20] The Spearman correlation coefficients for the change in the rate of unplanned readmission were: COPD r_s=-0.34, MJRLE r_s =-0.16, and CHF r_s =-0.25.

[21] This association was not statistically significant for COPD hospitals. The correlation coefficients for percent discharged home without any type of PAC were: COPD r_s=-0.01, MJRLE r_s =-0.19, and CHF r_s =-0.26.

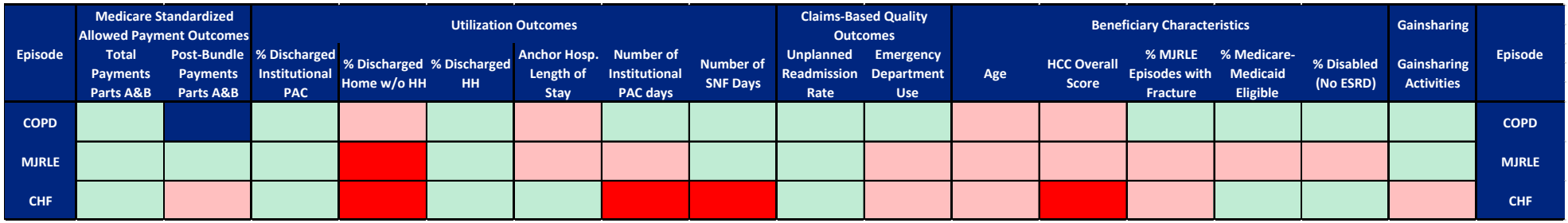
[22] We observed a similar pattern among COPD top performers, however, changes in case-mix were not statistically significant.

Exhibit 52: Spearman Rank-Order Correlation between Standardized NPRA and Changes in Payment, Utilization, and Quality Outcomes; COPD, MJRLE, and CHF Episodes; Model 2; Q4 2011 – Q2 2015[23,24]

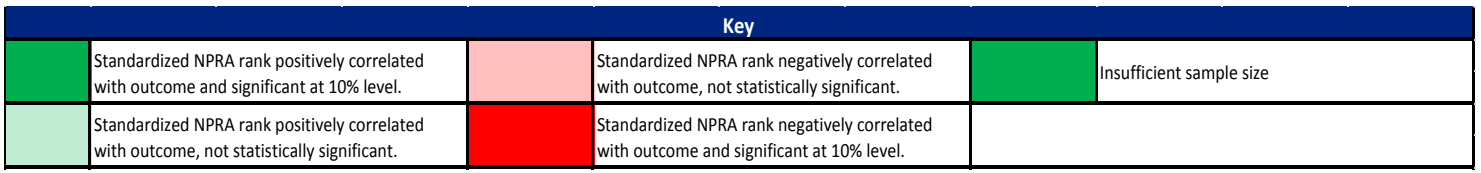


Source: Lewin analysis of Medicare claims, enrollment, and net payment reconciliation amount data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q2 2015 (intervention period).

Exhibit 53: Spearman Rank-Order Correlation between Standardized NPRA and Payment, Utilization, and Quality Outcomes at Baseline; COPD, MJRLE, and CHF Episodes; Model 2; Q4 2011 – Q3 2012



Source: Lewin analysis of Medicare claims, enrollment, and net payment reconciliation amount data for episodes that began Q4 2011 through Q3 2012 (baseline).



[23] Red indicates standardized NPRA was negatively correlated with an outcome (e.g., for all episodes, standardized NPRA is negatively correlated with the change in total Part A & B payments, meaning hospitals with high standardized NPRA had larger reductions in total payments). Green indicates standardized NPRA was positively correlated with an outcome. Dark shades indicate that the correlation was significant at the 10 percent level ($p \leq 0.10$).

[24] Gainsharing activities means a hospital has indicated it participated in the gainsharing waiver available under the BPCI initiative. Participation in the waiver does not guarantee that a hospital shared positive NPRA. Gainsharing participation only applies to the intervention period.

3. Changes in MJRLE Volume by Performance Group

Differences in financial gains may be correlated with changes in the volume of procedures. We hypothesize that providers have more control over the volume of scheduled, elective procedures than emergent procedures or admissions. Therefore, we focused on MJRLE volume because the majority of these procedures are elective and scheduled prior to the anchor hospitalization. We examined the correlation between standardized NPRA and change in MJRLE discharges from the baseline to the intervention period. We used discharges rather than episodes because the number of episodes ultimately assigned to a BPCI hospital depends on the mix and number of other active BPCI participants in the same market. Therefore any changes in the number of episodes may be due to the assignment of episodes to BPCI participants and changes in BPCI participants rather than provider responses to BPCI or volume changes throughout the market.

Higher performance was significantly correlated with having a higher number of MJRLE discharges in the baseline period.[25] There was also a statistically significant correlation between higher performance and an increase in volume of MJRLE discharges from the baseline to the intervention period.[26] There was a 12.1% increase in the average number of MJRLE discharges per quarter among all hospitals participating in the MJRLE episode, from 86.2 discharges per quarter in the baseline to 96.6 in the intervention period (see Exhibit 54). Top performers experienced the largest increase in MJRLE discharges. There was a 25.4% increase in MJRLE discharges among top performers, from 91.7 in the baseline to 115.0 in the intervention period.

Exhibit 54: Average Number of MJRLE Discharges per Quarter, by Performance Group, Model 2, Q4 2011 – Q2 2015[27]

Episode	Performance Group	EIs (N)	Average Number of Discharges per Quarter		% Change
			Baseline	Intervention	
MJRLE	Top Performers	24	91.7	115.0	25.4%
	Average Performers	48	95.5	104.5	9.5%
	Bottom Performers	25	62.9	63.8	1.4%
	Total	97	86.2	96.6	12.1%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q2 2015 (intervention period).

The changes we observed in volume do not imply that BPCI caused the rise in the number of MJRLE procedures. Without considering the volume of both BPCI and non-BPCI MJRLE discharges within markets, it is not possible to draw conclusions about whether BPCI-participating hospitals generated additional procedures. There are several possible explanations for the larger increase in MJRLE discharges among top performing hospitals. First, overall increases in MJRLE

[25] The Spearman correlation coefficient for the average number of MJRLE discharges per quarter in the baseline was r_s =0.20.

[26] The Spearman correlation coefficient for the change in the average number of MJRLE discharges per quarter from the baseline to the intervention period was r_s =0.22.

[27] BPCI providers were required to meet certain criteria to be included in this analysis such as having at least 50 MJRLE episodes during both the BPCI baseline and intervention periods. For more detail on all the inclusion and exclusion criteria, please see the Methods Section.

procedures in the market may have been concentrated among top performing hospitals. Second, volume may have shifted to higher performing BPCI hospitals (e.g., physicians who have admitting privileges at multiple hospitals could have performed surgeries at BPCI hospitals that they may have otherwise performed at non-BPCI hospitals or physicians who admit to top performing hospitals may have differentially developed positive reputations and attracted patients to higher performing hospitals). Alternatively, the increases in volume could be a result of induced demand, that is, an absolute increase in total number of procedures at BPCI hospitals that would not have occurred absent BPCI.[28]

As this analysis is correlational and does not control for market trends, it does not shed light on the relationship between NPRA and changes in volume. Because top performing hospitals participating in MJRLE realized an average of $2,490 of NPRA per episode (the highest average NPRA of top performers among the three clinical episodes), providers did have a financial incentive to increase the number of MJRLE discharges. However, we cannot determine with these data if that happened. Alternatively, it may be easier for hospitals to achieve positive NPRA in markets where the number of MJRLE procedures is growing, because increased volume may allow hospitals to standardize care more easily and provide more efficient care.

To investigate if the increases in the number of MJRLE discharges could be due to volume shifting to higher performing BPCI hospitals, we examined changes in the average number of MJRLE admissions to BPCI and non-BPCI hospitals among physicians who admitted MJRLE patients to BPCI hospitals during the intervention period (see Exhibits 55 and 56). We stratified admissions by fracture status. We hypothesized that if volume was shifting to higher performing BPCI hospitals, we would be more likely to observe changes among non-fracture patients because non-fracture MJRLE procedures are usually planned, whereas fracture procedures are not.

Higher standardized NPRA was correlated with an increase in non-fracture admissions to BPCI hospitals relative to non-BPCI hospitals (see Exhibit 55). Among physicians who admitted MJRLE patients to top, average, and bottom performing hospitals, there were increases in the number of non-fracture MJRLE admissions to both BPCI and non-BPCI hospitals from the baseline to the intervention period. Across all three performance groups, the increase in MJRLE admissions to BPCI hospitals was greater than the increase in admissions to non-BPCI hospitals. The largest increase in the number of non-fracture admissions to BPCI hospitals relative to the increase in the number of non-fracture admissions to non-BPCI hospitals was for physicians who admitted to top performing hospitals (a relative increase of 1.6 admissions per physician per quarter, see Exhibit 55).

At the same time, there were small increases in the number of fracture admissions to BPCI hospitals relative to non-BPCI hospitals from the baseline to the intervention period. Moreover, this increase was consistent across performance groups. Among physicians in all performance groups, there was an increase of 0.1 fracture admissions per physician per quarter to BPCI hospitals relative to non-BPCI hospitals (see Exhibit 56).

[28] These alternatives are not mutually exclusive, so any of these separately or in combination could potentially account for changes in volume.

As expected, the larger increase in MJRLE discharges among top performers was driven by an increase in non-fracture admissions. Because there were increases in non-fracture admissions to both BPCI and non-BPCI hospitals, it is not possible to conclude that physicians who admitted MJRLE patients to BPCI hospitals shifted non-fracture cases from non-BPCI hospitals to BPCI hospitals, though it remains a possibility. Disentangling shifts in admissions from overall market growth requires considering changes in both BPCI and non-BPCI volume within markets. The source of the increase in volume among BPCI hospitals (e.g., increases due to general market trends, shifts to BPCI hospitals, or potential induced demand) cannot be derived from these results alone as this analysis did not consider overall volume within markets. Our analysis that did consider volume within markets does not suggest that Model 2 hospitals participating in MJRLE increased their market share of MJRLE admissions. Together, these findings suggest that the increase in volume among BPCI hospitals was seemingly matched by an increase among non-BPCI hospitals. For additional detail on our analysis of changes in volume at the market level, please see the market dynamics chapter.

Exhibit 55: Average Number of Non-Fracture MJRLE Admissions per Physician per Quarter to BPCI and Non-BPCI Hospitals, by Performance Group, Q4 2011 – Q2 2015

Performance Group	Physicians (N)	Admissions to BPCI Hospitals		Admissions to Non-BPCI Hospitals		DiD estimate
		Baseline	Intervention	Baseline	Intervention	
Top Performers	411	3.9	5.6	0.6	0.7	1.6
Average Performers	656	5.2	6.3	0.4	0.4	1.1
Bottom Performers	361	3.0	3.3	0.9	1.2	0.0

Source: Lewin analysis of Medicare claims and enrollment data for admissions in Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q2 2015 (intervention period) for physicians admitting to BPCI and non-BPCI hospitals.

Note: This table reflects MJRLE admissions. Some MJRLE admissions may not be BPCI episodes.

Exhibit 56: Average Number of Fracture MJRLE Admissions per Physician per Quarter to BPCI and Non-BPCI Hospitals, by Performance Group, Q4 2011 – Q2 2015

Performance Group	Physicians (N)	Admissions to BPCI Hospitals		Admissions to Non-BPCI Hospitals		DiD estimate
		Baseline	Intervention	Baseline	Intervention	
Top Performers	411	0.5	0.6	0.1	0.1	0.1
Average Performers	656	0.5	0.6	0.1	0.1	0.1
Bottom Performers	361	0.5	0.7	0.1	0.1	0.1

Source: Lewin analysis of Medicare claims and enrollment data for admissions in Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q2 2015 (intervention period) for physicians admitting to BPCI and non-BPCI hospitals.

Note: This table reflects MJRLE admissions. Some MJRLE admissions may not be BPCI episodes.

4. Conclusions

Results from the first seven quarters indicate there was variation in financial performance, as measured by NPRA, within and across clinical episodes. Some strategies to reduce Medicare payments were common across hospitals with higher standardized NPRA for CHF, COPD, or MJRLE episodes, namely, discharging patients home with or without home health instead of to an institutional PAC setting and reducing the unplanned readmission rate. Among hospitals

participating in the MJRLE episode, high standardized NPRA was also correlated with a reduction in patient morbidity and an increase in the volume of non-fracture cases.

D. Model 2 Discussion

1. Participants

There were 422 acute care hospital (ACH) episode initiators (EIs) and 277 physician group practice (PGP) EIs under 215 awardees active in Model 2 during the first two years of the initiative. These participants initiated over 243,000 episodes, comprising about 85% of all Model 2, 3 and 4 episodes.

The composition of BPCI hospitals became more similar to non-participants as more entered the initiative during the second year, but differences between these two groups remain. Participants are still more likely to be larger. This is consistent with many interviewees citing the need for significant financial resources to successfully start and maintain participation. Participants also had nearly twice as many admissions for BPCI episode MS-DRGs compared to non-participants prior to joining. Moreover, the standardized Medicare allowed Part A payments for the hospitalization plus the 90-day post-discharge period were substantially higher for BPCI participants than non-participating providers, on average, for the clinical episodes they chose for participation. These characteristics suggest that hospitals evaluated their ability to achieve financial gains, the most common reason for joining BPCI, in deciding whether to participate.

Many BPCI participant interviewees also reported using consultants to advise them on Model and clinical episode choices. By year two of the initiative there were EIs in each of the 48 Model 2 clinical episodes. Hospital interviewees said they expected bundled payments to play a key role in future payment reforms and viewed BPCI as a learning opportunity. The majority of BPCI-participating hospitals had experience with care redesign and payment incentives that may have prepared them for the BPCI initiative. BPCI-participating PGPs, on the other hand, were less likely to have had prior experience with care redesign and payment incentives.

Interviews with participating hospital and PGP representatives indicated that an important strategy for improving the delivery of care under BPCI involved working with other providers throughout the full episode period. For hospitals, the focus was largely on post-acute care (PAC) providers. Interviewees from PGPs indicated having more control over the inpatient portion of the clinical episode as a main focus.

A number of participants stopped participating in some or all of their BPCI clinical episodes. About 14% of hospital and PGP EIs stopped participating in at least one clinical episode, while 27 hospitals (6%) and one PGP (0.4%) withdrew completely. Large financial requirements and financial risk were cited as the main reasons for withdrawals.

2. Impact of BPCI on Costs and Utilization

BPCI did not significantly affect changes in total Medicare standardized allowed payments from the baseline to intervention period relative to comparison providers for any of the 23 Model 2 ACH clinical episodes analyzed, with the exception of major joint replacement of the lower

extremity (MJRLE).[29] One reason for the lack of significant effects may be that a majority of EIs joined BPCI in the last two quarters of the current evaluation period (Q2 and Q3 of 2015). Thus, there may not have been enough time to see the impact on episode payments from any of their changes to care processes.

While measureable effects on total payments during the episode were limited, there were changes in patterns of payment by type of service across many of the clinical episodes that were consistent with changes in care delivery that participants aimed to employ. Namely, even though not statistically significant in most cases, SNF and IRF payments declined relative to the comparison group for the majority of clinical episode types. At the same time, HHA payments increased for most. These changes were driven by relative changes in utilization, as there was a shift in BPCI patient discharges following the anchor hospital stay away from institutional PAC settings and towards HHA care. Even for patients discharged to a SNF, the number of SNF days declined for BPCI patients relative to those in the comparison group in 16 of the 23 clinical episodes. The number of home health visits among BPCI patients who had at least one visit increased in 20 of the 23 clinical episodes.

Participants in three clinical episodes (COPD, MJRLE, and CHF) were studied more in-depth to better understand the variation in financial performance as measured by their net payment reconciliation amount (NPRA).[30] Participants with the largest NPRA were associated with larger reductions in the use of institutional PAC services. The top NPRA earners were also able to increase the proportion of patients discharged home without any type of PAC and reduce the rate of unplanned readmissions. The top earners, particularly for COPD and MJRLE, also tended to have higher proportions of patients receiving institutional PAC services prior to BPCI, which may have provided them more opportunity to reduce such care.

The one clinical episode that did exhibit a significant impact on total payments during the first two years of the initiative was MJRLE. Aside from reducing institutional PAC services, interviewees indicated patient education about discharge destination as a key to managing the episode. Patient education would be more impactful during planned admissions. Yet, participants were able to reduce per episode costs for fracture episodes, as well as elective non-fracture episodes.

3. Impact of BPCI on Quality

Various data sources were used to gather evidence on the impact of BPCI on the quality of care delivered to patients. In general, BPCI did not appear to have a systematic effect, either positive or negative, on the quality of care delivered.

There were few significant changes in mortality, unplanned readmissions, and emergency department visits from the baseline to intervention period. While more than half of the clinical episodes under Model 2 had positive point estimates for the change in emergency department use

[29] We were unable to evaluate PGP episodes due to inaccuracies in the list of participating physicians by PGP at the time of our analyses.

[30] Net payment reconciliation amount (NPRA) is the difference between the hospital's target amount and actual spending, across clinical episodes.

and unplanned readmission rates, the magnitudes tended to be small and they were statistically significant for only two clinical episodes.

There were also few significant changes in the improvement in activities of daily living (ADLs) as measured through patient assessments in HHA and SNF settings. These results were based on the approximately three-quarters of patients who received care long enough to have two assessments and may not be representative of all patients discharged to PAC. Taken together, these results suggest that across all Model 2 episodes, quality of care was generally maintained during the first two years of the BPCI initiative.

We might have expected to see reductions in the proportion of patients with improved ADLs during their PAC treatment, given that many participants were aiming to reduce institutional PAC referrals and the number of days patients used SNF care. Although the ADL results accounted for patient severity at PAC admission, the shorter SNF stay would result in a patient population with less time to recover. However, there was no strong relationship between changes in the number of SNF days and ADL outcomes across episodes. An analysis of the three episodes with the most volume (MJRLE, CHF, and sepsis) point to Model 2 EIs discharging a greater share of patients to highly rated SNFs and HHAs after they started participating in BPCI, which may have helped to avoid reductions in the quality of care.

MJRLE was one of the few clinical episodes that experienced a significant relative reduction in SNF utilization. At the same time, there were reduced proportions of patients improving on each of three SNF ADL measures among non-fracture episodes discharged to SNF. Given that MJRLE was associated with large NPRA and top earners were most successful in reducing institutional PAC use, it will be important to monitor the changes in ADL measures moving forward.

Patient survey data indicated that for the eight clinical episodes analyzed, BPCI and comparison respondents generally reported similar changes in functional status from before their hospitalization to after their episode as well as similar changes in health status and care experience. Where there were differences, few were consistent from one survey wave to the next. The two clinical episodes with the most consistent differences were MJRLE and sepsis. In the case of MJRLE, patients with BPCI episodes tended to have more positive effects across all three domains, including functional status. In contrast, sepsis patients with BPCI episodes reported more negative results than patients with episodes from comparison providers.

4. Unintended Effects

While the aim of bundled payments is to encourage providers to reduce costs through more efficient and higher quality care, providers may also respond to the incentives in unintended ways that potentially have adverse consequences. In particular, we investigated any increases in payments for services excluded from the bundle. We also evaluated whether participants selected a healthier, and therefore less costly, mix of patient. We paid particular attention to any evidence about patient selection for the planned non-fracture MJRLE episodes where patients are identified prior to their hospitalization for the procedure. Moving services outside of the bundle period and patient selection strategies may adversely affect patient care or increase total Medicare outlays, even if episode payments decline.

Among the 9 clinical episodes that are generally planned admissions, there were no indications of changes in the services delivered prior to the anchor hospitalization. However, for 8 out of the 10 episodes identified as chronic conditions, there were greater increases in payments in the 30 days after the end of the episode, relative to the comparison group (though the difference was statistically significant for only one clinical episode). More research will be needed to determine if those increases are due to participants postponing services or if the pattern reflects additional care management of the chronic conditions following a BPCI hospitalization, which may be indicative of better quality care.

For renal failure and gastrointestinal hemorrhage episodes, there was an increase in readmissions that were not included in the bundled payment, but not for readmissions that were included. These patterns will continue to be monitored as additional data becomes available.

Relative changes in patient characteristics between the baseline and intervention period indicate that for MJRLE non-fracture episodes BPCI patients were relatively healthier than comparison patients, from the baseline through the intervention period. Further, higher NPRA was correlated with healthier patients, as measured by a decline in the average number and intensity of comorbidities and lower patient age.

5. Conclusion

Through the first two years of the initiative, Model 2 of BPCI accounted for about 85% of all episodes in the three models evaluated here. While there has been a high level of participation in Model 2, the measurable impact of BPCI on cost and quality during this time has been limited.

The only episode to exhibit significant reductions in Medicare payments was MJRLE. The decrease was largely driven by reduced institutional PAC services. There did not appear to be a concomitant decline in quality. However, there was a relative decline in the proportion of patients treated by SNFs who received care long enough to have two completed SNF assessments that indicated an improvement in multiple functional status measures. Also, participants that did the best under Model 2 for MJRLE episodes, had a more severe mix of patients and higher than average institutional PAC use prior to BPCI. This raises questions about the potential for reducing payments among participants with a more typical mix of patients with average PAC use.

E. Clinical Episode Issue Briefs

1. Acute Myocardial Infarction (AMI)

a. Overview

Acute myocardial infarction (AMI), which occurs when the flow of blood to the heart is diminished, is a prevalent condition among Medicare beneficiaries. AMI was the sixth most common principal discharge diagnosis among Medicare fee-for-service beneficiaries hospitalized

in 2013.[31] In fiscal year 2014, over 133,000 Medicare discharges were classified in one of the three Medicare Severity-Diagnosis Related Groups (MS-DRGs) included in the AMI bundle for BPCI.[32]

Medicare beneficiaries received 2,684 AMI episodes of care in 78 BPCI-participating acute care hospitals in the first eight quarters of the BPCI initiative (from October 2013 through September 2015). We found no statistically significant change in total Medicare payments for the anchor hospitalization and the care provided during the following 90 days for BPCI episodes relative to comparison episodes. There is some indication that BPCI increased Medicare payments during the post-bundle period. Although such increases may raise concerns that BPCI providers are deferring services until after the bundle, this is likely not the case for AMI episodes as we only observed increases for the extended post-bundle periods. Analyses also indicate that the quality of care was maintained or improved for beneficiaries in episodes initiated by BPCI hospitals, as measured by readmission rates, emergency department visits, mortality, and functional status outcomes from patient assessments.

b. Participants

During the first eight quarters of the BPCI initiative, 78 hospitals (18% of all Model 2 BPCI hospitals) participated in the AMI episode, 75 of which chose 90-day episodes. There were 2,684 AMI episodes started in these hospitals during the BPCI initiative (approximately 2% of all episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the AMI results are based on an average of three quarters of experience. Approximately 26% of the hospitals began participating in AMI in Q3 2015; 72% joined in either Q2 or Q3 2015. Ten hospitals stopped participating in AMI within the first eight quarters of the initiative.

Compared to hospitals *not* participating in BPCI, hospitals that participated in the AMI episode were larger and more likely to be urban, non-profit facilities (see Exhibits 57a & 57b). Participating hospitals also had a lower share of Medicare days and greater teaching intensity as measured by the resident-to-bed ratio. Additionally, prior to joining BPCI, participating hospitals had more AMI discharges.

The 78 hospitals participating in AMI were more similar to all Model 2 BPCI hospitals than non-participating hospitals. There were no significant differences on key metrics between the 78 hospitals participating in AMI and all Model 2 BPCI hospitals.

Matched comparison hospitals were identified for 74 of the 78 BPCI-participating hospitals.[33] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals.

[31] Krumholz et al. (2015). Mortality, Hospitalizations, and Expenditures for the Medicare Population Aged 65 Years or Older, 1999–2013. *JAMA, 314*(4):355–365.

[32] Centers for Medicare and Medicaid Services. (2016). Notice of Proposed Rulemaking for Bundled Payment Models for High-Quality, Coordinated Cardiac and Hip Fracture Care. Retrieved from https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-07-25.html

[33] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a Propensity score model that considered market and hospital characteristics and baseline outcomes. Three hospitals were not

Exhibits 57a & 57b: Characteristics of BPCI-participating Hospitals with Acute Myocardial Infarction Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All AMI BPCI Hospitals (N=78)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched AMI BPCI Hospitals (N=74)		Comparison Hospitals (N=1,001+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	64	82%	1,594	57%*	321	77%	62	84%	811	81%
	Government	4	5%	542	20%*	32	8%	3	4%	40	4%
	For-Profit	10	13%	638	23%*	66	16%	9	12%	150	15%
Urban/Rural	Urban	73	94%	1,902	69%*	387	92%	69	93%	931	93%
Part of Chain	Yes	36	46%	1,469	53%	218	52%	36	49%	501	50%

Characteristic	All AMI BPCI Hospitals (N=78)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched AMI BPCI Hospitals (N=74)	Comparison Hospitals (N=1,001)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	294	175*	311	286	289
Medicare Days Percent	39%	42%*	39%	39%	38%
Resident-to-bed ratio	0.11	0.05*	0.12	0.11	0.12
Disproportionate Share Percent	28%	29%	27%	29%	29%
Number of AMI Discharges, 2011	65	34*	64	65	64
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, AMI, 2011	**$21,932**	**$21,620**	**$21,268**	**$22,041**	**$22,009**

PDP = post-discharge period; AMI = acute myocardial infarction

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. Non-participating hospitals are all hospitals that were not participating in BPCI from Q4 2013 – Q3 2015 and exclude Model 4 hospitals participating in BPCI during the first two years.

*Indicates the standardized mean difference between the two groups is greater than 0.2 (means are standardized by dividing them with the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All AMI BPCI" hospitals. "Comparison" hospitals are compared to "Matched AMI BPCI" hospitals.

**Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 596.

c. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. Among hospitals

included in the analysis because they either were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the AMI episode or had fewer than 5 discharges in 2011. One other participant was excluded because a match could not be found within the caliper.

participating in the AMI episode, there was a statistically significant decrease in the proportion of patients who were disabled relative to the change in the comparison group (p=0.05, see Exhibits 58a & 58b). There was also a relative decrease in the percentage of patients between the ages of 20 and 64 among BPCI hospitals (p=0.08), which is likely a reflection of the decrease in the proportion of patients who were disabled. There were no statistically significant changes in other beneficiary characteristics associated with health care spending, including the proportion of patients who were Medicaid eligible, the patients' number of HCC indicators, or measures of prior health care use.

Exhibits 58a & 58b: Characteristics of BPCI Beneficiaries and Matched Comparison Beneficiaries with a Hospitalization for an Acute Myocardial Infarction, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI AMI Hospital EIs				Matched Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	305	9%	183	8%	279	8%	209	9%	-1.8	-3.9	0.2	0.08
	65-79	1,241	35%	881	38%	1,318	37%	948	40%	-0.7	-4.2	2.9	0.71
	80+	1,978	56%	1,281	55%	1,927	55%	1,188	51%	2.5	-1.2	6.2	0.18
Gender	Female	1,913	54%	1,268	54%	1,921	55%	1,220	52%	2.3	-1.4	6.0	0.23
Medicaid and Disability	% Eligible for Medicaid	832	24%	466	20%	853	24%	515	22%	-1.5	-4.6	1.6	0.34
	% Disability, no ESRD	362	10%	207	9%	327	9%	235	10%	-2.2	-4.4	0.0	0.05
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	1,125	32%	724	31%	1,035	29%	649	28%	0.6	-2.7	4.0	0.71
	Emergency Room Admission	1,003	28%	712	30%	1,083	31%	762	32%	0.1	-3.3	3.5	0.94
	Home Health	754	21%	453	19%	702	20%	464	20%	-1.9	-4.9	1.0	0.20
	Inpatient Rehabilitation Facility	54	2%	37	2%	35	1%	22	1%	0.1	-0.7	0.9	0.81
	Skilled Nursing Facility	368	10%	251	11%	324	9%	211	9%	0.5	-1.7	2.7	0.68
	Psychiatric Hospital	31	1%	19	1%	30	1%	13	1%	0.2	-0.4	0.9	0.50
	Long-term Care Hospital	17	0%	7	0%	14	0%	9	0%	-0.2	-0.6	0.3	0.47
	Institutional Nursing Facility*	517	15%	307	13%	463	13%	273	12%	-0.1	-2.6	2.4	0.95
	No Institutional Care	2,324	66%	1,592	68%	2,409	68%	1,635	70%	0.6	-2.9	4.0	0.74
	No Post-acute Care	1,660	47%	1,138	49%	1,639	47%	1,112	47%	0.5	-3.2	4.2	0.79

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	3.01	2.21	3.00	2.21	2.87	2.09	2.86	2.18	-0.01	.0.17	0.15	0.94

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Note: "Count" represent the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

*Institutional Nursing Facility utilization is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

d. Medicare Payments

The change in total Medicare payments for the anchor hospitalization and all care delivered during the 90-day post-discharge period (PDP) was not statistically different between episodes initiated by BPCI and comparison hospitals (see Exhibit 59). Episode payments declined 6% for episodes initiated by BPCI hospitals and declined 7% for episodes initiated by comparison hospitals from the baseline to the intervention period.

Exhibit 59: Impact of BPCI on Total Payments for Inpatient Hospitalization and 90-day PDP, AMI Episodes, Model 2, Q4 2011-Q3 2015 ($2015)

	BPCI (N=2,311)		Comparison (N=2,308)		DiD estimate			
Measure	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Total allowed payment amount, IP and 90-day PDP	$27,850	$26,092	$28,199	$26,341	$101	-$1,633	$1,834	0.91

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison hospitals.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

Although there was no statistically significant change in total payments, payments for the inpatient hospitalization, which represent approximately one-third of total payments for 90-day episodes, decreased by $461 ($2015) for BPCI hospitals relative to the comparison group (p=0.07, see Exhibit 60).[34] Payments for the inpatient hospitalization decreased from $10,201 to $8,907 for BPCI hospitals from the baseline to the intervention period and decreased less, from $9,887 to $9,054, for comparison hospitals.

The relative decrease in payments for the inpatient stay is explained by a difference between BPCI and comparison hospitals in the proportion of episodes involving transfers, either to other acute care hospitals or to post-acute providers, and episodes eligible for an outlier payment, all of which can affect Medicare payments for inpatient hospital stays (see Exhibit 61).[35, 36] Our payment measure for the inpatient stay can include payments for the initial hospital stay as well as payments to any subsequent hospitals in the event of a transfer. Therefore, a change in the proportion of episodes in which a beneficiary was transferred to another acute care hospital would affect payments for the inpatient hospital stay. Additionally, a change in the proportion of short stay episodes involving transfers to post-acute care providers would affect payments for the inpatient stay because Medicare provides hospitals with a payment lower than the full MS-DRG payment in such cases. When we exclude transfers and outlier episodes from our analysis, the change in payments for the anchor hospitalization for BPCI hospitals relative to the comparison group increases from -$461 (see Exhibit 60) to $16 and is no longer statistically significant (p=0.75).

[34] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

[35] 42 CFR § 412.4 (2011)

[36] 42 CFR § 412.80 - § 412.86 (2011)

Exhibit 60: Impact of BPCI on Payments for Inpatient Hospitalization, AMI Episodes, Model 2, Q4 2011-Q3 2015 ($2015)

Measure	BPCI (N=2,327)		Comparison (N=2,320)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Total allowed payment amount, inpatient stay	$10,201	$8,907	$9,887	$9,054	-$461	-$962	$39	0.07

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison hospitals.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015.

Exhibit 61: Proportion of Outlier and Transfer Cases at BPCI and Comparison Hospitals, AMI, Model 2, Q4 2011-Q3 2015

Type of Hospital Stay	BPCI (N=2,327)		Comparison (N=2,320)		DiD estimate
	Baseline	Intervention	Baseline	Intervention	
Outlier	1.6%	1.0%	1.6%	1.8%	-0.8
Transfers to another hospital	12.3%	10.6%	11.6%	11.5%	-1.6
Short stay transfers to post-acute care	8.3%	8.4%	8.2%	9.9%	-1.6

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison hospitals.

We observed statistically significant increases in Medicare Part A and B payments during two of the post-bundle measurement periods for BPCI episodes relative to comparison episodes. Increases in payments in the post-bundle periods may raise concerns that BPCI providers are deferring services to the period following the bundle. Payments for such deferred services would not be considered when assessing payments for an episode of care relative to a provider's target price, and deferring services until after the bundle may make it easier for providers to reduce their payments below the target amount. For AMI episodes, though, we only observed increases in payments for the extended post-bundle periods and not for the time periods closer to the end of the bundle, so the increases may not be indicative of providers deferring services in this instance. As displayed in Exhibit 62, Medicare Part A and B payments for BPCI episodes increased by $3,241 during days 1 through 120 of the post-bundle period and by $4,018 during days 1 through 180 of the post-bundle period relative to the comparison group ($p<0.01$). In both instances, payments increased for BPCI episodes and decreased for the comparison group from the baseline to the intervention periods. For the measurement periods closer to the end of the bundle (days 1 through 30 and 1 through 90 of the post-bundle period), though, the change in payments was not statistically significant. This suggests that the relative increases in payments during days 1 through 120 and 1 through 180 of the post-bundle period are not indicative of providers deferring services, as we would anticipate that deferred services would be provided closer to the end of the bundle.

We investigated if the increases in post-bundle payments could be due to increased use of services intended to reduce the risk of future cardiac events, such as cardiac rehabilitation services. Cardiac rehabilitation services are associated with long-term improvements in patient outcomes, but have historically been underutilized. The increases in post-bundle payments observed for BPCI patients, however, were driven by increases in Part A payments. This suggests that the increases are not a result of increased use of cardiac rehabilitation services, which would be covered under Part B.

Exhibit 62: Impact of BPCI on Total Payments during the Post-bundle Period, AMI Episodes, Model 2, Q4 2011-Q3 2015 ($2015)

Measure	Intervention Episodes (N)		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
1-30 days post-bundle	1,499	1,508	$3,171	$3,097	$3,337	$3,361	-$99	-$736	$538	0.76
1-90 days post-bundle	1,179	1,203	$8,822	$9,548	$10,147	$9,737	$1,136	-$374	$2,646	0.14
1-120 days post-bundle	897	897	$11,082	$12,648	$13,697	$12,021	$3,241*	$880	$5,603	<0.01
1-180 days post-bundle	852	848	$15,732	$17,908	$18,686	$16,843	$4,018*	$1,007	$7,029	<0.01

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison hospitals.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015 for 30 days post-bundle, Q4 2013 through Q2 2015 for 90 days post-bundle, and Q4 2013 through Q1 2015 for 120 and 180 days post-bundle. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

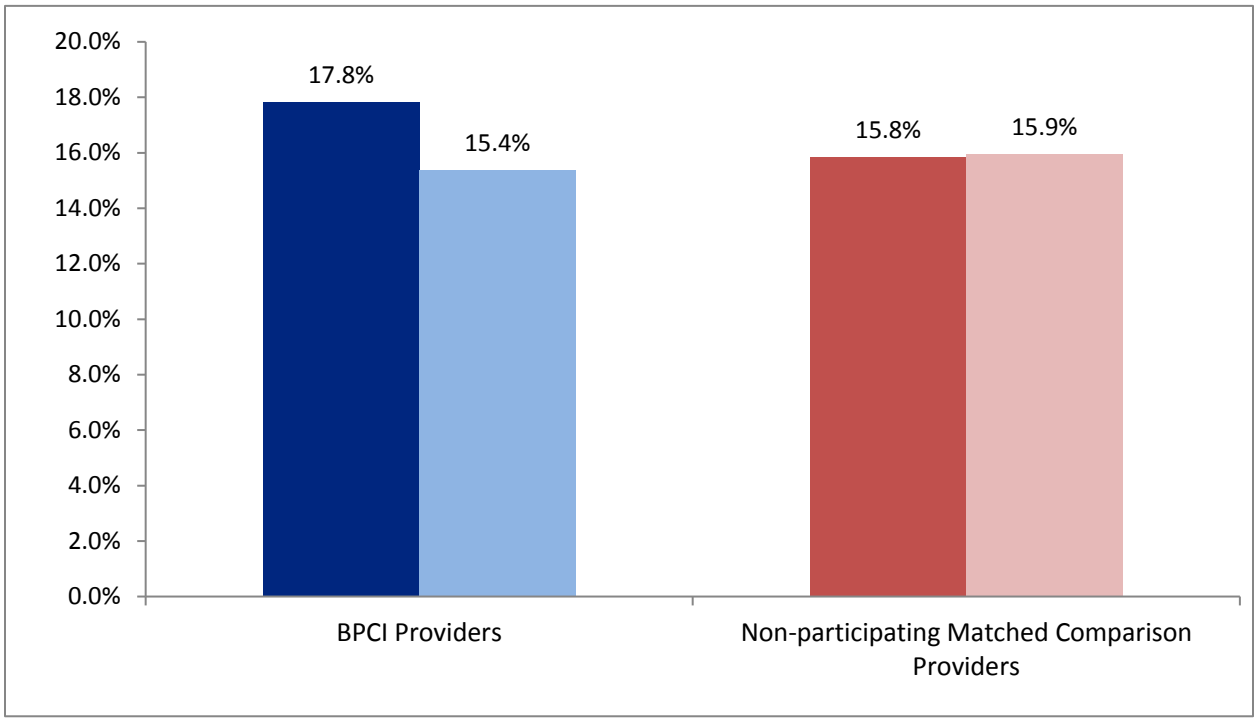
*Denotes statistical significance at the 5% level

e. Quality of Care

Quality of care was maintained or improved for beneficiaries in AMI episodes initiated by BPCI hospitals relative to beneficiaries treated by comparison hospitals. The mortality rate during the 90-day post-discharge period decreased by 2.5 percentage points from the baseline to the intervention period for patients treated by BPCI providers relative to the comparison group (p=0.05). BPCI hospitals, though, had higher baseline mortality rates and thus may have had more opportunity to improve quality. As shown in Exhibit 63, the mortality rate during the 90-day post-discharge period decreased from 17.8% to 15.4% for episodes initiated by BPCI hospitals, whereas the rates for comparison episodes remained stable at 15.8% and 15.9%.

From the baseline to the intervention period, the change in emergency department visits and the unplanned readmission rate following the anchor stay for AMI episodes was not statistically different between BPCI-participating and comparison hospitals. There were also no significant differences in measures of functional improvement as measured by patient assessments among beneficiaries who received post-acute care.

Exhibit 63: All-cause Mortality Rate, 90-day Post-Discharge Period, Matched BPCI and Comparison Providers



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison hospitals.

Notes: Results are risk adjusted. The intervention period is from the fourth quarter of 2013 through the third quarter of 2015. AMI = acute myocardial infarction.

f. Conclusions

The BPCI initiative did not have a statistically significant impact on total Medicare payments for AMI episodes. Additionally, based on the claim-based measures and assessment data used in the analysis, beneficiaries treated by BPCI providers experienced the same or improved level of

quality of care as patients treated by comparison hospitals. We did observe a relative increase in Medicare payments for services provided following the bundle. As we only observed increases for care provided during the 120 and 180 days following the end of the bundle, and not for any of the measurement periods closer to the end of the bundle, the increases are likely not indicative of providers deferring services to the post-bundle period. Additional research on the services provided following the bundle is needed to better understand the reasons for the relative increases among BPCI providers.

2. Cardiac Valve

Medicare beneficiaries received 2,541 cardiac valve episodes of care in 28 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015).[39] While there were no statistically significant changes in Medicare payments for the initial hospital stay and services during the subsequent 90-day post-discharge period (PDP),[40] there were changes in post-acute care (PAC) use during the 90-day PDP. The share of BPCI patients discharged to PAC settings increased, but the share of BPCI patients discharged to institutional PAC settings (skilled nursing facilities (SNF), inpatient rehabilitation facilities (IRF), and long term care hospitals (LTCH)) out of all PAC users declined relative to a matched comparison group, resulting in an increase in the share of BPCI patients discharged to home with home health agency (HHA) services. At the same time, there were no statistically significant changes in the quality of care among BPCI patients, as measured by readmission rates, emergency department visits, mortality rates, and measures of functional status from patient assessments, with the exception of one decline in one of five measures of functional status among BPCI patients discharged to HHA.

During this period, there was an increase in the use of minimally invasive or endoscopic surgical procedures, such as transcatheter aortic valve replacement (TAVR), for patients in the cardiac valve episode. BPCI participants were more likely than the comparison group to use these newer procedures. The share of cardiac valve patients treated using minimally invasive procedures increased by 23.4 percentage points more for the BPCI patients than the comparison group from the baseline through the intervention period. However, it is not clear how this change affected episode costs, quality outcomes, or PAC use, or why they were more prevalent in the BPCI-participating hospitals than in the matched comparison group.

a. Participants

During the first eight quarters of the BPCI initiative, 28 hospitals (7% of all Model 2 BPCI hospitals) participated in the cardiac valve episode, 82% of which chose 90-day episodes. There were 2,541 cardiac valve episodes initiated by these providers during the BPCI initiative (approximately 2% of episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the cardiac valve results are based on an average

[39] The cardiac valve episode under BPCI originally encompassed Medicare Severity-Diagnosis Related Groups (MS-DRGs) 216-221. In fiscal year 2015, MS-DRGs 266 and 267 were added as a result of changes in the grouping logic used by the MS-DRG system.

[40] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

of four quarters of experience. Approximately 33% of the hospitals began participating in cardiac valve in Q3 2015, and 44% joined in either Q2 or Q3 2015. Four hospitals stopped participating in cardiac valve within the first eight quarters of the initiative.

Interviewees that participated in BPCI quarterly interviews were asked about the choice of the cardiac valve episode. One interviewee selected the episode because of their strong cardiac team, which facilitated high patient volume, high success rates, and low complication rates. Another interviewee decided against joining the cardiac valve episode because they did not see opportunities to influence costs. They explained that cardiac valve patients go through a process that does not vary and, therefore, the hospital cannot control costs throughout the 90-day PDP.

Compared with hospitals that did not participate in BPCI, hospitals that participated in the cardiac valve episode were more likely to be non-profit, urban, larger, and part of a chain. Participating hospitals also had a greater teaching intensity as measured by the resident-to-bed ratio and had a higher number of cardiac valve discharges prior to joining BPCI (see Exhibits 64a and 64b). Matched comparison hospitals were identified for 18 of the 28 BPCI-participating hospitals in the sample.[41] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals.

[41] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. Four participants were not included in the analysis because there was not a match within the caliper. Three participants were not included because they were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the cardiac valve episode, and three additional participants were not included because they had fewer than five relevant discharges in 2011 or 2012.

Exhibits 64a and 64b: Characteristics of BPCI-participating Hospitals with Cardiac Valve Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Cardiac Valve BPCI Hospitals (N=28)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)		Matched Cardiac Valve BPCI Hospitals (N=18)		Comparison Hospitals (N=158)+	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	22	79%	1,594	57%*	321	77%	14	78%	125	79%
	Government	4	14%	542	20%	32	8%*	3	17%	25	16%
	For-Profit	2	7%	638	23%*	66	16%*	1	6%	8	5%
Urban/Rural	Urban	27	96%	1,902	69%*	387	92%	17	94%	146	94%
Part of Chain	Yes	19	68%	1,469	53%*	218	52%*	12	67%	112	71%

Characteristic	All Cardiac Valve BPCI Hospitals (N=28)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)	Matched Cardiac Valve BPCI Hospitals (N=18)	Comparison Hospitals (N=158)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	500	175*	311*	570	586
Medicare Days Percent	37%	42%*	39%	34%	33%
Resident-to-bed ratio	0.24	0.05*	0.12*	0.27	0.27
Disproportionate Share Percent	28%	29%	27%	27%	29%
Number of Cardiac Valve Discharges, 2011	69	11*	54	99	108
Standardized Part A Allowed Payment inpatient stay plus 90-day post-discharge period, Cardiac Valve episodes, 2011	**$50,273**	**$50,914**	**$49,884**	**$51,696**	**$51,711**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI during the first year.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All cardiac valve BPCI" hospitals. "Comparison" hospitals are compared to "Matched cardiac valve BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 124.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality controls for changes in patient characteristics, it does not directly examine changes in patient mix. A comparison of the cardiac valve patient population of BPCI participants from baseline to intervention relative to the same time periods for the cardiac valve patient population of the comparison providers does not suggest that BPCI providers treated a healthier patient population during the intervention period (see Exhibits 65a and 65b).

Exhibits 65a and 65b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Cardiac Valve, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	103	6%	104	5%	94	5%	107	5%	-0.7	-2.7	1.3	0.52
	65-79	961	56%	1,122	53%	952	55%	1,160	54%	-2.3	-6.8	2.2	0.31
	80+	663	38%	909	43%	681	39%	868	41%	3.0	-1.4	7.4	0.19
Gender	Female	791	46%	968	45%	763	44%	881	41%	2.5	-2.0	6.9	0.28
Medicaid and Disability	% Eligible for Medicaid	198	11%	206	10%	192	11%	207	10%	-0.4	-3.1	2.3	0.78
	% Disability, no ESRD	135	8%	127	6%	113	7%	123	6%	-1.1	-3.3	1.1	0.33
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	625	36%	709	33%	575	33%	704	33%	-2.7	-6.9	1.6	0.22
	Emergency Room Admission	346	20%	477	22%	405	23%	580	27%	-1.4	-5.2	2.4	0.47
	Home Health	231	13%	263	12%	203	12%	275	13%	-2.2	-5.2	0.8	0.15
	Inpatient Rehabilitation Facility	17	1%	15	1%	16	1%	13	1%	0.0	-0.8	0.8	0.93
	Skilled Nursing Facility	88	5%	98	5%	70	4%	100	5%	-1.1	-3.0	0.7	0.24
	Psychiatric Hospital	4	0%	2	0%	5	0%	4	0%	0.0	-0.4	0.4	0.86
	Long-term Care Hospital	4	0%	2	0%	2	0%	4	0%	-0.2	-0.6	0.1	0.25
	Institutional Nursing Facility**	86	5%	93	4%	71	4%	97	5%	-1.1	-2.9	0.8	0.27
	No Institutional Care	1,087	63%	1,416	66%	1,143	66%	1,414	66%	3.3	-0.9	7.6	0.13
	No Post-acute Care	888	51%	1,125	53%	918	53%	1,089	51%	3.4	-1.1	7.9	0.13

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.58	2.07	2.57	2.12	2.42	2.05	2.47	2.00	-0.07	-0.25	0.12	0.50

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare Payments and Post-acute Care Utilization

Total Medicare payments ($2015) for the inpatient hospitalization plus all care delivered during the 90-day post-discharge period did not change for cardiac valve episodes initiated in BPCI-participating hospitals relative to episodes from a matched comparison group of hospitals not participating in BPCI (see Exhibit 66). Total payments declined by 5% for BPCI episodes and by 6% for the matched comparison group from before the initiative.

Exhibit 66: Impact of BPCI on Total Standardized Allowed Payment, Inpatient Hospitalization plus 90-day Post-discharge Period ($2015) for Cardiac Valve Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=2,086)		Comparison (N=2,089)		DiD			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-Value
Total Payment, Inpatient Hospitalization plus 90-day post-discharge period	$67,200	$63,745	$68,708	$64,334	$919	-$1,191	$3,029	0.39

DiD = Difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Although there was no statistically significant change in total Medicare payments for BPCI episodes, there were shifts in PAC use. The proportion of Medicare cardiac valve patients treated at BPCI-participating hospitals discharged to any PAC provider increased 6.5 percentage points (p=0.04) relative to the comparison group (see Exhibit 67).[42,43] The proportion of cardiac valve patients discharged to PAC following their stay in a BPCI-participating hospital increased from 73.4% to 78.8% from the baseline to intervention periods. The proportion of patients discharged to PAC following their stay in a matched comparison hospital remained virtually steady around 80%.

At the same time, the proportion of patients treated at BPCI hospitals discharged to an institutional PAC setting among PAC users declined 10.9 percentage points (p=0.10), from 52.8% to 39.1% , between the baseline and intervention period.[44] The proportion of cardiac valve patients discharged to institutional PAC among PAC users at matched comparison hospitals decreased less, from 52.1% to 49.4%.

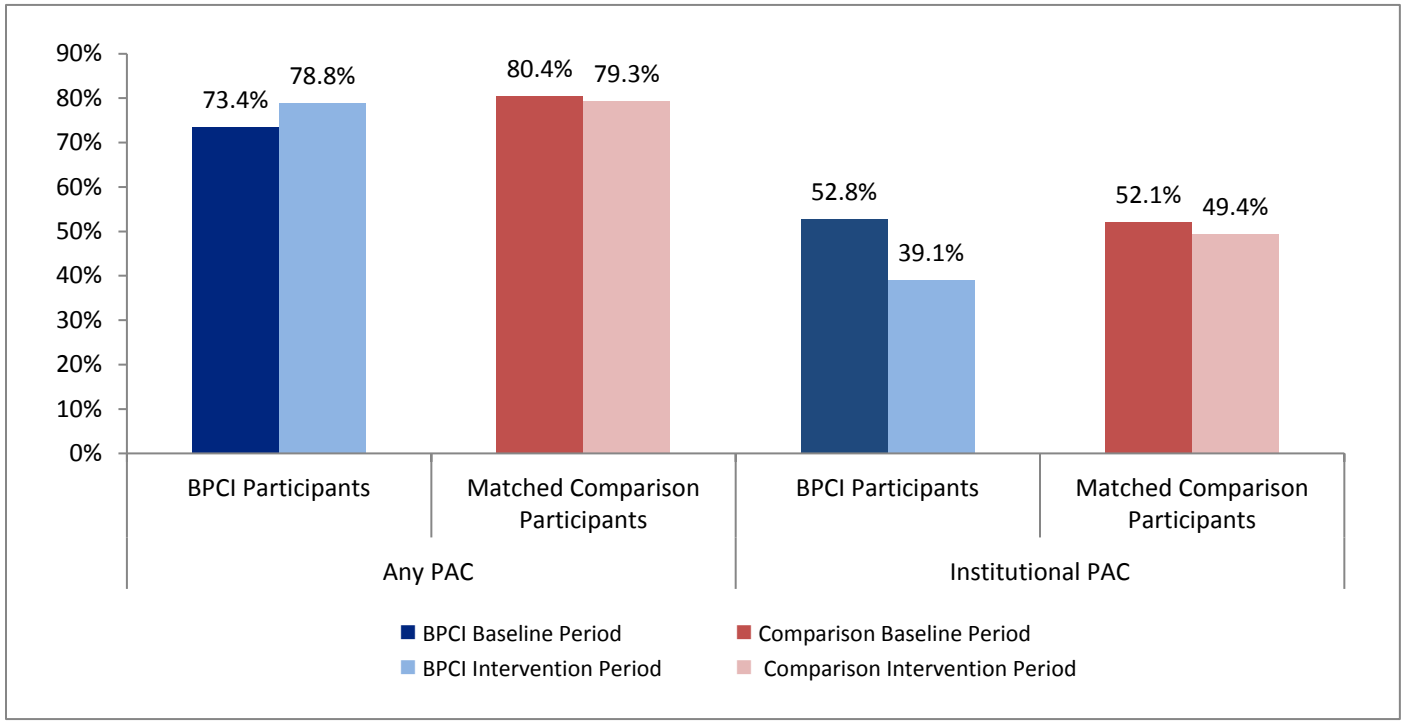
These findings indicate that a larger share of BPCI patients were discharged home with HHA services relative to the comparison group. Although not statistically significant, this is consistent with a relative increase in HHA payments (p=0.19).

[42] PAC settings include SNFs, IRFs, LTCHs, and HHAs.

[43] These estimates are developed using a Difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

[44] Institutional PAC settings include SNFs, IRFs, and LTCHs. PAC users include beneficiaries discharged to any PAC setting (SNF, IRF, LTCH, and HHA).

Exhibit 67: Impact of BPCI on the Proportion of Cardiac Valve Episodes Discharged to any Post-Acute Care Setting and the Proportion of Cardiac Valve Episodes Discharged to Institutional PAC Settings among PAC users, Model 2, Q4 2011-Q3 2015



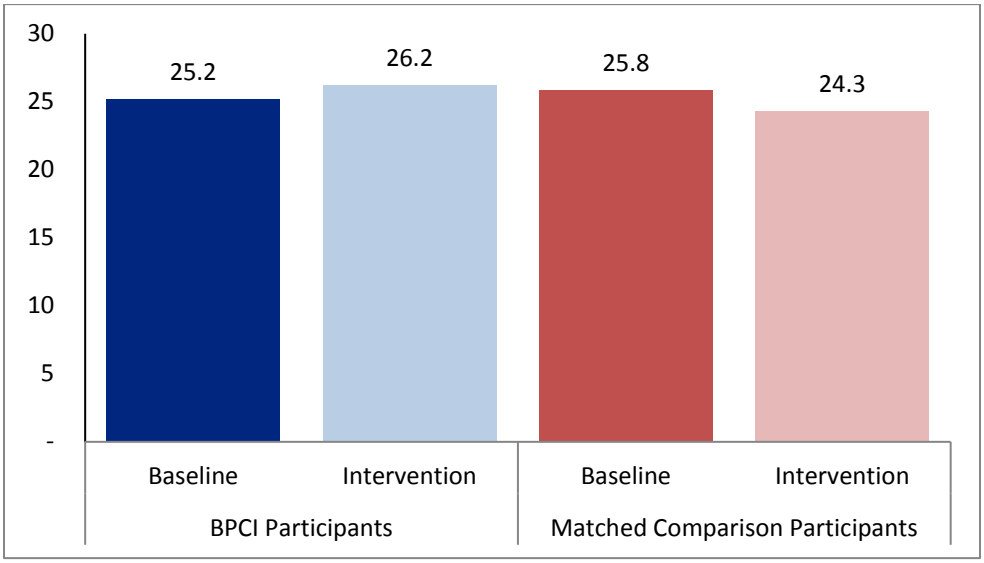
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Notes: Post-acute care (PAC) settings includes skilled nursing facility (SNF), inpatient rehabilitation facility (IRF), long term care hospital (LTCH), and home health agencies (HHA). Institutional PAC includes SNF, IRF, and LTCH.

In addition, there was an increase in institutional PAC days. The number of days spent in an institutional PAC setting among all institutional PAC users increased by 3.9 days for BPCI episodes relative to the comparison group ($p<0.01$). Among BPCI episodes discharged to a PAC setting, institutional PAC days increased from 22.4 to 24.3 between the baseline and intervention period, and they decreased from 23.9 to 21.9 days among comparison episodes. The number of SNF days among SNF users increased by 2.6 for BPCI episodes relative to the comparison group ($p=0.08$) (see Exhibit 68). SNF days increased from 25.1 to 26.2 for BPCI episodes and decreased from 25.8 to 24.3 for comparison group episodes. The smaller share of patients discharged to institutional PAC and the higher number of institutional PAC days (particularly SNF days) suggests that less intensive or severe BPCI patients were less likely to receive institutional PAC during the intervention, resulting in more intensive or severe patients using institutional PAC.

Another indicator of more resource-intensive patients using SNF is the relative increase in the SNF nursing case-mix index. The nursing case-mix index of patients discharged to SNF increased from 1.33 to 1.39 for BPCI patients from the baseline to the intervention period, while it remained relatively stable at 1.26 and 1.24 for cardiac valve patients in the comparison group (significant at the 5% level).

Exhibit 68: Impact of BPCI on the Average Number of Days in a Skilled Nursing Facility among Cardiac Valve Episodes with at Least One SNF Day, Model 2, Q4 2011-Q3 2015



Note: This measure includes number of days spent in a skilled nursing facility (SNF) among all episodes with at least one day spent in a SNF.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Corroborating these findings, BPCI interviewees stated that fewer patients received institutional PAC.[45] One interviewee reported they have set the expectation that every valve patient will be discharged home with HHA services rather than to an IRF, as they did prior to BPCI, and only the most ill patients would be discharged to IRF or SNF, which are higher cost than HHAs. Two interviewees reported that one reason for the reduced need for institutional PAC was the development of the minimally invasive procedures, discussed further below.

Despite significant shifts in PAC utilization, there were no statistically significant changes in total payments during the PAC phase. Declines in IRF payments were offset by increases in SNF and HHA payments and therefore did not result in a statistically significant change in total payments. There was a statistically significant decrease of $32 in Part B payments for therapy services during the 90-day PDP for BPCI episodes (p=0.01). Because Part B therapy payments represent 0.1% of total episode payments, this decline did not result in a statistically significant decline in total payments.

d. Quality of Care

There was no statistically significant change in the quality of care for Medicare patients receiving cardiac valve surgery at BPCI-participating hospitals compared to patients at comparison hospitals as measured by unplanned readmission rates, emergency department visits, mortality rates, and measures of functional improvement from patient assessments, with the exception of one decline in the share of BPCI patients discharged home with HHA services who experienced an improvement in lower body dressing.

[45] See Methods section for more information on the BPCI participant site visits and interviews.

The share of BPCI patients discharged home with HHA services who experienced an improvement in lower body dressing decreased by 3.8 percentage points relative to the comparison group (p=0.09) (see Exhibit 69). Though not significant, there were also declines in the other HHA functional status measures.

Exhibit 69: Impact of BPCI on Functional Improvements among Patients Discharged Home with Home Health Agency Services for Cardiac Valve Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=864)		Comparison (N=714)		DiD			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-Value
Improved lower body dressing	90.9%	89.7%	88.9%	91.5%	-3.8	-8.2	0.6	0.09
Improved upper body dressing	91.0%	89.7%	89.3%	90.8%	-2.7	-6.7	1.2	0.18
Improved bathing	88.6%	87.6%	87.5%	89.7%	-3.3	-8.5	2.0	0.23
Improved ambulation	86.9%	86.2%	87.1%	87.7%	-1.2	-5.5	3.1	0.57
Improved bed transferring	83.7%	80.6%	86.1%	83.1%	-0.1	-6.0	5.8	0.97

DiD = Difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers

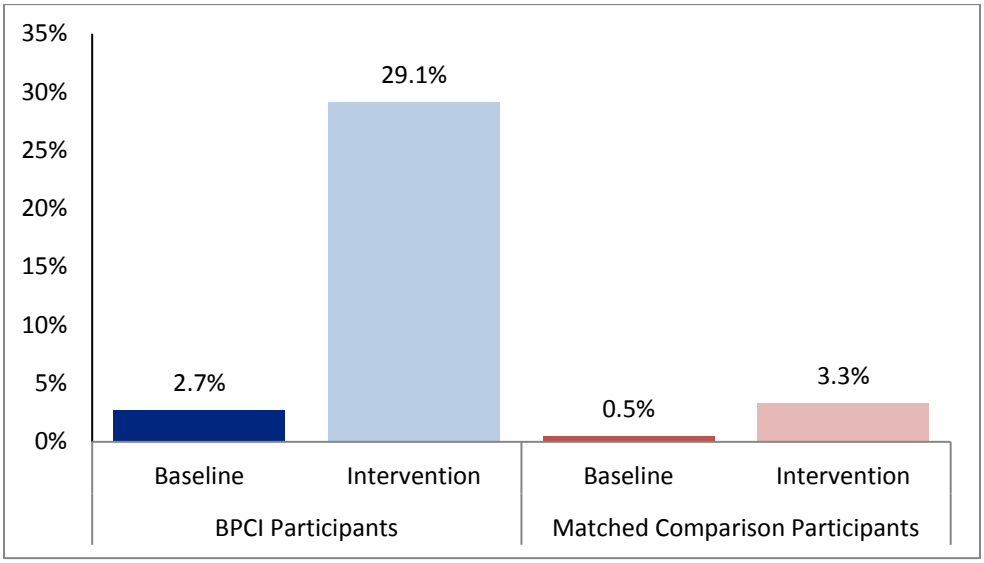
Note: "Improvement" means improvement or staying in the best possible status.

e. Increases in Minimally Invasive Procedures

BPCI hospitals treated a larger share of their patients using minimally invasive procedures during the baseline, and they increased this share by 23.4 percentage points more than the comparison group during the intervention (see Exhibit 70).[1] We do not know why BPCI hospitals were more likely to implement these new procedures than comparison providers. However, as noted elsewhere in this annual report, many BPCI hospital representatives we interviewed indicated that they decided to participate in BPCI because they viewed episode-based payments as the wave of the future and that they wanted the opportunity to learn from the initiative. These responses may suggest that BPCI participants are forward thinking, a characteristic that may apply to clinical innovations as well as payment policies, although our does not explicate the relationship between performing minimally invasive valve procedures and BPCI.

[1] There are currently endoscopic procedures available to treat three out of the heart's four valves that are included in the cardiac valve episode (MS-DRGs 216-221 and 266-267). These are transcatheter aortic valve replacement (TAVR) (also known as percutaneous aortic valve replacement (PAVR) or transcatheter aortic valve implantation (TAVI)), and transcatheter pulmonary valve implantation (TPVI). The minimally invasive procedures are identified in Medicare claims and enrollment data using Healthcare Common Procedure Coding System (HCPCS) codes.

Exhibit 70: Impact of BPCI on the Share of Cardiac Valve Patients treated with a Minimally Invasive Procedure, Model 2, Q4 2011-Q3 2015



Note: Minimally invasive procedures captured in the cardiac valve episode (MS-DRGs 216-221 and 266-267) include transcatheter aortic valve replacement (TAVR) (also known as percutaneous aortic valve replacement (PAVR) or transcatheter aortic valve implantation (TAVI)), and transcatheter pulmonary valve implantation (TPVI). These procedures are identified in claims data using Healthcare Common Procedure Coding System (HCPCS) codes.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Representatives of BPCI-participating hospitals who we interviewed indicated that changes in cardiac valve treatment, specifically, the use of minimally invasive procedures, have affected their ability to reduce internal costs.[46] Minimally invasive procedures allow treatment of the valve without open heart surgery. As a result, they present a viable option for patients with high surgical risk who are not candidates for the traditional open surgical approach. In addition, interviewees indicated that the cost of the device was a large share of the hospital's internal cost of providing the minimally invasive procedure. One interviewee noted that they were able to reduce Medicare payments through lower readmissions and lower PAC use during the post-discharge period for these patients, but the tradeoff with the higher cost of the device itself made it hard for them to save money internally. They expressed hope that internal hospital costs would decline as they continued to increase minimally invasive procedures. It is not clear whether being able to provide cardiac valve surgery to patients who were not candidates for the open surgical procedure has increased episode costs. Nor is it known if the use of minimally invasive procedures has affected quality outcomes or PAC utilization.[47]

[46] Hospitals are paid based on the MS-DRG of the episode and therefore receive the same payment for all procedures regardless of whether they are minimally invasive.

[47] In Q4 2015, CMS introduced new MS-DRGs for minimally invasive procedures, which will allow us to examine these questions for episodes initiated after that time.

f. Conclusions

During the first eight quarters of the BPCI initiative, no statistically significant changes were observed in total payments or payments during the 90-day PDP. BPCI hospitals did make substantial changes in the use of PAC for their cardiac valve patients, but there were no corresponding changes in payments. Because minimally invasive cardiac value procedures are still a relatively new innovation, it may take more time to better understand and distinguish the impact of BPCI from the impact of these procedures.

3. Chronic Obstructive Pulmonary Disease (COPD), Bronchitis and Asthma

Medicare beneficiaries received 8,770 chronic obstructive pulmonary disease (COPD), bronchitis and asthma[48] episodes of care in 110 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015).[49] BPCI had no statistically significant impact on total Medicare payments, post-acute care (PAC) utilization, or the quality of care.[50] Although in the aggregate there was no statistically significant change in these measures for COPD episodes, there was variation in the financial performance of hospitals for this episode, as measured by the Net Payment Reconciliation Amount (NPRA), that is, the difference between the hospital's target amount and actual spending. Top performers had a mean average NPRA of $1,752 per COPD episode, whereas bottom performers had a mean average NPRA of - $2,433 per episode. Top performers reduced both the share of beneficiaries discharged to institutional PAC and unplanned readmissions, while bottom performers did not.

a. Participants

During the first eight quarters of the BPCI initiative, 110 hospitals (26% of all Model 2 BPCI hospitals) participated in the COPD episode and the majority chose 90-day episodes.[51] There were 8,770 COPD episodes initiated by these providers during this period (approximately 6% of episodes across the 48 BPCI clinical episodes). The COPD results are based on an average of three quarters of experience because approximately 23% of the hospitals began participating in COPD in Q3 2015; 69% joined in either Q2 or Q3 2015. Furthermore, of the 110 hospitals, 14 stopped participating in COPD within the first eight quarters of the initiative.

Compared to hospitals *not* participating in BPCI, hospitals that participated in the COPD episode were larger and more likely to be non-profit hospitals, with higher resident-bed ratios, and with a lower share of Medicare days (see Exhibits 71a & 71b). COPD hospitals were also more likely to be located in urban areas. Prior to joining BPCI, participating hospitals had more COPD discharges

[48] Although this clinical episode consists of COPD, Bronchitis and Asthma diseases, for the remainder of the brief we will refer to all of them as COPD.

[49] As stated below, because participants were allowed to join BPCI over an extended period, these participants had an average of 3 quarters of experience in Model 2.

[50] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and expressed in 2015 dollars. These results do not take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

[51] 108 hospitals chose 90-day episodes and 2 hospitals chose 30-day episodes.

and had higher average 2011 standardized Part A payments for COPD episodes during the anchor hospitalization plus 90-days post-discharge period than non-participants ($15,330 vs. $13,772).

Matched comparison hospitals were identified for 108 of the 110 BPCI-participating hospitals in the sample.[52] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals with respect to the characteristics we considered.

Exhibits 71a & 71b: Characteristics of BPCI-participating Hospitals with COPD Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospital EIs, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All COPD BPCI Hospitals (N=110)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched COPD BPCI Hospitals (N=108)		Comparison Hospitals (N=1,537+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	87	79%	1,594	57%*	321	77%	86	80%	1,226	80%
	Government	4	4%	542	20%*	32	8%	4	4%	58	4%
	For-Profit	19	17%	638	23%	66	16%	18	17%	253	16%
Urban/Rural	Urban	103	94%	1,902	69%*	387	92%	101	94%	1,444	94%
Part of Chain	Yes	59	54%	1,469	53%	218	52%	59	55%	883	57%

Characteristic	All COPD BPCI Hospitals (N=110)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched COPD BPCI Hospitals (N=108)	Comparison Hospitals (N=1,537+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	300	175*	311	303	313
Medicare Days Percent	38%	42%*	39%	38%	38%
Resident-to-bed ratio	0.10	0.05*	0.12	0.10	0.11
Disproportionate Share Percent	30%	29%	27%*	30%	30%
Number of COPD Discharges, 2011	189	118*	191	191	196
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, COPD, 2011	**$15,330**	**$13,772***	**$14,317***	**$15,330**	**$15,233**

PDP=post-discharge period; COPD=Chronic obstructive pulmonary disease, bronchitis, asthma.

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All COPD BPCI" hospitals. "Comparison" hospitals are compared to "Matched COPD BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 686.

[52] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were two hospitals that were not included in the analysis because they were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the COPD episode.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. There was no statistically significant difference in the change in patients by age group or in patients' number of HCC indicators, relative to the comparison group. For BPCI hospitals, there was, however, a statistically significant increase in the proportion of the population that was Medicaid eligible ($p=0.02$) and the population that was disabled ($p=0.06$), relative to the comparison group.[53] With respect to prior health care utilization metrics, there was a statistically significant decrease in the percentage COPD patients in BPCI hospitals who used an IRF ($p=0.10$) and a long-term care hospital ($p<0.01$) in the six months prior to their index hospitalization (see Exhibits 72a & 72b). These changes suggest a shift to a more healthy mix of COPD patients in BPCI hospitals. However, increases in the Medicaid eligible and disabled populations suggest the changes in the health status of COPD patients treated at BPCI hospitals were mixed.

[53] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

Exhibits 72a & 72b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for COPD, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospital				Comparison Hospitals				DiD*	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	2,535	18%	1,616	20%	2,700	19%	1,647	20%	0.8	-0.7	2.3	0.29
	65-79	6,290	45%	3,875	47%	6,307	45%	3,837	46%	0.6	-1.3	2.5	0.55
	80+	5,049	36%	2,795	34%	4,867	35%	2,802	34%	-1.4	-3.2	0.4	0.14
Gender	Female	8,652	62%	5,211	63%	8,485	61%	5,079	61%	0.4	-1.5	2.3	0.68
Medicaid and Disability	% Eligible for Medicaid	4,826	35%	2,742	33%	4,879	35%	2,591	31%	2.2	0.4	4.0	0.02
	% Disability, no ESRD	2,837	20%	1,785	22%	3,033	22%	1,778	21%	1.5	-0.1	3.1	0.06
Utilization Six Months Prior to Index Hospitalization+	Inpatient Acute Care Hospital	5,014	36%	3,025	37%	4,597	33%	2,685	32%	1.1	-0.7	2.9	0.24
	Emergency Room Admission	4,959	36%	3,317	40%	4,893	35%	3,250	39%	0.3	-1.5	2.2	0.73
	Home Health	3,767	27%	2,247	27%	3,429	25%	2,017	24%	0.3	-1.3	2.0	0.69
	Inpatient Rehabilitation Facility	285	2%	135	2%	205	1%	122	1%	-0.4	-0.9	0.1	0.10
	Skilled Nursing Facility	1,508	11%	929	11%	1,178	8%	668	8%	0.8	-0.4	1.9	0.18
	Psychiatric Hospital	203	1%	140	2%	154	1%	85	1%	0.3	-0.1	0.7	0.16
	Long-term Care Hospital	140	1%	46	1%	90	1%	63	1%	-0.6	-0.9	-0.2	0.00
	Institutional Nursing Facility	1,882	14%	1,090	13%	1,526	11%	845	10%	0.4	-0.9	1.6	0.54
	No Institutional Care	8,577	62%	5,089	61%	8,981	65%	5,431	66%	-1.2	-3.1	0.6	0.20
	No Post-acute Care	5,414	39%	3,066	37%	5,699	41%	3,284	40%	-0.6	-2.5	1.3	0.55

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.93	2.21	3.02	2.28	2.75	2.09	2.80	2.19	0.04	-0.05	0.12	0.41

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

*These estimates are developed using a Difference-in-Differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: "Count" represents the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

+Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Long Term Care Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare Payments

There was no statistically significant difference in total Medicare payments for the anchor hospitalization plus all care delivered during the 90-day post-discharge period for BPCI-COPD episodes compared to episodes initiated at comparison hospitals (see Exhibit 73). Total Medicare payments ($2015)[54] for a COPD episode declined by 4.6% for BPCI hospitals and declined by 4.3% for comparison hospitals.

Exhibit 73: Impact of BPCI on Total Payments ($2015) for Inpatient Anchor Hospitalization and 90-day PDP for COPD Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=8,088)		Comparison (N=8,079)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total allowed payment amount, inpatient stay plus 90-day PDP	$19,699	$18,796	$19,654	$18,811	-$59	-$776	$657	0.87

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; PDP = Post-discharge period.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and matched comparison hospitals.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

The impact of BPCI on PAC payments and utilization was modest. Relative to the comparison group, average HHA payments for COPD episodes increased by $71 for BPCI participants (p=0.10). However, this change in HHA payments did not result in a statistically significant change in the overall payments for the episode.

d. Quality of Care

There was no indication that claim- and assessment-based measures of quality of care changed in a statistically significant way for BPCI COPD patients relative to patients treated by comparison hospitals. From the baseline to the intervention period, the change in hospital readmission rates, emergency department visits, and mortality following the COPD anchor hospitalization was not statistically different between BPCI-participating and comparison hospitals.

In addition, for SNF, HHA, and IRF users, there was no difference in functional status between those with episodes initiated at a BPCI-participating hospital and those with episodes beginning at a comparison provider, according to an analysis of patient assessment data.

Based on two Waves (4 and 5) of survey data, there was no consistent pattern of significant differences between BPCI and matched comparison COPD respondents in their self-reported functional status, health, or experiences.[55] In Wave 4, no differences between the two groups

[54] Medicare payments are inflated to 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

[55] Episodes of COPD were sampled twice in each survey wave over two consecutive months. Each of the two samples were drawn from hospital discharges that occurred within a single month: Wave 4, May-June 2015; Wave 5, October-November 2015.

were significant at the 10% level. In Wave 5, BPCI respondents reported higher rates of decline in bathing, dressing, using the toilet, or eating (8.0 percentage points, p<0.01) from before to after the episode (see Exhibit 74). Conversely, BPCI respondents reported higher rates of improvement and lower rates of decline from before to after the episode in whether pain limited their regular activities (8.0 percentage points, p=0.03 and -6.2 percentage points, p=0.07, respectively). These were the only significant differences between BPCI and comparison respondents in Wave 5. Given the mixed results across survey waves, there is limited evidence to conclude that BPCI- COPD beneficiaries were better or worse off with regards to self-reported functional status and health relative to comparison beneficiaries. We will continue to monitor survey outcomes for evidence of relative changes in function due to BPCI.

Exhibit 74: Impact of BPCI on Select Survey Outcomes for COPD Episodes, Model 2, May/June 2015 and October/November 2015

Measure	Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
		BPCI	Comparison	BPCI	Comparison	Point Estimate	95% LCI	95% UCI	p-value
Improvement in pain limiting regular activities	4	281	299	45.5	48.8	-3.2	-10.7	4.2	0.39
	5	351	335	48.9	40.8	8.0	0.7	15.4	0.03
Decline in bathing, dressing, using the toilet, or eating	4	280	293	13.0	12.6	0.4	-4.7	5.5	0.88
	5	346	331	18.3	10.3	8.0	3.0	13.0	0.00
Decline in pain limiting regular activities	4	281	299	21.7	21.6	0.1	-6.4	6.5	0.99
	5	351	335	22.6	28.8	-6.2	-12.8	0.4	0.07

This table reports all functional status and care experience measures that have a statistically significant difference between the intervention and comparison groups at the 10% or 5% level in at least one wave.

The improvement indicator takes a value of 1 if a patient moved to a better functional status level after the episode of care or if the patient recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response. The decline indicator takes a value of 1 if the patient moved to a worse functional status group after the episode of care or if the patient recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey. These indicators are assigned a value of 0 otherwise.

Source: Lewin analysis of beneficiary survey data for episodes that began May, June, October, and November 2015 for BPCI and the matched comparison providers.

e. Variation in NPRA among BPCI Providers

There was variation in financial performance among BPCI hospitals for their COPD episodes, as measured by average NPRA per episode, as shown in Exhibit 75.[56] Top performers were defined as hospitals with standardized NPRA for COPD episodes above the 75th percentile of the standardized NPRA[57] distribution, average performers were defined as hospitals with standardized NPRA for COPD episodes between the 25th and 75th percentiles of the standardized NPRA distribution, and bottom performers were those with standardized NPRA at or below the 25th percentile of the distribution. Top performers had a mean average NPRA of $1,752 per episode. Bottom performers had a mean average NPRA of -$2,433.[58]

Exhibit 75: Average NPRA/Episode by Performance Group, Model 2 COPD, Q4 2013- Q2 2015

Performance Group	N	Mean	Median	Min	Max
Top Performers (Above 75th Percentile)	7	$1,752	$1,616	$1,386	$2,377
Average Performers (Between 25th-75th Percentile)	16	-$95	$36	-$1,777	$1,266
Bottom Performers (At or below the 25th Percentile)	8	-$2,433	-$2,122	-$4,580	-$1,304

Source: Lewin analysis of Net Payment Reconciliation Amount data for the Q4 2013- Q2 2015 period.

Consistent with the variation in NPRA, there were differences in the changes in total payments, PAC utilization, and the rate of unplanned readmissions, from baseline to intervention period, across the three performance groups.[59] The average total payment for a COPD episode among top performers decreased by $3,621 from baseline to intervention period, while the average total payment for a COPD episode among bottom performers increased by $116, as shown in Exhibit 76.

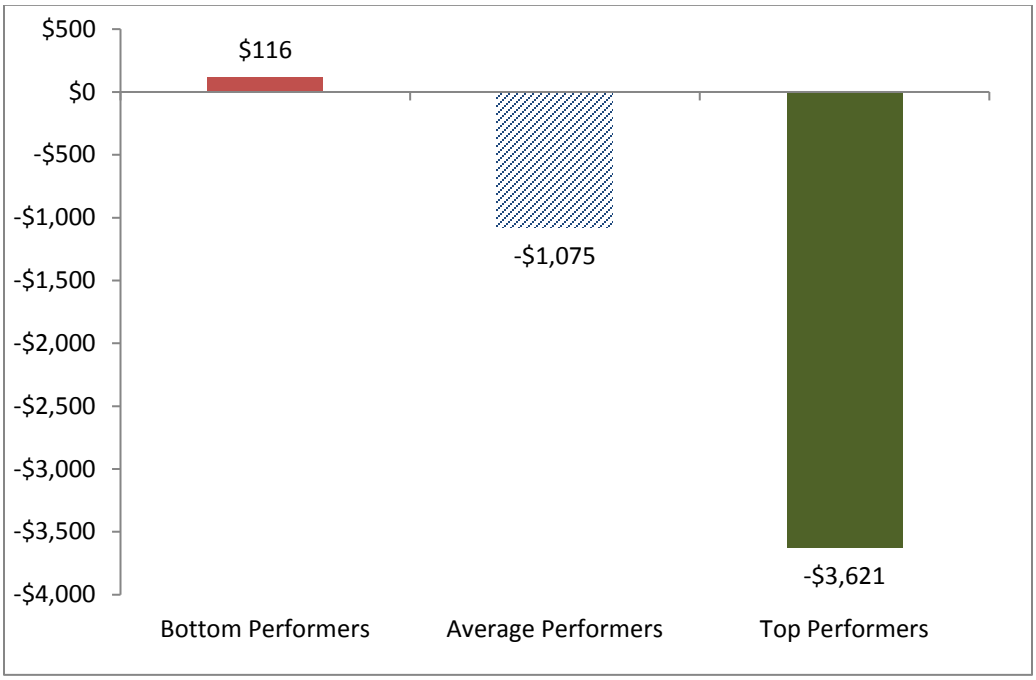
[56] It is important to distinguish between the goals of this analysis and those of the DiD analysis. While the DiD analysis aims to measure the impact of BPCI on total Medicare payments, the goal of this analysis is to understand the factors driving the variation in financial performance among BPCI hospitals. Unlike the DiD analysis that uses total Medicare Part A and B payments during the inpatient hospitalization and 90-day PDP, this analysis focuses on provider's average standardized NPRA per episode. NPRA measures the difference between the target price for services provided during the episode of care and the total dollar amount of Medicare fee-for service expenditures for that episode. When a provider's episode payments are below the target price (i.e., positive NPRA), the provider receives this amount from CMS. When a provider's episode payments exceed the target price (i.e., negative NPRA) the provider may need to return the amount to CMS. For more detail on the NPRA performance measure, please see the Methods section.

[57] To adjust for differences in baseline payments and thus target prices among BPCI providers, we standardize NPRA by dividing it by the provider's average target price.

[58] BPCI providers were required to meet certain criteria to be included in this analysis such as having at least 50 COPD episodes during both the BPCI baseline and intervention periods. For more detail on all the inclusion and exclusion criteria, please see the Methods Section.

[59] The analyses presented in this section encompass the first seven quarters of the initiative (from October 2013 through June 2015) and a baseline period spanning from October 2011 through September 2012.

Exhibit 76: Mean Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, by Performance Group, Model 2 COPD, Q4 2011- Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

The differences in performance across the three groups is also evident in heat maps where red indicates a negative deviation in the mean performance for the group from the overall mean across all BPCI- COPD providers (see Exhibits 77, 78, 79 and 81). Green represents a positive deviation from the mean. [60] Lighter shades represent smaller deviations between the group and total mean and darker shades represent larger deviations. For example, the top right red cell in the heat map in Exhibit 77 shows that total payments declined, from the baseline to intervention period, for top performers relative to all BPCI- COPD providers. The green cell indicates that total payments increased for bottom performers relative to the change in total payments for all BPCI-COPD providers. Although there is some variation in performance among individual providers within these three groups, in general, the performance of individual providers within each group followed similar patterns. As shown in Exhibit 78, bottom performers had increases in payments relative to the mean of all BPCI-COPD providers, with only one exception. However, the different shades of green indicate the variability in relative payment increases within the bottom performance group. Similarly, a majority of top performers had decreases in total payments relative to the average of all of BPCI-COPD providers. However, as indicated by the different shades of red, some top performers experienced large decreases and others experienced smaller decreases.

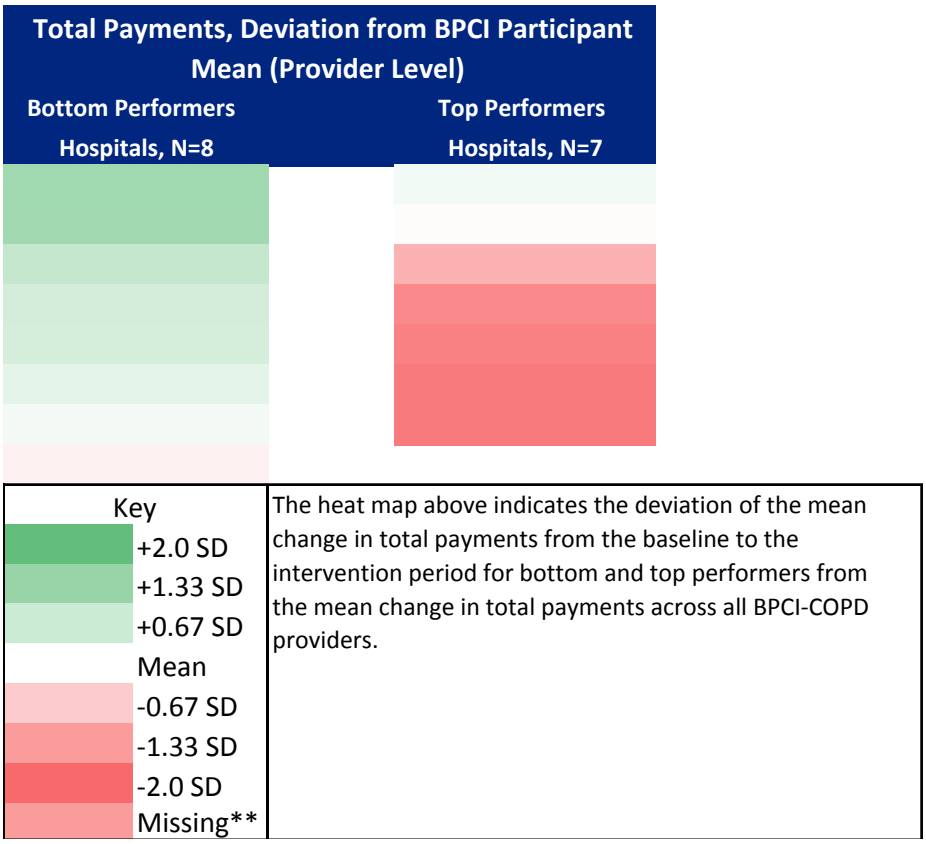
[60] The heat map shown in Exhibit 78 represents deviations of the mean performance of each individual provider from the mean performance across all BPCI- COPD providers.

Exhibit 77: Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, Deviation from BPCI Participant Mean, Model 2 COPD, Q4 2011- Q2 2015

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	**Change in Total Payments**
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in total payments from baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 78: Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, Deviation from BPCI Participant Mean, Model 2 COPD, Q4 2011- Q2 2015



** Provider did not have sufficient sample size for outcome to be risk adjusted.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Relative to all BPCI-COPD hospitals, top performers reduced transfers to institutional PAC, an expensive site for care, and reduced readmissions, an expensive component of an episode.

Changes in both the share of beneficiaries discharged to institutional PAC and the unplanned readmissions rate may account for the decreases in average total payments for a COPD episode observed among top performers. As shown by the red and green colors in the heat map presented in Exhibit 79, relative to the mean performance of all BPCI-COPD providers in our sample, top performers decreased the percent of beneficiaries discharged to institutional PAC while average and bottom performers increased the percent of institutional PAC discharges.

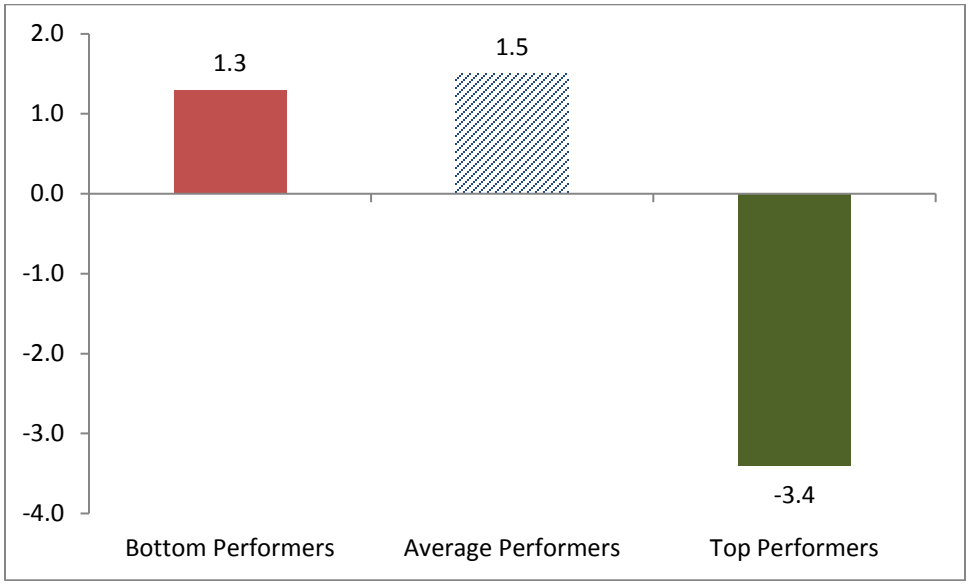
Exhibit 79: Change in Percent Discharged to Institutional PAC, Deviation from BPCI Participant Mean, Model 2 COPD, Q4 2011- Q2 2015

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	**Change in Percent Discharged to Institutional PAC**
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in percent discharged to institutional PAC from baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

The share of patients discharged to institutional PAC declined by 3.4 percentage points, from baseline to intervention period, for top performers. The share of patients discharged to institutional PAC for bottom performers increased by 1.3 percentage points (see Exhibit 80).

Exhibit 80: Mean Change in Percent Discharged to Institutional PAC by Performance Group, Model 2 COPD, Q4 2011- Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

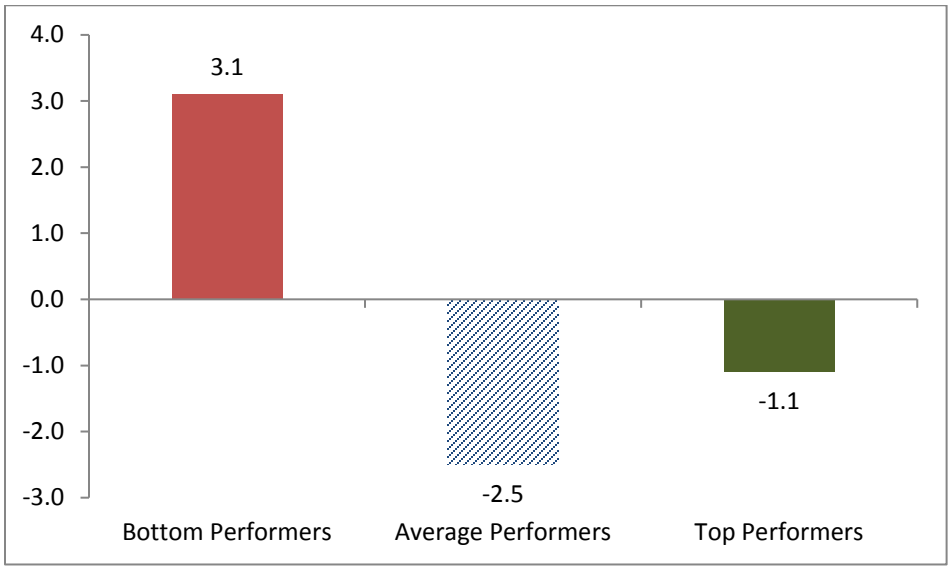
The rate of unplanned readmissions declined for average and top performers, relative to the average, while it increased for bottom performers (see Exhibit 81). From baseline to intervention period, the average readmission rate for top performers declined by 1.1 percentage points.[61] It increased by 3.1 percentage points for bottom performers (see Exhibit 82). Anecdotally, representatives of top performers that we interviewed mentioned using risk assessment tools and case management programs for high risk patients to reduce the likelihood of readmission. A representative of a bottom performer indicated that it was difficult to manage readmissions due to non-compliant and complex patients with multiple co-morbidities. However, our analysis does not indicate that bottom performers were treating sicker patients.

Exhibit 81: Change in Unplanned Readmission Rate, 90-day PDP, Deviation from BPCI Participant Mean, Model 2 COPD, Q4 2011- Q2 2015

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	**Change in Unplanned Readmission Rate, 90-day PDP**
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in unplanned readmission rate from baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 82: Mean Change in Unplanned Readmission Rate, 90-day PDP, by Performance Group, Model 2 COPD, Q4 2011- Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

[61] The Spearman correlation coefficient between standardized NPRA and the change in overall HCC score was not statistically significant, suggesting that the decrease in unplanned readmission rates among top performers did not stem from a change in the health status of the population they served.

The comparisons across performance groups indicates the reasons for the variation in success of the BPCI initiative among BPCI-COPD providers ─ with top performers effectively decreasing average payments for a COPD episode and attaining positive NPRA and bottom performers failing to realize positive NPRA. Relative declines in use of institutional PAC and unplanned readmissions are effective ways to reduce total episode costs, which is consistent with what we see in the top performers.

f. Conclusions

Results from the first eight quarters indicate that the BPCI initiative did not have a statistically significant impact on overall Medicare payments or quality of care for COPD episodes. However, the variation in NPRA within BPCI participants appears to have been driven by the success of top performers in reducing the use of institutional PAC and the rate of unplanned readmissions.

4. Congestive Heart Failure

Medicare beneficiaries received 14,179 congestive heart failure (CHF) episodes of care in 136 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015). BPCI had no statistically significant impact on Medicare payments ($2015)[62] for the anchor hospital stay and the 90-day post-discharge period (PDP).[63] However, there is some evidence that post-acute care (PAC) services changed for BPCI patients relative to the comparison group. Part A payments for inpatient rehabilitation facility (IRF) services declined for beneficiaries treated at BPCI hospitals relative to the comparison group, while payments for home health agency (HHA) services increased. Furthermore, the number of HHA visits among HHA users increased for BPCI patients relative to a matched comparison group.

There were no statistically significant changes in quality of care as measured by claim-based outcomes (readmission rates, emergency department visits, and mortality), or functional status outcomes from patient assessments. Data from the beneficiary survey indicated BPCI and comparison group respondents generally reported similar changes in functional improvement from before their episode of care to after and similar care experiences, with few statistically significant differences.

BPCI-participating hospital representatives indicated through interviews and site visits that they implemented care redesign activities,[64] but as a whole, BPCI hospitals did not successfully lower total payments or reduce readmission rates relative to a matched comparison group. Several interviewees explained that CHF patients tend to be severe (patients are complex to treat and have high numbers of comorbidities), making it difficult to control costs after hospital discharge and making CHF a financially risky episode. In fact, some interviewees indicated they experienced negative financial results. In addition, participants on the Technical Expert Panel (TEP) held on April 12, 2016 speculated that BPCI hospitals might not have been able to reduce

[62] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

[63] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

[64] See Methods section for more information on the BPCI participant site visits and interviews.

readmission rates further to bring down episode payments because providers nationwide had already lowered readmission rates under the Hospital Readmissions Reduction Program under the Affordable Care Act of 2010 where readmissions for CHF are included in assessing hospitals for a penalty.

Some hospitals reduced episode payments below their target amount and achieved positive Net Payment Reconciliation Amounts (NPRA). The BPCI-participating hospitals that had achieved the highest NPRAs lowered episode payments below their target price by reducing readmissions and by discharging patients home with no HHA services rather than to institutional PAC (skilled nursing facilities (SNF), IRF, and long-term care hospitals). In addition, top performers treated healthier CHF patients during the baseline period and had lower discharge rates to home with no HHA compared to bottom performers. This may be an indication they did not use PAC as efficiently as other BPCI hospitals, and so they may have had an easier time reducing PAC use to achieve savings without changing quality of care.

a. Participants

During the first eight quarters of the BPCI initiative, 136 hospitals (32% of all Model 2 BPCI-participating hospitals) participated in the Model 2 CHF episode. Of these, 127 chose 90-day episodes. There were 14,179 CHF episodes initiated by BPCI providers during the initiative (approximately 10% of episodes across the 48 BPCI clinical episodes). The CHF results are based on an average of three quarters of experience because approximately 26% of the hospitals began participating in CHF in Q3 2015, the most recent quarter of claims data available, and 63% joined in either Q2 or Q3 2015. BPCI hospital representatives who we interviewed reported various reasons for deciding to participate in Model 2 CHF, such as high volumes of CHF patients and high baseline readmission rates. A few interviewees were also motivated to join CHF because it aligned with their participation in other federal, local, or private initiatives, such as CMS's readmissions reduction program, local quality improvement initiatives, or private grants to manage CHF patients.

Of the 136 BPCI hospitals, 26 stopped participating in CHF within the first eight quarters of the initiative. During site visits and interviews with representatives from BPCI-participating hospitals, interviewees noted particular challenges with managing CHF episodes, which may help explain the withdrawals from this episode. Some indicated that they had lower episode volumes during the intervention than expected. There were also difficulties identifying the patients who were in CHF episodes because many of these patients had other conditions, such as pneumonia or chronic obstructive pulmonary disease (COPD), which could complicate or delay the MS-DRG assignment. Some interviewees indicated that the presence of multiple comorbidities also made it difficult to control costs after the hospitalization, so that these patients were financially risky. Some said that they had negative financial results for these episodes.

Compared to hospitals that did not participate in BPCI, hospitals that participated in Model 2 CHF episodes were larger, less likely to be government-owned, and more likely to be non-profit, located in urban areas and have higher resident-to-bed ratios (see Exhibits 83a & 83b). They also had more CHF discharges in 2011 and higher average 2011 standardized Part A payments for CHF episodes during the inpatient hospitalization plus 90-day PDP than non-participants. Matched comparison hospitals were identified for 131 of the 136 BPCI-participating hospitals in

the sample.[65] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals with respect to any notable measures.

Exhibits 83a & 83b: Characteristics of BPCI-participating Hospitals with Congestive Heart Failure Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Matched Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All CHF BPCI Hospitals (N=136)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched CHF BPCI Hospitals (N=131)		Comparison Hospitals (N=1,898+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	109	80%	1,594	57%*	321	77%	105	80%	1,551	82%
	Government	5	4%	542	20%*	32	8%	5	4%	82	4%
	For-Profit	22	16%	638	23%	66	16%	21	16%	265	14%
Urban/Rural	Urban	126	93%	1,902	69%*	387	92%	121	92%	1,735	91%
Part of Chain	Yes	69	51%	1,469	53%	218	52%	68	52%	1,014	53%

Characteristic	All CHF BPCI Hospitals (N=136)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched CHF BPCI Hospitals (N=131)	Comparison Hospitals (N=1,898+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	310	175*	311	312	317
Medicare Days Percent	38%	42%*	39%	38%	38%
Resident-to-bed ratio	0.11	0.05*	0.12	0.11	0.11
Disproportionate Share Percent	29%	29%	27%	29%	29%
Number of CHF Discharges, 2011	202	109*	199	203	210
Standardized Part A Allowed Payment, Inpatient Stay plus 90-day Post-discharge Period, CHF, 2011	**$19,645**	**$18,499***	**$18,825***	**$19,757**	**$19,694**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All CHF BPCI" hospitals. "Comparison" hospitals are compared to "Matched CHF BPCI" hospitals.

** BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 779.

[65] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. Five participants were not included in the analysis because either they were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the CHF episode or had fewer than five relevant discharges in 2011 or 2012.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. The changes in CHF patient characteristics treated at BPCI hospitals between the baseline and intervention periods suggest a shift to a more healthy patient mix in BPCI hospitals during the intervention period relative to the comparison group. There was a 2.2 percentage point increase in the share of female patients ($p = 0.01$) and small but statistically significant decreases in three out of ten measures of health care utilization in the six months prior to their index hospitalization for BPCI patients relative to the comparison (see Exhibits 84a & 84b). The percent of BPCI patients admitted to the emergency room, the percent that used a long-term care hospital (LTCH), and the percent that used an IRF all declined, relative to the comparison group.

Exhibits 84a & 84b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Congestive Heart Failure, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	1,561	8%	1,141	8%	1,637	9%	1,265	9%	-0.5	-1.4	0.4	0.27
	65-79	6,158	33%	4,578	34%	6,226	34%	4,531	33%	0.7	-0.8	2.2	0.34
	80+	10,693	58%	7,833	58%	10,549	57%	7,756	57%	-0.2	-1.8	1.3	0.79
Gender	Female	10,129	55%	7,405	55%	10,226	56%	7,185	53%	2.2	0.6	3.7	0.01
Medicaid and Disability	% Eligible for Medicaid	4,462	24%	3,026	22%	4,562	25%	3,070	23%	0.2	-1.1	1.6	0.75
	% Disability, no ESRD	1,792	10%	1,293	10%	1,898	10%	1,414	10%	-0.3	-1.3	0.6	0.51
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	8,988	49%	6,629	49%	8,141	44%	5,950	44%	0.4	-1.2	2.0	0.61
	Emergency Room Admission	6,047	33%	4,788	35%	6,099	33%	5,019	37%	-1.4	-2.9	0.1	0.06
	Home Health	6,880	37%	5,087	38%	6,395	35%	4,738	35%	-0.1	-1.6	1.4	0.94
	Inpatient Rehabilitation Facility	636	3%	383	3%	531	3%	440	3%	-1.0	-1.5	-0.4	<0.01
	Skilled Nursing Facility	3,413	19%	2,671	20%	2,853	15%	2,155	16%	0.8	-0.4	2.0	0.21
	Psychiatric Hospital	115	1%	76	1%	102	1%	79	1%	-0.1	-0.3	0.1	0.45
	Long-term Care Hospital	218	1%	97	1%	176	1%	126	1%	-0.4	-0.7	-0.1	<0.01
	Institutional Nursing Facility**	3,927	21%	2,910	21%	3,406	18%	2,433	18%	0.7	-0.6	1.9	0.28
	No Institutional Care	9,046	49%	6,664	49%	9,835	53%	7,307	54%	-0.5	-2.0	1.1	0.57
	No Post-acute Care	5,771	31%	4,166	31%	6,172	34%	4,390	32%	0.5	-0.9	2.0	0.48

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	4.00	2.46	4.10	2.57	3.81	2.38	3.86	2.45	0.05	-0.02	0.13	0.20

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; ESRD = end stage renal disease.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and matched comparison hospitals.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare Payments and Utilization

During the first eight quarters of the BPCI initiative, the change in total Medicare payments ($2015) for the inpatient hospitalization and all care delivered during the 90-day PDP for BPCI episodes was not statistically different from the change for episodes from a matched comparison group of providers not participating in BPCI (see Exhibit 85). Total payments declined 4% for BPCI hospitals and 3% for the matched comparison group from the baseline to the intervention period.

Exhibit 85: Impact of BPCI on Total Standardized Allowed Payment ($2015) Inpatient Hospitalization plus 90-day Post-discharge Period for Congestive Heart Failure Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=13,347)		Comparison (N=13,328)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total Payment, Inpatient stay plus 90-day post-discharge period	$25,903	$24,799	$25,417	$24,566	-$253	-$873	$367	0.42

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

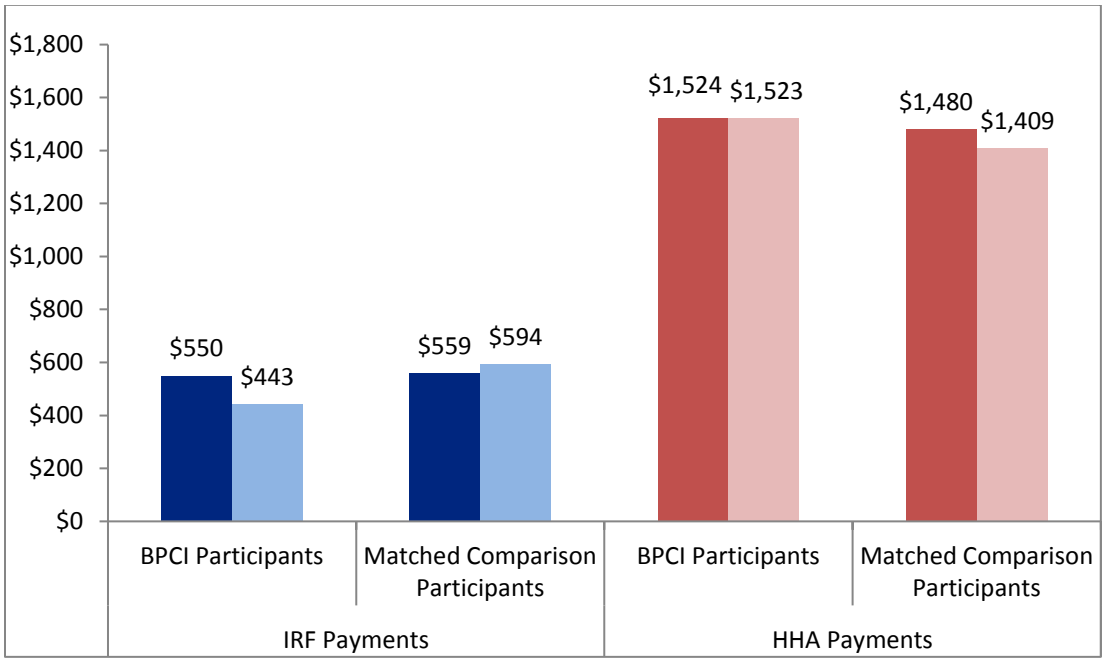
Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

While there was no differential change in total payments, there were differences for particular services. Medicare IRF payments decreased by $143 (p = 0.03) for BPCI episodes relative to the comparison group, while HHA payments increased by $70 (p = 0.10) for BPCI episodes relative to the comparison group (see Exhibit 86).[66] Because IRF and HHA payments account for approximately 2% and 6% of total episode spending, respectively, these relatively small changes did not appreciably affect total payments for the episode.

[66] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Exhibit 86: Impact of BPCI on Payments ($2015) for Inpatient Rehabilitation Facilities (IRF) and Home Health Agencies (HHA) for Congestive Heart Failure Episodes, Model 2, Q4 2011-Q3 2015



Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

In addition, the number of HHA visits among patients with any HHA care increased for BPCI patients relative to the comparison group. The number of home health visits among HHA users increased from 18.5 to 19.6 for BPCI episodes between the baseline and intervention period and from 18.7 to 19.0 for the comparison group.

d. Quality of Care

There was no statistically significant change in the quality of care for CHF Medicare beneficiaries as measured by readmission rates, emergency department visits, and mortality. There were also no statistically significant changes in functional status based on patient assessments for those receiving PAC.

BPCI and comparison group respondents generally reported similar changes in functional improvement from before their episode of care to after. However, results for two out of seven functional improvement measures varied between two different survey waves.[67,68] BPCI respondents were 6.4 percentage points (p=0.03) more likely to report that pain limited their regular activities relative to comparison respondents in the first wave but not the second wave of the survey. In terms of dependence on a mobility-assisted device, BPCI respondents reported a 9.0 percentage point (p<0.01) higher rate of decline and a 5.8 percentage point (p=0.03) lower rate of improvement in the second wave, but not the first wave. There was one difference in results between the two waves regarding patient experience as well. BPCI respondents were 6.9 percentage points (p=0.08) more likely than comparison respondents to report they never received conflicting medical advice in the second wave but not the first. Because these results do not provide a consistent view of either worsening or improving experience of BPCI patients, more data is needed to clarify these issues.

Despite the incentives to reduce readmissions under BPCI, panelists on the TEP speculated that hospitals might not have been able to make additional reductions because they had already achieved reductions in readmissions prior to BPCI (from 2010 to 2012) due to incentives under the CHF readmission penalty requirement of the Hospital Readmissions Reduction Program of the Affordable Care Act of 2010.[69]

Although there were no observable changes in quality of care, BPCI-hospital representatives who we interviewed discussed a variety of care redesign activities aimed at improving the quality of care. While these care redesign activities may be important for improving patient experience, they may not be detected in our outcomes. Interviewees described increasing post-discharge follow up. For example, several hospital representatives indicated that they now call patients at regular intervals after discharge to monitor their progress, while some use care navigators to monitor and coordinate care for their CHF patients. Interviewees also indicated they are focused on ensuring patients receive appropriate PAC. Some have implemented protocols to help determine the best PAC setting for each patient and some work with patients, physicians, and PAC providers to set expectations about PAC. Interviewees shared that they have developed patient education initiatives around topics such as nutrition and medication. They have also pursued other initiatives to help patients manage their medications, such as providing a medication dispenser, using telemonitoring, or asking pharmacists to contact patients after discharge. Interviewees also discussed palliative care use. One interviewee indicated they implemented a palliative care initiative at the same time they entered BPCI, coincidentally. Another interviewee indicated that as part of their care redesign efforts, they introduce palliative care earlier in the episode of care.

[67] CHF patients were surveyed in two waves. The sample for the first wave was drawn from hospital discharges that occurred within a single month in May-June 2015. The second wave was drawn from a single month in October-November 2015.

[68] For each measure, there are two binary indicators, one for both improvement and one for decline. The improvement indicator takes a value of 1 if the respondent moved to a better functional status group after the hospitalization or if the respondent recalled having the highest functional status prior to hospitalization and remained in that status. The indicator is assigned a value of 0 otherwise. The decline indicator follows the same rule.

[69] For more information, see https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html.

e. Variation in NPRA among BPCI Providers

There was variation in financial performance among BPCI hospitals for CHF episodes, as measured by average NPRA per episode (see Exhibit 87).[70,71] Top performers were defined as hospitals with standardized NPRA for CHF episodes above the 75th percentile of the *standardized* NPRA[72] distribution, average performers were defined as hospitals with standardized NPRA for CHF episodes between the 25th and 75th percentiles of the standardized NPRA distribution, and bottom performers were those with standardized NPRA at or below the 25th percentile of the distribution. Top performers had a mean average NPRA of $2,002 per episode. Bottom performers had a mean average NPRA of -$2,361.

Exhibit 87: Average NPRA/Episode by Performance Group, CHF Episodes, Model 2, Q4 2013-Q2 2015

Performance Group	N	Mean	Median	Min	Max
Top Performers (Above 75th Percentile)	12	$2,002	$1,692	$816	$5,172
Average Performers (25th-75th Percentile)	24	-$251	-$153	-$1,734	$940
Bottom Performers (At or below the 25th Percentile)	12	-$2,361	-$2,423	-$3,917	-$950

Source: Lewin analysis of Net Payment Reconciliation Amount data for the Q4 2013- Q2 2015 period.

There was also variation across the three performance groups in the change from baseline to intervention for total Medicare payments during the inpatient hospitalization plus 90-day PDP relative to the average for all BPCI hospitals. On average, mean total CHF episode payments for top performers decreased relative to episode payments for all CHF BPCI hospitals, while total payments for bottom performers increased (see Exhibit 88). In the heat map in Exhibit 88, red represents a negative deviation from the mean group performance from the mean across all CHF BPCI hospitals. Green represents a positive deviation from the mean. Lighter shades represent smaller deviations, i.e., closer distance to the mean performance across all CHF BPCI hospitals, and darker shades represent larger deviations.

[70] Unlike the DiD analysis that uses total Medicare Part A and B payments during the inpatient hospitalization and 90-day PDP, this analysis focuses on provider's average standardized NPRA per episode. NPRA measures the difference between the target price for services provided during the episode of care and the total dollar amount of Medicare fee-for service expenditures for that episode. When a provider's episode payments are below the target price (i.e., positive NPRA), the provider receives this amount from CMS. When a provider's episode payments exceed the target price (i.e., negative NPRA) the provider may need to return the amount to CMS. For more detail on the NPRA performance measure, please see the Methods section.

[71] The analyses presented in this section encompass the first seven quarters of the initiative (from October 2013 through June 2015) and a baseline period spanning from October 2011 through September 2012. BPCI providers were required to meet certain criteria to be included in this analysis, such as having at least 50 CHF episodes during both the BPCI baseline and intervention periods. For more detail on the inclusion and exclusion criteria, please see the Methods section.

[72] To adjust for differences in baseline payments and thus target prices among BPCI providers, we standardize NPRA by dividing it by the provider's average target price.

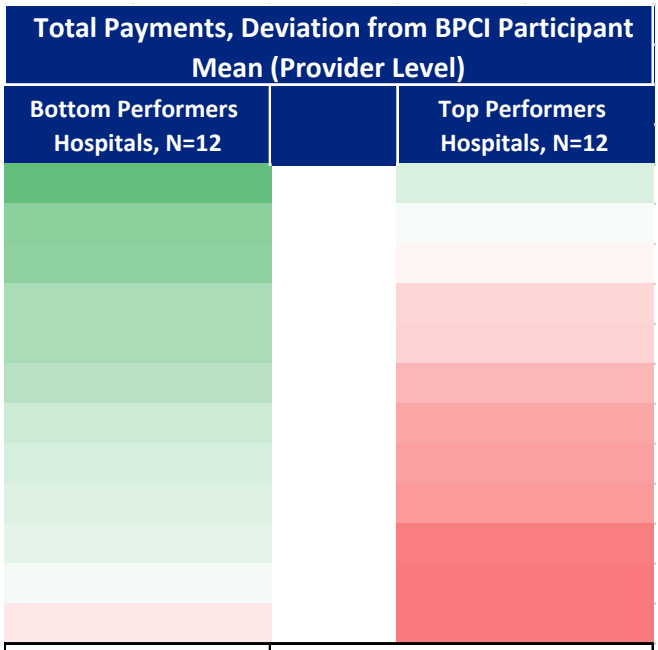
Exhibit 88: Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, Deviation from BPCI Participant Mean by Performance Group, CHF Episodes, Model 2, Q4 2011-Q2 2015

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	**Change in Total Payments**
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in total payment from the baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Within performance groups, there was variation among hospitals in the difference from the mean. Total episode payments declined from the baseline to the intervention period for 9 out of 12 top performers, and they increased for 11 out of 12 bottom performers. In the heat map shown in Exhibit 89, the different shades of red indicate the variability in the relative payment decreases among top performers, and the different shades of green represent the variability in the relative increases among bottom performers. The average total payment for CHF episodes among top performers declined by $1,915 from the baseline to the intervention period, while the average total CHF episode payment among bottom performers increased by $1,736 (see Exhibit 90). The baseline payment levels were similar across the three groups.

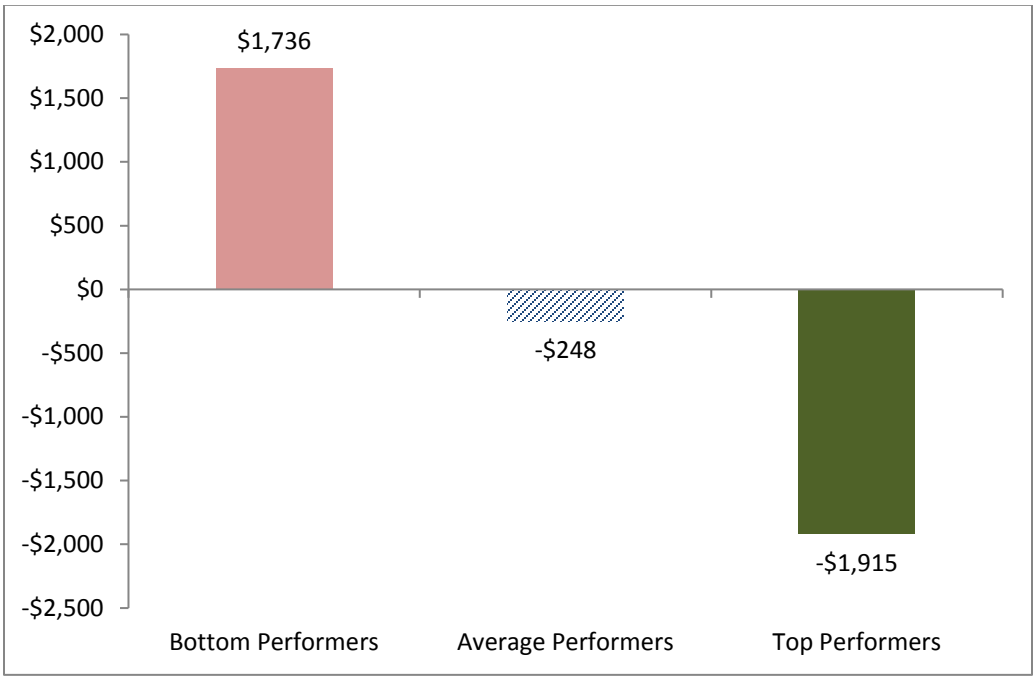
Exhibit 89: Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, Deviation from BPCI Participant Mean by Hospital, CHF Episodes, Model 2, Q4 2011-Q2 2015



** Provider did not have sufficient sample size for outcome to be risk adjusted

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 90: Mean Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, by Performance Group, CHF Episodes, Model 2, Q4 2011-Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

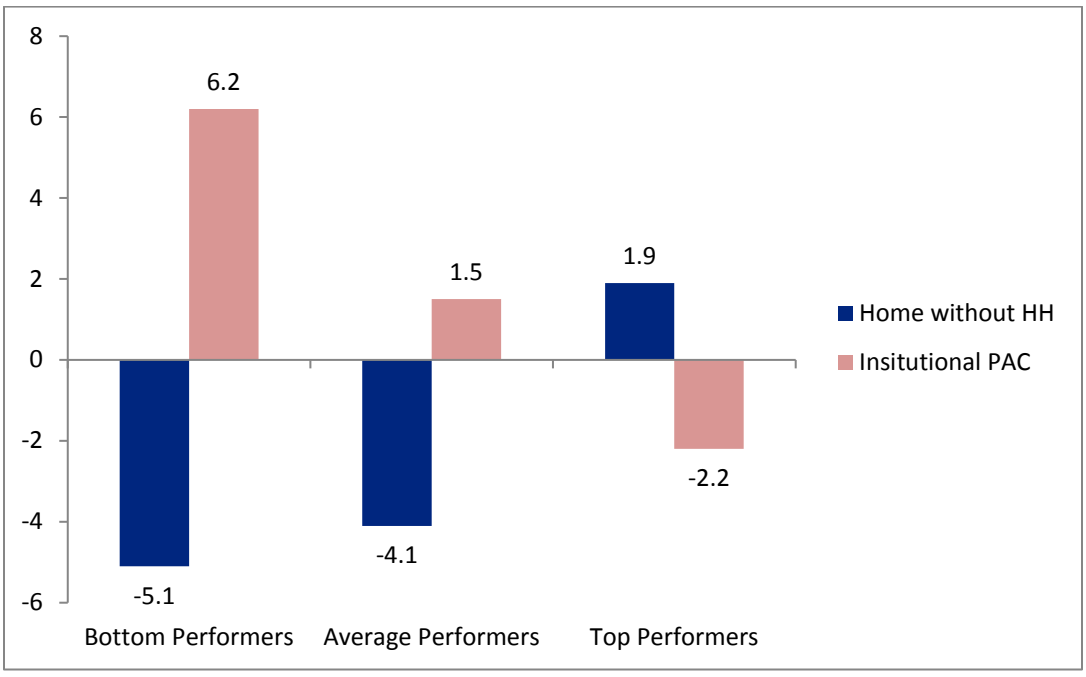
Top performers reduced episode costs by reducing the intensity of PAC care and reducing readmissions more than other BPCI hospitals. As shown by the red and green colors in the heat map presented in Exhibit 91, top performers reduced the proportion of patients discharged to institutional PAC and increased the proportion discharged home without HHA services relative to all BPCI hospitals participating in the CHF episode. Institutional PAC discharges decreased by 2 percentage points and HHA discharges increased by 2 percentage points for top performers (see Exhibit 92). Conversely, bottom performers increased institutional PAC discharges (by 6 percentage points) and decreased HHA discharges (by 5 percentage points). In addition, top performers decreased the rate of unplanned readmissions by 4 percentage points, while bottom performed increased by 2 percentage points (see Exhibit 93).

Exhibit 91: Change in Proportion Discharged to Institutional PAC, Deviation from BPCI Participant Mean by Performance Group, CHF Episodes, Model 2, Q4 2011-Q2 2015

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	Change in Percent Discharged to Institutional PAC
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in total payment from the baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

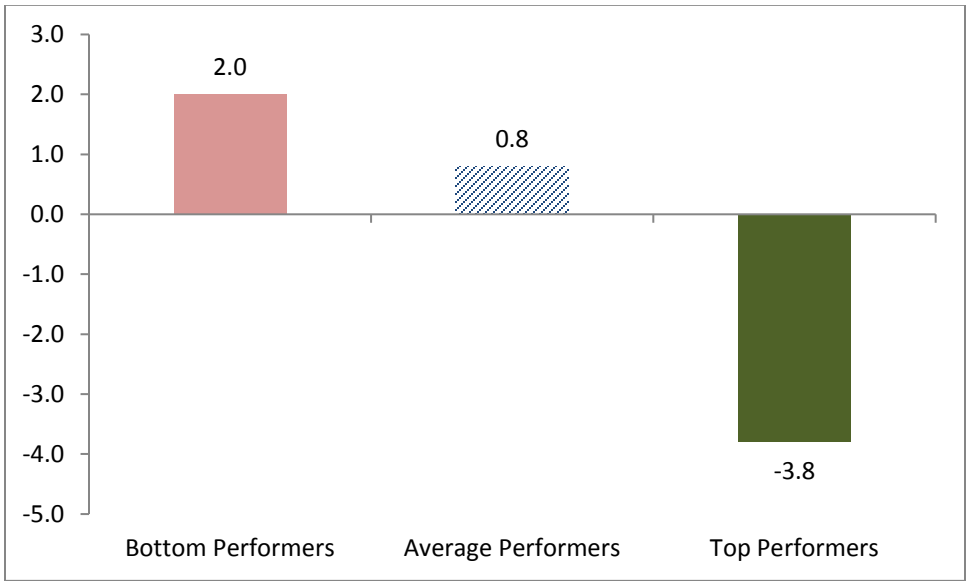
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 92: Mean Changes in Proportions Discharged to Home without Home Health Agency Care and to Institutional PAC, by Performance Group, CHF Episodes, Model 2, Q4 2011-Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 93: Mean Change in Unplanned Readmission Rate, 90-day Post-discharge Period, by Performance Group, CHF Episodes, Model 2, Q4 2011-Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Top performers appeared to have a healthier CHF patient population and a higher share of PAC discharges during the baseline relative to all hospitals that participated in the CHF episode, which may have contributed to their ability to reduce institutional PAC use and readmissions. They had a lower average Hierarchical Condition Category (HCC) score, as indicated in Exhibit 94.[73] They also discharged a smaller share of patients home without any PAC services than the two other performance groups (see Exhibit 95). These initial conditions, patients with fewer comorbidities and greater use of PAC, may have made it easier for these hospitals to reduce institutional PAC discharges to achieve positive NPRA during the intervention period.

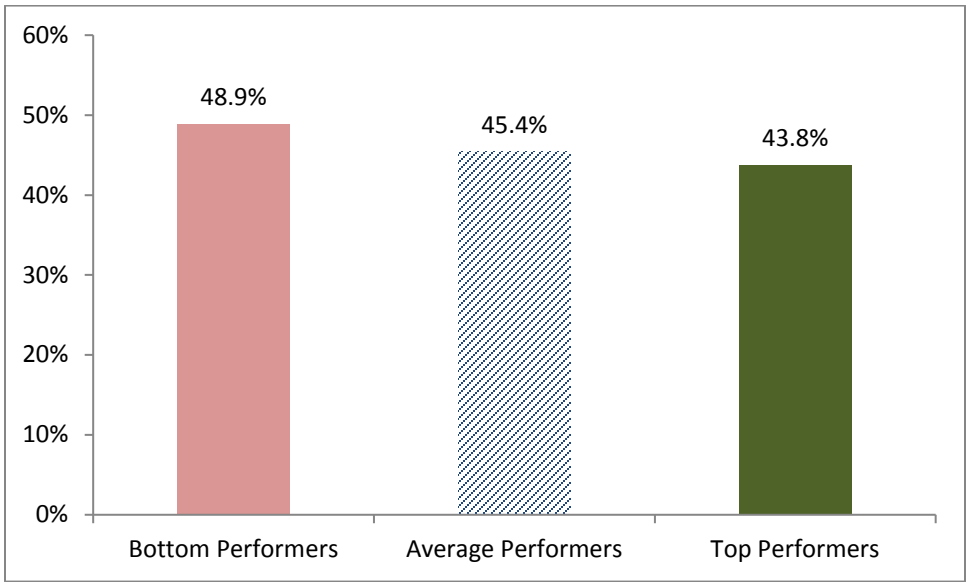
[73] A beneficiary's HCC score is based on observed medical diagnoses and reflects the expected cost of their health care requirements. Higher scores suggest a less healthy patient and greater future spending.

Exhibit 94: Hierarchical Condition Category (HCC) Overall Score, Baseline, Deviation from BPCI Participant Mean, CHF Episodes, Model 2, Q4 2011-Q3 2012

Deviation from BPCI Participant Mean (Performance Group Level)	
Performance Groups Based on Performance Measure	HCC Overall Score, Baseline
Top Performers (Above 75th Percentile)	
Average Performers (Between 75th-25th Percentile)	
Bottom Performers (Under 25th Percentile)	
Red indicates a reduction in total payment from the baseline to intervention period, relative to the overall average. Green indicates a relative increase.	

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 for BPCI providers.

Exhibit 95: Mean Percent Discharged Home without Home Health Agency Care, by Performance Group, Baseline, CHF Episodes, Model 2, Q4 2011-Q3 2012



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 for BPCI providers.

We interviewed a sample of top and bottom performing hospital representatives about their performance under BPCI and any challenges related to the CHF episode. Top and bottom performers reported employing similar care redesign approaches and strategies to reduce readmissions. Interviewees expressed a variety of challenges related to patient populations and relationships with PAC partners. One important difference between the top and bottom performing groups was their ability to identify CHF patients before discharge. Two out of three top performing hospitals interviewed indicated that they identified CHF patients prior to

discharge through sophisticated software algorithms, while both bottom performers that were interviewed discussed difficulty in identifying CHF patients prior to discharge.

Top and bottom performers discussed similar activities to engage with PAC providers, such as meeting regularly to share data and other information. One interviewee from a top performing hospital noted that the preferred PAC network providers had lower readmissions than other PAC providers. A bottom performer interviewee indicated that they had lowered the SNF length of stay and readmissions as a result of PAC partnerships under BPCI. Another bottom performing interviewee noted that successful partnerships with PAC providers would be achievable if they could limit patient choice to a list of preferred providers, allowing the hospital to work with specific facilities that were better able to provide quality care and control costs.

f. Conclusions

Overall, there were few statistically significant changes in Medicare payments, the utilization of health care services, and the quality of care for CHF episodes initiated in BPCI-participating hospitals in Model 2 relative to a matched comparison group. While a sample of BPCI hospital representatives that we interviewed discussed care redesign activities aimed at lowering readmission rates, improving coordination with PAC providers, and providing better education, there have not been statistically significant changes in total payments or quality of care.

A comparison of top and bottom performing Model 2 CHF BPCI hospitals, measured by standardized NPRA, indicates that top performing hospitals may have achieved positive NPRA through shifting to a smaller share of institutional PAC discharges and reducing unplanned readmissions. They may have had an easier time than bottom performing hospitals in making these changes because of their higher PAC use and healthier patient mix in the baseline period. These factors might have given the top performing hospitals an advantage, as they were able to become more efficient in their PAC use during the BPCI intervention period. In addition, a select sample of top and bottom performers indicated that the ability to identify CHF patients before discharge and discharging patients to preferred PAC providers were important for managing CHF episodes successfully under BPCI.

5. Coronary Artery Bypass Graft (CABG)

a. Summary

This analysis distinguishes between elective and urgent/emergent CABG procedures based on recommendations from a clinical technical expert panel (TEP). Average payments for urgent/emergent episodes were substantially higher than for elective episodes. For both urgent/emergent and elective episodes, we observed no statistically significant change in total Medicare payments for BPCI episodes compared to a matched comparison group, though there were some differences in the changes from the baseline to intervention period for payments by type of service within each group.

There was a significant increase in the anchor stay payments and pre- and post-bundle episode payments for urgent/emergent episodes. The higher anchor stay payments for BPCI episodes was due to more outlier payments for high cost inpatient stays and fewer short stay transfer cases. There was also a statistically significant increase in the 30-day mortality rate for BPCI urgent/emergent episodes relative to the change for the comparison group, but this outcome was

not statistically significant when measured over the 90-day post-discharge time period. Further, the 30-day mortality rate may be biased as we were unable to test the null hypothesis that there were parallel trends for BPCI participants and the comparison group during the baseline period, which is required for an unbiased estimate. There were no statistically significant changes in claim-based measures of quality of care for elective episodes.

We examined whether the increased outlier payments might be an indication of a more severe patient mix treated by BPCI-participating hospitals, which could explain the higher pre- and post-episode payments as well as elevated 30-day mortality rate. We found that the BPCI impact on the 30-day mortality rate and payments for the 120- and 180-day post-bundle periods were no longer statistically significant after controlling for outlier payments. The pre- and 30- and 90-day post-bundle payments still had a statistically significant increase for BPCI relative to the change for the comparison group. These mixed results are not definitive with respect to whether BPCI participants are treating a more severe mix of patients or why BPCI participants are likelier to incur outlier payments than comparison episodes.

b. Background

CABG, a surgical procedure that can improve blood flow to the heart, is used to treat severe coronary heart disease. In fiscal year 2014, over 60,000 Medicare beneficiaries had a CABG, as indicated by a hospital discharge for one of six Medicare Severity-Diagnosis Related Groups (MS-DRGs) in the BPCI CABG clinical episode, resulting in nearly $2 billion in Medicare hospital payments.[74]

In this issue brief, we summarize key findings from the evaluation of CABG episodes for participating hospitals under Model 2 of BPCI. First, we explain how we characterized CABG episodes as elective or urgent/emergent and why this is important. This is followed by a description of participating hospitals and how they differ from hospitals that did not participate in the CABG episode. The next section describes the impact of BPCI on claim-based cost and quality outcomes for CABG episodes. Additionally, assessment data was used to determine the impact of BPCI on activities of daily living (ADL) for patients who were treated by home health agencies (HHAs), skilled nursing facilities (SNFs) or inpatient rehabilitation facilities (IRFs) long enough to have two patient assessments. Characteristics of BPCI and comparison patients are also compared to determine whether there was a change in patient mix.

c. Defining Elective versus Urgent/Emergent CABG Episodes

A clinical TEP that we convened highlighted the importance of distinguishing between elective and urgent/emergent episodes when examining the potential impact of BPCI on costs and quality.[75] This is because elective and urgent/emergent procedures would have different quality, utilization and payment outcomes. Further, providers are likely to have more control over the care delivered during elective episodes.

[74] Centers for Medicare and Medicaid Services. (2014). Inpatient Charge Data FY 2014. Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Inpatient2014.html

[75] See Appendix I for a summary of the CABG TEP.

There is no indicator on the claims (or any other data available for this study) to definitively categorize CABG episodes as elective or urgent/emergent. The TEP recommended using the MS-DRG of the anchor inpatient hospital stay to distinguish episodes with percutaneous transluminal coronary angioplasty (PTCA) or cardiac catheterization (MS-DRGs 231-234) during the admission from those without (MS–DRGs 235 and 236) as shown in Exhibit 96. Hospitalizations with PTCA or cardiac catheterization are more likely to be urgent/emergent. Hospitalizations without one of these procedures are more likely to be elective, with the PTCA or cardiac catheterization occurring prior to the hospitalization for the CABG.

Exhibit 96: MS-DRGs of the anchor inpatient hospital stay for Urgent/Emergent and Elective Coronary Artery Bypass Graft

CABG Subgroup	MS-DRGs
Urgent/Emergent	MS-DRG 231 CABG with percutaneous transluminal coronary angioplasty (PTCA) with major complications or comorbidities
	MS-DRG 232 CABG with PTCA without major complications or comorbidities
	MS-DRG 233 CABG with cardiac catheterization with major complications or comorbidities
	MS-DRG 234 CABG with cardiac catheterization without major complications or comorbidities
Elective	MS-DRG 235 CABG without cardiac catheterization with major complications or comorbidities
	MS-DRG 236 CABG without cardiac catheterization without major complications or comorbidities

The TEP acknowledged limitations with this approach. Some patients with long commutes to the hospital, for example, may undergo a catheterization and a CABG a day or two apart during the same hospital stay. The TEP also acknowledged that there may be different types of urgent/emergent cases. For example, there are patients who require an emergent procedure in less than 24 hours and those who require a longer pre-operative timeframe to evaluate and stabilize active conditions, as well as longer stays post-surgery. Despite these limitations, the TEP felt that the MS-DRG based approximation would improve the analysis. No Model 2 participants mentioned different care redesign strategies for CABG patients based on whether or not the episode was elective in any of the interviews or site visits we conducted.[76]

d. Characteristics of Participating Hospitals

During the first eight quarters of the BPCI initiative, 41 hospitals (10% of all Model 2 BPCI hospitals) participated in the CABG episode, 9 of which chose 30-day episodes and 32 of which chose 90-day episodes. There were 2,302 CABG episodes during this period (approximately 2% of episodes across the 48 BPCI clinical episodes).

Because participants were allowed to join BPCI over an extended period, the CABG results are based on an average of four quarters of experience. Approximately 37% of the hospitals began participating in CABG in Q3 2015. Of the 41 hospitals, three stopped participating in CABG within the first eight quarters of the initiative.

Compared to hospitals *not* participating in BPCI, hospitals that participated in the CABG episode were more likely to be non-profit (Exhibits 97a & 97b). They also had more beds, a lower

[76] See Methods section for more information on the BPCI participant site visits and interviews.

proportion of Medicare days and a higher resident-to-bed ratio, on average. Participants were almost exclusively in urban locations. Prior to joining BPCI, participating hospitals had more CABG discharges and had higher average 2011 standardized Part A payments for services during the anchor CABG hospitalization plus 90-days post-discharge than non-participants ($39,003 compared to $37,174).

In general, hospitals participating in CABG were more similar to all BPCI hospitals, regardless of whether they participated in CABG, compared to hospitals not participating in BPCI. However, CABG participants had markedly higher Medicare Part A payments for services during the anchor CABG hospitalization plus 90 days post-discharge than all BPCI participants ($39,003 compared to $36,909).

Matched comparison hospitals were identified for 33 of the 41 BPCI-participating hospitals in the sample.[77] The characteristics of CABG hospitals and the matched comparison hospitals were very similar.

[77] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were four participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter or had fewer than five relevant discharges in 2011 and 2012. There were four other participants for which there was not a match within the caliper.

Exhibits 97a & 97b: Characteristics of BPCI-participating Hospitals with Coronary Artery Bypass Graft Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospital EIs, and Comparison Hospitals Model 2, Q4 2013 - Q3 2015

Characteristic		All CABG BPCI Hospitals (N=41)**		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)		Matched CABG BPCI hospitals (N=33)		Comparison Hospitals (N=387+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	32	78%	1,594	57%*	321	77%	26	79%	308	79%
	Government	3	7%	542	20%*	32	8%	2	6%	31	8%
	For-Profit	6	15%	638	23%*	66	16%	5	15%	48	12%
Urban/Rural	Urban	40	98%	1,902	69%*	387	92%*	32	97%	372	96%
Part of Chain	Yes	23	56%	1,469	53%	218	52%	19	58%	221	57%

Characteristic	All CABG BPCI Hospitals (N=41)**	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)	Matched CABG BPCI hospitals (N=33)	Comparison Hospitals (N=387+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	498	175*	311*	491	489
Medicare Days Percent	35%	42%*	39%*	34%	33%
Resident-to-bed ratio	0.24	0.05*	0.12*	0.23	0.23
Disproportionate Share Percent	30%	29%	27%*	31%	32%
Number of CABG Discharges, 2011	56	14*	56	59	59
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, CABG episodes, 2011	**39,003**	**37,174***	**36,909***	**38,276**	**38,517**

PDP=post-discharge period; CABG=coronary artery bypass graft

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All CABG BPCI" hospitals. "Comparison" hospitals are compared to "Matched CABG BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers were matched to multiple BPCI hospitals. The unique number of matched comparison providers is 225.

e. Medicare Payments and Utilization

Total Episode Payments

Medicare payments were analyzed separately for elective and urgent/emergent episodes.[78] There was no statistically significant difference in the change in total payments for the anchor stay plus 90-day post-discharge period (PDP) for BPCI participants from the baseline to the intervention

[78] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

period relative to the comparison group. Medicare payments for services provided during the anchor CABG acute care hospital stay plus 90-day PDP averaged $45,065 ($2015) for elective BPCI episodes, declining by 5.7% in the intervention period to $42,509 (Exhibit 98). For matched comparison episodes, total payments declined by 5.4%, from $44,798 to $42,369.

Average payments for urgent/emergent episodes were substantially higher than for elective episodes. Medicare payments for services provided during the anchor stay plus 90-days post discharge remained relatively stable from the baseline ($55,330) to intervention period ($55,538) for urgent/emergent BPCI episodes. The average payment declined by 3.9 percent (from $55,696 to $53,500) for comparison episodes, although the difference in the change between BPCI and comparison episodes was not statistically significant.

Exhibit 98: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day PDP, Elective and Urgent/Emergent CABG, Model 2, Q4 2011-Q3 2015

Type of Episode	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Elective	792	792	$45,065	$42,509	$44,798	$42,369	-$127	-$2,961	$2,707	0.93
Urgent/ Emergent	829	843	$55,330	$55,538	$55,696	$53,500	$2,404	-$603	$5,411	0.12

PDP = post-discharge period; DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome. Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

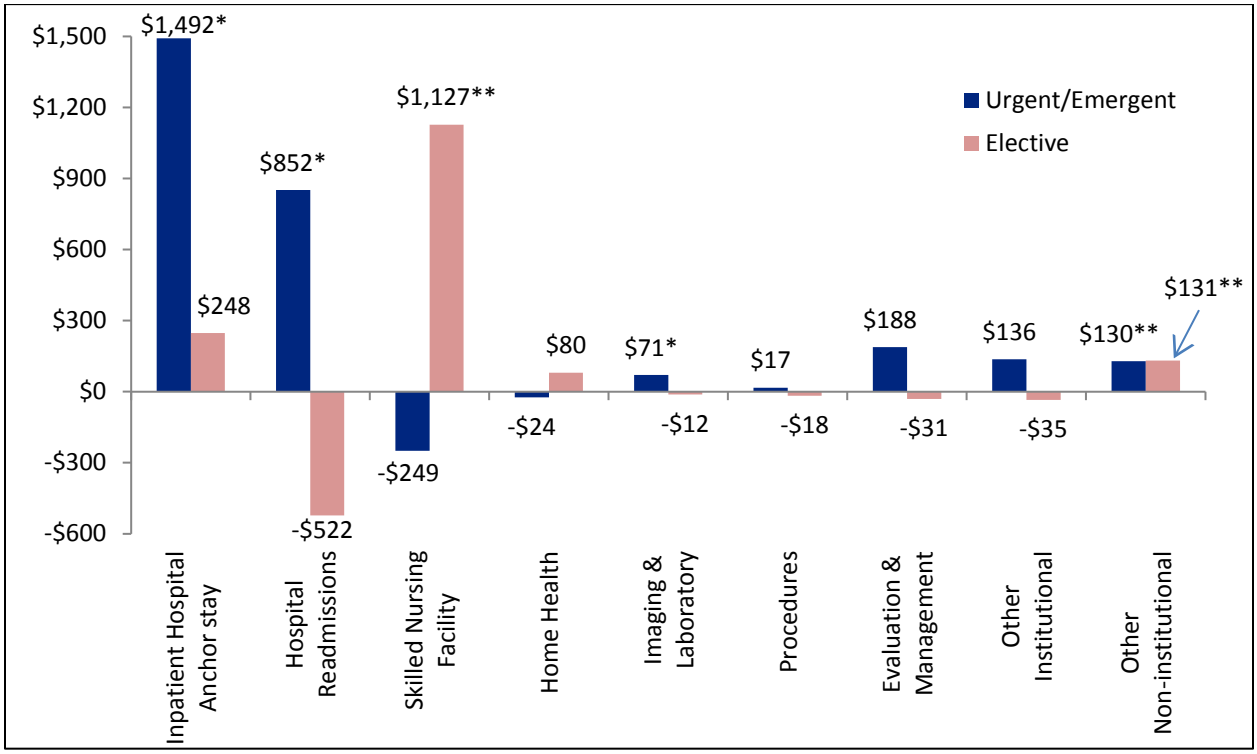
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison hospitals.

Shifts in Payments by Type of Service within Episodes

Even though changes from the baseline to intervention period in total episode payments were not significantly different between BPCI and matched comparison episodes, there were statistically significant differences in payments by type of service (Exhibit 99). The largest difference in payment changes for elective episodes was for SNF payments. BPCI SNF payments per elective episode increased from $2,452 in the baseline to $2,975 (21.3%) in the intervention period, compared to a decrease from $3,210 to $2,605 (-18.8%) for comparison episodes, resulting in a relative increase of $1,127 for SNF payments due to BPCI (p=0.02).[79] There was a 5.7 day increase in SNF days for elective BPCI episodes (p<0.01). The average number of SNF days for BPCI episodes increased 11.8% from 21.7 in the baseline to 24.3 days in the intervention period. SNFs days decreased 13.8% from 22.8 to 19.6 for comparison episodes.

[79] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Exhibit 99: BPCI Impact on Payments ($2015) by Service Type within the 90-day Bundle, Elective and Urgent/Emergent CABG, Model 2, Q4 2011-Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

* Indicates $p < 0.10$.

** Indicates $p < 0.05$.

Other non-institutional care payments per elective BPCI episode increased 52.3%, from $225 to $343, while payments decreased by 4.4% ($304 to $291) for matched comparison episodes ($131, p=0.02) from the baseline to intervention period.[80] The relative increases for SNF and other non-institutional care were offset by relative decreases in payments for readmissions and other types of services in the aggregate (although the differences in other services were not statistically significant).

Except for the relative increase in the number of SNF days (discussed above), there were no statistically significant changes in utilization for elective CABG episodes. This includes the hospital length of stay as well as the number of home health visits and the proportion of episodes that involved any post-acute care (PAC).

The Medicare payment for the anchor hospitalization for BPCI urgent/emergent CABG episodes increased $1,492 (p=0.07) from the baseline to intervention periods relative to the anchor hospitalization payment for comparison episodes. This difference was largely due to a decline in the average payment for comparison hospitals. Anchor hospitalization payments for BPCI episodes

[80] The other non-institutional category of services includes services such as ambulance, chiropractic, vision, hearing, and speech services.

were relatively stable at $34,282 in the baseline period and $34,365 in the intervention period. Payments for comparison episodes declined by 4.0% during this time, from $35,207 to $33,798.

BPCI and comparison episodes differed in whether the payment for the anchor hospitalization was increased because of exceptionally costly outlier cases or reduced because of transfers, either to other acute care hospitals or to PAC providers (see Exhibit 100). Consistent with the relative increase in BPCI payments for the anchor hospitalization, the proportion of BPCI episodes incurring outlier payments increased from 8.8% in the baseline to 13.4% in the intervention period; while there was little change in the proportion for comparison episodes (13.6% and 13.9%). At the same time, there was an increase in the proportion of comparison episodes with short-stay transfers to PAC providers (17.8% to 25.4%), which resulted in a reduction to the anchor hospitalization payment, and a slight decrease for BPCI episodes (21.3% to 20.2%). While there were few of these cases for urgent/emergent CABG episodes, our payment measure for the inpatient stay can include payments for the initial hospital stay as well as payments to any subsequent hospitals in the event of a transfer to another hospital. As expected, when we exclude transfers and outlier episodes from our analysis, the change in payments for the anchor hospitalization for BPCI episodes relative to the comparison group is no longer statistically significant.[81]

Exhibit 100: Proportion of Outlier cases and Short stay transfers at BPCI and Comparison Hospitals, Urgent/Emergent CABG, Model 2, Q4 2011-Q3 2015

	BPCI		Comparison	
Type of Hospital Stay	Baseline	Intervention	Baseline	Intervention
Outlier	8.8%	13.4%	13.6%	13.9%
Short stays transferred to post-acute care	21.3%	20.2%	17.8%	25.4%
Short stays transferred to another hospital	0.0%	0.1%	0.2%	0.2%
Other transfers to another hospital	0.1%	0.2%	0.2%	0.5%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

There was also an increase in readmission payments for urgent/emergent BPCI episodes relative to comparison episodes ($852, p=0.10). Readmission payments for BPCI episodes increased 19.0%, from $2,817 in the baseline to $3,353 in the intervention period, compared with a 10.7% decline, from $2,956 to $2,639, for comparison episodes. The relative increases in payments for the anchor stay and readmissions were the reasons that total payment (the anchor stay plus 90-day PDP) increases were higher (by $2,404, which was not statistically significant) for BPCI urgent/emergent episodes relative to the comparison.

There were also differences between BPCI and comparison episodes in payment changes for imaging and laboratory ($71, p=0.07) and other non-institutional services ($130, p<0.01), both of which comprise a small share of total Medicare payments for CABG episodes.

There were no statistically significant changes in utilization (e.g., proportion of patients discharged to SNF, HHA, IRF; number of SNF days; number of HHA visits) for urgent/emergent episodes.

[81] The DiD estimate decreased from $1,492 to -$263 (p=0.34).

Pre- and Post-Bundle Payments for Urgent/Emergent Episodes

BPCI participants may attempt to reduce episode costs by changing the timing of services so that they are delivered before the anchor hospitalization or after the end of the bundle period. This tactic would be more likely for episodes that are planned because the episode initiator (EI) could change the anchor hospitalization date or timing of other services to minimize the payments included in the bundle. We monitor the pre- and post-bundle payments because although changing the timing of services could reduce episode payments, it would not necessarily reduce total Medicare payments for the beneficiary and might even result in payment increases.

There were no differences in the change in pre-bundle payments, measured as Part B payments in the 30 days prior to the anchor stay, between BPCI and comparison elective episodes. There also were no differences in the change in post-bundle payments, measured as Part A and B payments within 30, 60, 90 and 180 days after the 90-day episode, for elective episodes. In contrast, there were statistically significant increases for all measures of pre- and post-bundle payments for urgent/emergent episodes (Exhibit 101). For all pre- and post-bundle payment measures, payments for BPCI episodes increased from the baseline to the intervention period, while payments for the matched comparison episodes decreased.

It seems unlikely that providers would be able to shift substantial amounts of care prior to the bundle for urgent/emergent cases. The increase in the post-bundle payment outcomes may indicate that BPCI participants had patients postpone services until after the end of the episode. Alternatively, BPCI participants may have arranged for patients to receive additional services to prevent a relapse or maintain their function past the end of the episode, such as cardiac rehabilitation. In fact, site visit interviews with several BPCI-participating hospitals with CABG episodes mentioned enhancing the use of rehabilitation services post-discharge as well as encouraging more active patient follow-up with the surgeon and care managers as part of their strategy for improved care to CABG patients.

Exhibit 101: Impact of BPCI on Pre- and Post-Bundle Payments ($2015) for Urgent/Emergent Episodes

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Part B, 30-day pre-bundle period	840	848	$1,608	$1,839	$1,602	$1,585	$249	$9	$488	0.04
Part A&B, 30-day post-bundle period	697	723	$1,919	$2,851	$2,505	$2,194	$1,243	$539	$1,948	<0.01
Part A&B, 90-day post-bundle period	572	569	$5,734	$7,740	$6,239	$6,188	$2,057	$2	$4,112	0.05
Part A&B, 120-day post-bundle period	466	471	$7,810	$9,934	$7,419	$7,039	$2,503	-$301	$5,308	0.08
Part A&B, 180-day post-bundle period	433	445	$10,083	$13,222	$9,843	$9,747	$3,236	-$78	$6,549	0.06

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

f. Quality of Care

The findings on the quality of care measures differed, as with the payment measures, for elective and urgent/emergent episodes (Exhibit 102). There was no statistically significant change in unplanned readmission, emergency department use, or mortality rates for elective episodes. For urgent/emergent episodes, however, the 30-day mortality rate for BPCI episodes increased from 0.8% to 1.4% from the baseline to the intervention period, while the rate decreased from 1.2% to 0.8% for matched comparison episodes (0.9, p=0.09). However, this estimate may be biased as we did not have sufficient sample size to test the null hypothesis that the BPCI and comparison groups had parallel trends during the baseline period, which is required for an unbiased estimate.

Exhibit 102: Impact of BPCI on Unplanned Readmission, Emergency Department and Mortality Rates, Urgent/Emergent CABG, Model 2, Q4 2011-Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Emergency Department Use, 30-day PDP	867	865	9.2%	13.6%	13.4%	14.4%	3.4	-1.0	7.7	0.13
Emergency Department Use, 90-day PDP	833	847	19.1%	24.0%	25.2%	25.6%	4.5	-1.1	10.0	0.11
Unplanned readmission rate, 30-day PDP	867	865	13.8%	13.6%	15.6%	12.9%	2.6	-1.3	6.4	0.19
Unplanned readmission rate, 90-day PDP	833	847	21.2%	20.5%	22.6%	19.2%	2.7	-1.8	7.2	0.24
All-cause mortality rate, 30-day PDP	864	865	0.8%	1.4%	1.2%	0.8%	0.9**	-0.2	2.0	0.09
All-cause mortality rate, 90-day PDP	830	847	2.1%	2.5%	2.5%	1.8%	1.2	-0.5	2.9	0.18

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; PDP = post-discharge period.

Note: Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

**There was insufficient sample during the baseline period to test if the BPCI and comparison providers were on parallel trends for this outcome. Equal trends test was conducted for total allowed payment amount IP through 90-day PDP, emergency department visits, readmission, and mortality outcomes.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Among the patients discharged to PAC, there was only one statistically significant difference between BPCI and the matched comparison group among the 10 ADL measures investigated for elective cases and none for urgent/emergent cases. The share of BPCI patients with an elective CABG who were discharged to a SNF and experienced an improvement in mobility functioning decreased from 78.6% to 68.5%, while the comparison group share increased from 60.2% to 65.1%, resulting in a relative decrease of 15.0 percentage points (p=0.06) due to BPCI.

g. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the payment, utilization, and quality impact estimates presented above control for changes in these patient characteristics, they do not directly examine changes in patient mix. We observed no consistent pattern of changes in patient characteristics from the baseline to the intervention period for either elective or urgent/emergent BPCI episodes relative to comparison episodes (Exhibits 103a, 103b, 103c & 103d). For both elective and urgent/emergent episodes, there were no statistically significant differences in the change in age, Medicaid eligibility, or the average number of comorbidities as represented by Hierarchical Condition Category (HCC) indicators between BPCI and comparison episodes. For urgent/emergent cases, emergency room (ER) visits increased 6.1 percentage points more for BPCI than comparison episodes (p=0.03) in the six months prior to the index hospitalization.

Exhibits 103a & 103b: Characteristics of BPCI and Comparison Beneficiaries with an Elective Hospitalization for CABG, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI CABG Hospital				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	89	12%	70	9%	73	10%	70	9%	-2.2	-6.3	2.0	0.31
	65-79	534	72%	606	75%	539	73%	605	74%	0.8	-5.4	7.0	0.80
	80+	119	16%	137	17%	130	18%	138	17%	1.4	-3.9	6.6	0.61
Gender	Female	211	28%	216	27%	220	30%	200	25%	3.2	-3.1	9.4	0.32
Medicaid and Disability	% Eligible for Medicaid	118	16%	109	13%	90	12%	91	11%	-1.6	-6.3	3.2	0.52
	% Disability, no ESRD	113	15%	84	10%	84	11%	83	10%	-3.8	-8.3	0.7	0.10
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	284	38%	310	38%	253	34%	247	30%	-3.6	-3.1	10.3	0.30
	Emergency Room Admission	171	23%	214	26%	158	21%	188	23%	1.4	-4.5	7.4	0.63
	Home Health	51	7%	65	8%	38	5%	43	5%	1.0	-2.5	4.4	0.59
	Inpatient Rehabilitation Facility	9	1%	7	1%	4	1%	5	1%	-0.4	-1.7	0.8	0.50
	Skilled Nursing Facility	19	3%	23	3%	11	1%	11	1%	0.4	-1.6	2.4	0.70
	Psychiatric Hospital	3	0%	5	1%	1	0%	0	0%	0.3	-0.4	1.1	0.37
	Long-term Care Hospital	3	0%	0	0%	0	0%	0	0%	-0.4	-0.8	0.0	0.07
	Institutional Nursing Facility*	21	3%	22	3%	10	1%	10	1%	0.0	-2.0	2.0	0.99
	No Institutional Care	454	61%	498	61%	486	65%	566	70%	-4.1	-10.8	2.7	0.24
	No Post-acute Care	356	48%	383	47%	404	54%	439	54%	-0.4	-7.4	6.6	0.91

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.25	1.79	2.34	1.80	2.17	1.66	2.08	1.72	0.18	-0.06	0.43	0.15

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

*Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Exhibits 103c & 103d: Characteristics of BPCI and Comparison Beneficiaries with an Urgent/Emergent Hospitalization for CABG, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI CABG Hospital EIs				Matched Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	108	12%	110	13%	85	10%	90	9%	-0.2	-4.4	3.9	0.91
	65-79	635	70%	591	68%	637	70%	611	70%	-2.1	-8.1	4.0	0.50
	80+	162	18%	167	19%	183	20%	167	20%	2.3	-2.9	7.5	0.38
Gender	Female	283	31%	270	31%	297	33%	252	33%	3.6	-2.5	9.7	0.24
Medicaid and Disability	% Eligible for Medicaid	175	19%	175	20%	130	14%	102	14%	3.4	-1.4	8.3	0.17
	% Disability, no ESRD	123	14%	121	14%	106	12%	102	12%	0.3	-4.1	4.7	0.89
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	180	20%	168	19%	129	14%	121	14%	-0.2	-5.1	4.7	0.93
	Emergency Room Admission	206	23%	257	30%	199	22%	197	22%	6.1	0.5	11.8	0.03
	Home Health	44	5%	50	6%	42	5%	38	5%	1.2	-1.7	4.0	0.42
	Inpatient Rehabilitation Facility	5	1%	7	1%	3	0%	3	0%	0.2	-0.7	1.2	0.62
	Skilled Nursing Facility	7	1%	13	1%	17	2%	15	2%	0.9	-0.7	2.5	0.28
	Psychiatric Hospital	4	0%	3	0%	4	0%	0	0%	0.3	-0.4	1.1	0.36
	Long-term Care Hospital	0	0%	1	0%	0	0%	2	0%	-0.1	-0.5	0.3	0.56
	Institutional Nursing Facility*	8	1%	13	1%	19	2%	14	2%	1.1	-0.5	2.7	0.18
	No Institutional Care	715	79%	690	79%	767	85%	742	85%	-0.2	-5.3	4.8	0.92
	No Post-acute Care	567	63%	492	57%	610	67%	574	67%	-4.7	-11.0	1.6	0.15

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	1.91	1.45	2.05	1.68	1.79	1.46	1.79	1.45	0.14	-0.06	0.34	0.18

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

*Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

h. Quality, patient mix, and outlier payments

We conducted sensitivity analyses to determine if the differential proportion of anchor stays that received outlier payments may help explain the BPCI impact of increased pre-episode payments, post-episode payments, and mortality for urgent/emergent episodes. Patients that trigger outlier payments tend to be more complicated, with longer hospitalizations and higher mortality than patients who do not trigger outlier payments.[83] So, outlier payments may indicate unaccounted for severity differences between the BPCI and comparison groups. Indeed, the CABG TEP indicated that hospitals 'cherry pick' less expensive CABG patients, although the BPCI participants have stronger incentives than non-participants to try to avoid more expensive patients and favor the less expensive. Further, selecting patients would be more feasible for planned admissions, not urgent/emergent CABG episodes. If, however, the BPCI participants were building a cardiac program, they might receive the more severe patients. Outlier payments may also indicate a decline in the quality of inpatient care that resulted in the need for a longer stay or more intensive interventions.

We controlled for the presence of an outlier payment in the risk adjustment models for pre-episode payments, post-episode payments, and mortality for urgent/emergent episodes. After controlling for outlier payments, the BPCI impact on payments for the 120- and 180-days post-bundle periods and 30-day mortality rate were no longer statistically significant at the 0.10 level. This suggests that the relative increase in 120- and 180-day post-bundle payments and 30-day mortality rate were because of the relative increase in BPCI episodes with outlier payments. However, even after controlling for outlier payments, there were still statistically significant increases in the 30-day pre-bundle, 30-day post-bundle, and 90-day post-bundle payments relative to the change for the comparison group. We cannot establish whether the increase in outlier payments is due to a change in patient mix not currently captured by our claim-based risk factors or lower quality of care, or both.[84]

i. Conclusion

Results differ markedly between elective and urgent/emergent CABG episodes. In general, there were few statistically significant changes in Medicare payments and quality of care for elective episodes. In contrast, there were several notable results for BPCI urgent/emergent episodes. There was a statistically significant increase in Medicare payments for the anchor hospital stay. The change in anchor stay payments contributed to the relative increase (which was not statistically significant) in Medicare payments for the anchor stay plus 90-day PDP. There were also statistically significant relative increases in pre- and post-bundle spending for urgent/emergent episodes initiated by BPCI hospitals compared to matched comparison episodes. With regards to quality outcomes, there was a statistically significant increase in the 30-day mortality rate for BPCI urgent/emergent episodes relative to the change for the comparison group, but the difference in mortality was not statistically significant when measured over the 90-day post-discharge period. The proportion of BPCI episodes incurring outlier payments increased from 8.8% in the baseline to 13.4% in the intervention period; compared to little change in the proportion of comparison episodes which remained around 14%.

[83] Baser et al. (2009). Outlier Payments For Cardiac Surgery And Hospital Quality. *Health Affairs, 28*(4), 1154-1160.
[84] Baser et al. (2009). Outlier Payments For Cardiac Surgery And Hospital Quality. *Health Affairs, 28*(4), 1154-1160.

The increase in the proportion of BPCI episodes that incurred outlier payments was associated with the increase in mortality and 30- and 90-day post-bundle spending, however it was not associated with the increase in pre-bundle and 120- and 180-day post-bundle spending. These mixed results raise questions about potential changes in patient mix and quality of care for BPCI episodes. We will continue to examine these potential effects of BPCI on CABG episodes to be able to draw more definitive conclusions.

6. Gastrointestinal Hemorrhage

Medicare beneficiaries received 2,465 gastrointestinal (GI) hemorrhage episodes of care in 43 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015). BPCI had no statistically significant impact on Medicare payments for the initial hospital stay at participating providers and services during the subsequent 90-day post-discharge period (PDP).[85] At the same time, quality of care was the same between the two groups, as measured by unplanned readmission rates, emergency department visits, and mortality, with the exception of a decline in 30-day readmissions. Additional analyses with more time and episode experience under the initiative will reveal whether BPCI participation will affect Medicare payments for these episodes.

a. Participants

During the first eight quarters of the BPCI initiative, 43 hospitals (10% of all Model 2 BPCI hospitals) participated in the GI hemorrhage episode, all of which chose 90-day episodes. There were 2,465 GI hemorrhage episodes initiated by these providers during the BPCI initiative (approximately 2% of episodes across the 48 BPCI clinical episodes). Because a large share of participants joined BPCI during the latter part of the evaluation period, the GI hemorrhage results are based on an average of three quarters of experience. Approximately 58% of the hospitals began participating in GI hemorrhage in either Q2 or Q3 2015. Furthermore, of the 43 hospitals, nine stopped participating in the GI hemorrhage episode within the first eight quarters of the initiative.

Compared to hospitals that did not participate in BPCI, hospitals that participated in the GI hemorrhage episode were larger, and more likely to be non-profit, to have higher resident-to-bed ratios and to be located in urban areas (see Exhibits 104a & 104b). Prior to joining BPCI, participating hospitals had more GI hemorrhage discharges and higher average 2011 standardized Part A payments for GI hemorrhage episodes during the anchor hospitalization plus 90-days PDP than non-participants. Matched comparison hospitals were identified for 38 of the 43 BPCI-participating hospitals in the sample.[86] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals.

[85] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf

[86] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were five hospitals that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the GI hemorrhage episode or had fewer than five relevant discharges in 2011.

Exhibits 104a & 104b: Characteristics of BPCI-participating Hospitals with Gastrointestinal Hemorrhage Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals Model 2, Q4 2013 - Q3 2015

Characteristic		All GI Hemorrhage BPCI Hospitals (N=43)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched GI Hemorrhage BPCI Hospitals (N=38)		Comparison Hospitals (N=535+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	42	98%	1,594	57%*	321	77%*	37	97%	514	96%
	Government	0	0%	542	20%*	32	8%*	0	0%	0	0%
	For-Profit	1	2%	638	23%*	66	16%*	1	3%	21	4%
Urban/Rural	Urban	39	91%	1,902	69%*	387	92%	32	89%	476	89%
Part of Chain	Yes	21	49%	1,469	53%	218	52%	13	53%	284	53%

Characteristic	All GI Hemorrhage BPCI Hospitals (N=43)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched GI Hemorrhage BPCI Hospitals (N=38)	Comparison Hospitals (N=535+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	279	175*	311	285	312
Medicare Days Percent	40%	42%	39%	39%	40%
Resident-to-bed ratio	0.16	0.05*	0.12	0.17	0.18
Disproportionate Share Percent	27%	29%	27%	28%	29%
Number of GI Hemorrhage Discharges, 2011	94	53*	101	94	101
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, GI Hemorrhage, 2011	**$15,914**	**$15,102***	**$15,209***	**$15,878**	**$15,834**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All gastrointestinal hemorrhage BPCI" hospitals. "Comparison" hospitals are compared to "Matched gastrointestinal hemorrhage BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 324.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. A comparison of the GI hemorrhage patient population of BPCI participants from baseline to intervention relative to the same time periods for the GI hemorrhage patient population of the comparison providers does not suggest that BPCI providers treated a healthier patient population during the intervention period (see Exhibits 105a & 105b).

Exhibits 105a & 105b: Characteristics of BPCI and Comparison Patients with a Hospitalization for Gastrointestinal Hemorrhage, Model 2, Q4 2013- Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	192	7%	191	8%	209	8%	199	9%	0.3	-1.9	2.4	0.80
	65-79	969	35%	837	37%	1,031	38%	885	39%	0.2	-3.6	4.0	0.94
	80+	1,573	58%	1,246	55%	1,494	55%	1190	52%	-0.4	-4.3	3.5	0.83
Gender	Female	1,567	57%	1,247	55%	1,520	56%	1242	55%	-1.5	-5.4	2.4	0.45
Medicaid and Disability	% Eligible for Medicaid	513	19%	438	19%	635	23%	461	20%	3.5	0.3	6.6	0.03
	% Disability, no ESRD	221	8%	210	9%	242	9%	214	9%	0.6	-1.6	2.8	0.60
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	934	34%	825	36%	941	34%	811	36%	0.9	-2.9	4.6	0.65
	Emergency Room Admission	760	28%	741	33%	777	28%	725	32%	1.3	-2.3	4.9	0.47
	Home Health	584	21%	543	24%	638	23%	551	24%	1.6	-1.7	4.9	0.34
	Inpatient Rehabilitation Facility	57	2%	36	2%	48	2%	42	2%	-0.6	-1.6	0.5	0.27
	Skilled Nursing Facility	401	15%	339	15%	321	12%	293	13%	-0.9	-3.6	1.8	0.51
	Psychiatric Hospital	20	1%	22	1%	23	1%	17	1%	0.3	-0.4	1.0	0.36
	Long-term Care Hospital	9	0%	11	0%	16	1%	16	1%	0.0	-0.5	0.6	0.90
	Institutional Nursing Facility**	586	21%	436	19%	471	17%	389	17%	-2.1	-5.2	0.9	0.17
	No Institutional Care	1,720	63%	1,405	62%	1,726	63%	1424	63%	-0.6	-4.4	3.2	0.75
	No Post-acute Care	1,286	47%	971	43%	1,220	45%	984	43%	-3.0	-6.9	0.9	0.13

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.47	2.48	2.63	2.60	2.36	2.35	2.53	2.51	-0.01	-0.21	0.18	0.90

DiD=difference-in-differences.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

c. Medicare Payments and Utilization

Total Medicare payments ($2015) for the inpatient hospitalization plus all care delivered during the 90-day post-discharge period did not change for BPCI hospitals relative to episodes from a matched comparison group of hospitals not participating in BPCI (see Exhibit 106).[87] Total payments declined by 5% for BPCI hospitals and by 3% for the matched comparison group from before the initiative.

Exhibit 106: Impact of BPCI on Total Payments for Inpatient Hospitalization and 90-day PDP ($2015) for Gastrointestinal Hemorrhage Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=2,248)		Comparison (N=2,255)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Total allowed payment amount, inpatient plus 90-day PDP	$21,500	$20,484	$21,107	$20,458	-$367	-$1,828	$1,095	0.62

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; PDP = post-discharge period.

These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

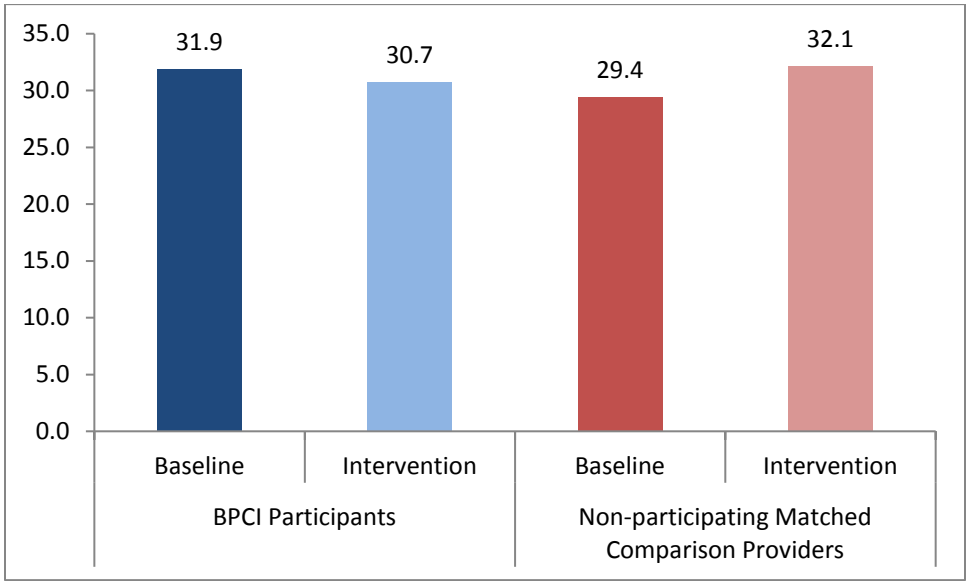
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

The number of institutional post-acute care (PAC) days among institutional PAC users during the 90-day PDP declined by 3.8 days (p=0.07) for BPCI episodes relative to comparison users (see Exhibit 107).[88] The number of institutional PAC days decreased from 31.9 to 30.7 for BPCI GI hemorrhage patients and increased from 29.4 to 32.1 for the comparison group. There were no other statistically significant relative differences among the other payment and PAC utilization outcomes.

[87] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

[88] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Exhibit 107: Impact of BPCI on Institutional PAC Days among Institutional PAC users for Gastrointestinal Hemorrhage Episodes, Model 2, Q4 2011-Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

We monitor changes in payments before the bundle begins and after the bundle period ends, because BPCI participants have incentives to shift payments outside of the bundle to reduce their episode payments below their target. We do not see a change in pre-bundle payments or payments during the 1-30 day post-bundle measurement period, but we do see a *decrease* in post-bundle payments for the 1-90, 1-120, and 1-180 post-bundle periods (see Exhibit 108). However, the results involving longer time periods are affected by changes in the composition of the BPCI subsample because the post-bundle outcomes are measured with a one to two quarter lag in order to allow sufficient time for the claims to be submitted. Thus, the hospitals that joined BPCI in the most recent quarter (Q3 2015) are not included in these measures, and the intervention period does not include any episodes initiated in Q3 2015.[89]

[89] The 1-90 post-bundle outcome is measured with a one quarter lag and therefore does not include hospitals that joined BPCI after Q2 2015, and does not include episodes initiated by any hospital after Q2 2015. The 1-120 and 1-180 post-bundle measures use a two quarter lag and therefore do not include hospitals that joined after Q1 2015, and the intervention period does not include episodes initiated after Q1 2015.

Exhibit 108: Impact of BPCI on Payments ($2015) for Services After the Bundle, Gastrointestinal Hemorrhage Episodes, Model 2, Q4 2011-Q3 2015

Post bundle Time period	Number of intervention episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
1-30 days	1,678	1,687	$3,055	$2,758	$2,918	$3,104	-$484	-$1,121	$153	0.14
1-90 days	1,359	1,381	$9,181	$7,462	$8,822	$8,887	-$1,784	-$3,291	-$277	0.02
1-120 days	1,060	1,069	$12,195	$9,633	$11,448	$11,518	-$2,631	-$4,749	-$514	0.02
1-180 days	971	964	$16,842	$13,346	$15,894	$16,308	-$3,909	-$6,522	-$1,297	<0.01

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015 for the 1-30 post-bundle period, through Q2 2015 for the 1-90 post-bundle period, and through Q12015 for the 1-120 and 1-180 post-bundle periods. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

d. Quality of Care

There was no difference in the quality of care for gastrointestinal hemorrhage Medicare patients at BPCI-participating hospitals and patients at comparison hospitals, except for a decrease in unplanned readmission rates within the first 30 days post-discharge for BPCI hospitals relative to the comparison group. The share of BPCI patients with a readmission within the 30-day PDP decreased 2.5 percentage points relative to patients in the comparison group (p=0.08). The share of BPCI patients with a readmission decreased from 16.1% to 14.1%, while the share of comparison group patients with a readmission increased from 13.4% to 13.9%. The relative change in readmission rate was no longer significant when measured during the 90-day PDP. There were no significant differences in measures of functional improvement as measured by patient assessments.

e. Conclusions

During the first eight quarters of the initiative, there were few statistically significant changes in Medicare payments, utilization of health care services, and quality of care for the gastrointestinal hemorrhage episodes in Model 2. Additional analyses with more time and episode experience under the initiative may be needed to determine whether BPCI has had an impact.

7. Hip and Femur Procedures

Medicare beneficiaries received 3,077 surgical hip and femur episodes of care in 74 BPCI-participating acute care hospitals in the first eight quarters of the initiative (October 2013 through September 2015).[90] Under BPCI, the hip and femur procedures except major joint replacement of the lower extremity episode (hip and femur) includes only surgical procedures.

Based on the subset of BPCI hip and femur episodes for which we found comparison providers and patients, average Medicare payments for the hospitalization and the subsequent 90 days of care did not change from baseline to intervention relative to the matched sample of episodes from hospitals that did not participate in BPCI.[91] Payments for home health care increased for BPCI episodes because the share of episodes that received home health care rose. There was no difference in the change in emergency department visits and mortality. There was a relative increase in readmission rates among BPCI episodes during the 90-day post-discharge period (PDP). Even though the readmission rate declined for both BPCI and comparison episodes, the decline was greater among comparison episodes. Overall, the first two years of evaluation suggest that BPCI has had little effect on the cost or quality of care during hip and femur episodes.

a. Participants

During the first eight quarters of the BPCI initiative, 74 hospitals (18% of all Model 2 BPCI hospitals) participated in hip and femur episodes; all but one hospital chose 90-day episodes. These hospitals initiated 3,077 hip and femur episodes during the BPCI initiative (approximately 2% of episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, hip and femur episode results are based on hospitals with an average of three quarters of experience. Over half (58%) of the BPCI hospitals began participating in hip and femur procedure episodes in Q2 2015 or Q3 2015; 27% of hip and femur participating hospitals joined in Q3 2015. Six hospitals stopped participating in the hip and femur episode in the first eight quarters of the initiative. Of the 74 hospitals participating in hip and femur episodes, 68 (92%) also participated in lower extremity joint replacement episodes.

Compared to hospitals that were not participating in the initiative, BPCI episode initiators (EIs) with hip and femur episodes were less likely to be government-owned and more likely to be located in urban areas (Exhibits 109a & 109b). Additionally, BPCI hospitals had higher bed counts and a smaller percentage of Medicare days. Prior to the initiative, BPCI-participating hospitals had relatively more discharges for hip and femur procedures. Standardized 2011 Part A payments during the hospitalization plus the 90-day PDP also tended to be higher for hip and femur discharges from BPCI hospitals.

Compared with all Model 2-participating hospitals, those with hip and femur episodes were more likely to be for-profit and urban. Hip and femur participants also had lower resident-to-bed

[90] Because participants were allowed to join BPCI over an extended period, these hospitals had an average of three quarters of experience in the Model 2 hip and femur episode.

[91] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do not take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf

ratios. Standardized Medicare Part A payments during the hospitalization and 90-day PDP for hip and femur discharges were higher among participating hospitals with hip and femur episodes than among all Model 2 BPCI hospitals (Exhibits 109a & 109b).

Matched comparison hospitals were identified for 66 of the 74 BPCI-participating hospitals.[92] There were no statistically significant differences between BPCI hospitals with hip and femur episodes and the comparison group with respect to post-acute care (PAC) use, readmission rates, or Medicare payments for the hospitalization and the 90-day PDP.

[92] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. Six participating hospitals were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in hip and femur episodes or had fewer than five relevant discharges in 2011 or 2012. There were two other participants for which there was not a match within the caliper (5% of the standard deviation of the log odds propensity score).

Exhibits 109a &109b: Characteristics of BPCI-participating Hospitals with Hip and Femur Episodes compared with Non-participating Hospitals, All BPCI-participating Hospitals and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Hip and Femur BPCI Hospitals (N=74)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched Hip and Femur BPCI hospitals (N=66)		Comparison Hospitals (N=865)+	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	48	65%	1,594	57%	321	77%*	44	67%	570	66%
	Government	5	7%	542	20%*	32	8%	3	5%	45	5%
	For-Profit	21	28%	638	23%	66	16%*	19	29%	250	29%
Urban/Rural	Urban	73	99%	1,902	69%*	387	92%*	65	98%	853	99%
Part of Chain	Yes	35	47%	1,469	53%	218	52%	31	47%	438	51%

Characteristic	All Hip and Femur BPCI Hospitals (N=74)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched Hip and Femur BPCI hospitals (N=66)	Comparison Hospitals (N=865)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	294	175*	311	284	301
Medicare Days Percent	38%	42%*	39%	39%	39%
Resident-to-bed Ratio	0.08	0.05	0.12*	0.09	0.10
Disproportionate Share Percent	28%	29%	27%	28%	29%
Number of Hip and femur Discharges, 2011	59	33*	62	59	61
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, Hip and femur, 2011	**$38,235**	**$36,275***	**$36,661***	**$38,335**	**$38,211**

PDP=post-discharge period

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All hip and femur BPCI" hospitals. "Comparison" hospitals are compared to "Matched hip and femur BPCI" hospitals.

** BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+These columns represent the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. There are 469 unique matched comparison providers in this clinical episode.

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI-participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. Throughout the baseline and intervention periods, the hip and femur patients in BPCI and comparison hospitals were similar across key metrics (Exhibits 110a & 110b). There was, however, a statistically significant

increase in the share of patients who were eligible for Medicaid ($p<0.01$) in BPCI participating hospitals, which may be an indicator of increased patient complexity. With the exception of a statistically significant decrease in the percent of patients that had been in an inpatient rehabilitation facility (IRF) ($p=0.09$), there were no significant differences with respect to the change in health care utilization in the six months prior to the episode.

Exhibits 110a & 110b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Hip and Femur Procedure, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hip and Femur Hospitals				Comparison Hospitals				DiD	Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Age	20-64	168	5%	170	6%	199	6%	160	6%	1.3	-0.4	2.9	0.12
	65-79	884	27%	799	28%	899	27%	860	30%	-1.7	-4.9	1.5	0.29
	80+	2,270	68%	1,854	66%	2,224	67%	1,803	64%	0.4	-2.9	3.8	0.81
Gender	Female	2,531	76%	2,131	75%	2,453	74%	2,098	74%	-1.2	-4.3	1.9	0.45
Medicaid and Disability	% Eligible for Medicaid	602	18%	579	21%	618	19%	431	15%	5.7	3.0	8.5	<0.01
	% Disability, no ESRD	197	6%	185	7%	219	7%	182	6%	0.8	-1.0	2.5	0.38
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	763	23%	671	24%	709	21%	612	22%	0.5	-2.5	3.4	0.76
	Emergency Room Admission	876	26%	818	29%	899	27%	851	30%	-0.5	-3.7	2.7	0.77
	Home Health	749	23%	673	24%	738	22%	671	24%	-0.3	-3.2	2.7	0.86
	Inpatient Rehabilitation Facility	76	2%	52	2%	76	2%	78	3%	-0.9	-2.0	0.1	0.09
	Skilled Nursing Facility	422	13%	386	14%	357	11%	334	12%	-0.1	-2.4	2.2	0.92
	Psychiatric Hospital	34	1%	34	1%	20	1%	31	1%	-0.3	-1.0	0.4	0.37
	Long-term Care Hospital	22	1%	19	1%	19	1%	16	1%	0.0	-0.5	0.6	0.96
	Institutional Nursing Facility*	652	20%	571	20%	549	17%	501	18%	-0.6	-3.4	2.1	0.66
	No Institutional Care	2,446	74%	2,079	74%	2,517	76%	2,122	75%	0.6	-2.5	3.7	0.70
	No Post-acute Care	1,684	51%	1,414	50%	1,772	53%	1,409	50%	2.8	-0.7	6.4	0.12

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.54	1.87	2.66	2.05	2.50	1.87	2.56	2.00	0.06	-0.08	0.19	0.42

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; ESRD = end stage renal disease.

These estimates are developed using a difference-in-differences (DiD) framework, which quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. Changes are observed before and after the intervention began.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention) for BPCI EIs and matched comparison providers.

Note: "Count" represent the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values

***** Institutional Nursing Facility utilization is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare Payments and Utilization

There is no indication that BPCI had a statistically significant impact on total Medicare payments ($2015) for hip and femur episodes. Though total payments for services during the inpatient stay and 90-day PDP declined among BPCI episodes from baseline to intervention, payments for the comparison group episodes also declined. The difference in the change was not significant (Exhibit 111).

Exhibit 111: Impact of BPCI on Total Payments for Inpatient Hospitalization and 90-day PDP ($2015), Hip and Femur Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=2,789)		Comparison (N=2,791)		DiD Estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total allowed payment amount, IP plus 90-day PDP	$47,483	$44,766	$47,780	$45,228	-$164	-$1,650	$1,322	0.83

DiD = difference-in-differences, LCI = lower confidence interval at the 5% level, UCI = upper confidence interval at the 5% level, IP = inpatient, PDP = post-discharge period.

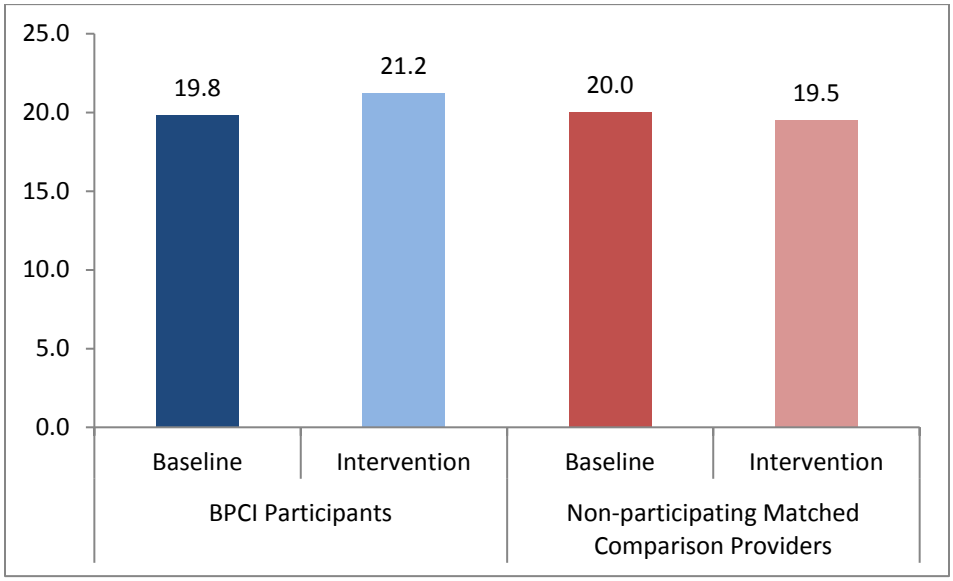
Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors. Medicare payments are expressed in 2015 dollars, the result of adjusting actual dollar amounts based on changed in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

With respect to PAC use, BPCI and comparison providers exhibited little change in the proportion of beneficiaries discharged to PAC (of all beneficiaries) or the proportion of beneficiaries discharged to institutional PAC (of all discharged to any PAC). In the intervention period, BPCI providers continued to discharge the vast majority of their hip and femur patients to institutional PAC, particularly to skilled nursing facilities (SNFs). However, there is evidence that BPCI patients altered their home health agency (HHA) utilization in the 90 days following the anchor hospitalization. Over half of BPCI hip and femur episodes received some home health care during the 90-day PDP and this share increased from baseline to intervention relative to the comparison group (6.1 percentage points).[93] The number of home health visits also increased for BPCI episodes relative to comparison episodes (Exhibit 112). From baseline to intervention, BPCI patients who received home health care had 1.9 more visits than comparison patients ($p<0.01$).

[93] This DiD estimate is based on unadjusted episode utilization. It does not account for differences in patient mix. BPCI average increased from 53.6% to 58.3% from baseline to intervention while the comparison average decreased (55.0% to 53.5%).

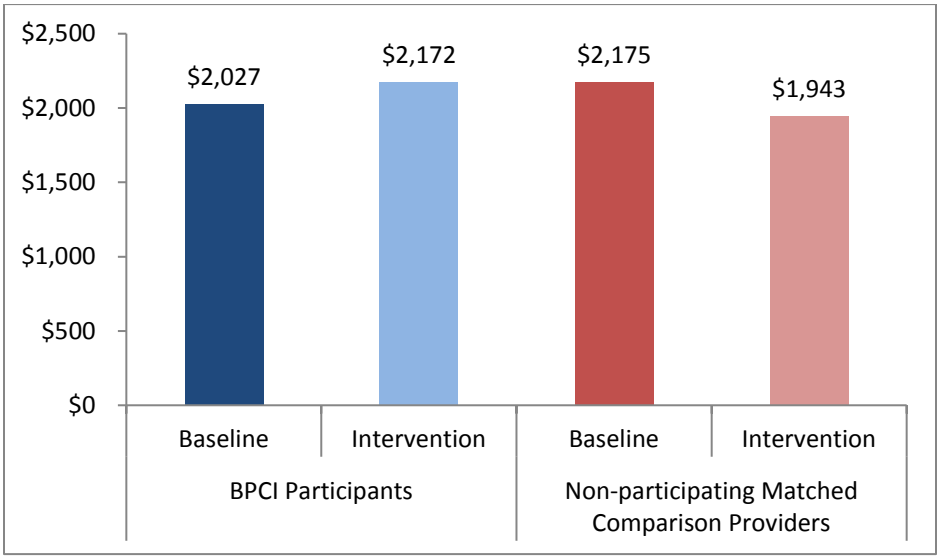
Exhibit 112: Impact of BPCI on Number of Home Health Visits, Hip and Femur Episodes, Model 2, Q4 2011 – Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Consistent with these changes in utilization during the 90-day PDP, Medicare Part A and B payments for home health care increased $377 more for BPCI episodes than for comparison episodes between the baseline and intervention periods (Exhibit 113, $p<0.01$).

Exhibit 113: Impact of BPCI on Home Health Payments ($2015), Hip and Femur Episodes, Model 2, Q4 2011 – Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Notes: Medicare payments for the anchor hospitalization plus 90-day PDP are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and trended to 2015.

There is no indication that BPCI participants attempted to reduce episode costs by changing the timing of services so that they are delivered before the anchor hospitalization or after the end of the bundle period. The change in Medicare payments during the post- and pre-bundle period from baseline to intervention was not significantly different between BPCI and comparison episodes.

d. Quality of Care

There was no difference in the change in emergency department (ED) use and mortality between BPCI and comparison hip and femur episodes. In the case of readmissions, comparison episodes had a larger decline in the 90-day unplanned readmission rate from baseline to intervention, resulting in a relative increase of 2.7 percentage points for BPCI episodes (Exhibit 114, p=0.06). Readmission rates within 30 days of inpatient discharge, during which hospitals can likely exercise greater influence, declined more for BPCI than the comparison group from baseline to intervention, though the result is not statistically significant. There were also no statistically significant differences in assessment-based functional status measures between BPCI and comparison patients who received institutional PAC and home health care.

Exhibit 114: Impact of BPCI on Unplanned Readmission Rates during the Post-discharge Period, Hip and Femur Episodes, Model 2, Q4 2011-Q3 2015

Measurement Period	Intervention Episodes (N)		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
30 days post-discharge	2,822	2,822	12.7%	11.3%	11.3%	10.4%	-0.5	-2.7	1.8	0.69
90 days post-discharge	2,798	2,801	22.2%	21.9%	21.4%	18.4%	2.7	-0.1	5.6	0.06

DiD = difference-in-differences, 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 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Note: Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

e. Conclusion

Results from the first two years of BPCI indicate that the initiative did not have a statistically significant impact on total Medicare payments for hip and femur episodes. There was, however, a relative increase in payments for home health care during the 90-day PDP. This was driven by a difference in the proportion of episodes that received home health services between the BPCI and comparison providers. We also observed a relative increase in readmission rates during the 90 days following the anchor hospitalization. Though BPCI providers decreased readmissions from baseline to intervention, they were unable to match the improvement observed among the comparison group. Given that these findings are drawn from 66 hospitals with an average of three quarters of experience in the hip and femur bundle, results should be interpreted with caution.

8. Major Joint Replacement of the Lower Extremity

Medicare beneficiaries received 53,622 episodes of care for major joint replacement of the lower extremity (MJRLE) in 294 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015).[94] MJRLE was the only episode within Model 2 that had a significant decline in the average Medicare payments for the inpatient hospitalization and the services during the 90-day post-discharge period relative to the comparison group. Medicare payments declined by $1,273 from the baseline to the intervention period for BPCI episodes relative to the comparison group.

Within this clinical episode we distinguish between planned joint replacements not due to fracture (MJRLE-NF) and emergent replacements due to fracture (MJRLE-F), because care needs and types of patients are quite different according to members of the Technical Expert Panel on MJRLE episodes. Total episode payments declined by $1,105 for non-fracture and $1,924 for fracture episodes. The decline in payments was generally due to reductions in institutional post-acute care (PAC) payments, particularly skilled nursing facility (SNF) payments.

In general, BPCI has not had a detrimental impact on quality of care. There were no statistically significant changes in readmission rates, emergency department visits, or mortality rates during the 90 days following discharge from the anchor hospitalization. For some patients who received institutional PAC, there were relative declines in measures of functional status. Beneficiary survey results indicated relative improvements in functional status for BPCI patients from before to after their episode of care. In particular, self-reported mobility improved for BPCI patients.

Under BPCI, hospitals have strong incentives to admit patients who will likely be less resource-intensive. Because non-fracture MJRLE are planned procedures, hospitals have considerable ability to select patients for admission. There is some evidence that hospitals selected a healthier mix of non-fracture patients under BPCI.

Some hospitals were more successful than others in reducing MJRLE episode payments below their target, as indicated by their Net Payment Reconciliation Amounts (NPRAs). The BPCI hospitals that achieved the greatest NPRA reduced the proportion of their patients discharged to institutional PAC more than hospitals that were less successful. The top performers also increased the use of home health agency (HHA) care more than bottom performers. It is possible that the

[94] Because participants were allowed to join BPCI over an extended period, these participants had an average of three quarters of experience in Model 2.

substitution of HHA care for institutional PAC among top performers may have been easier to achieve because of changes in their mix of patients. Under BPCI, top performers treated healthier patients, younger patients, and a greater proportion of non-fracture patients.

a. Participants

During the first eight quarters of the BPCI initiative, 294 hospitals (70% of all Model 2 BPCI hospitals) participated in the MJRLE episode, 88% of which chose 90-day episodes. There were 53,622 MJRLE episodes initiated by these providers during the BPCI initiative (approximately 36% of episodes across the 48 BPCI clinical episodes). The MJRLE results are based on an average of three quarters of experience because approximately 36% of the hospitals began participating in MJRLE in Q3 2015 and 66% joined in either Q2 or Q3 2015. Of the 294 hospitals, 20 stopped participating in MJRLE within the first eight quarters of the initiative.

Compared to hospitals not participating in BPCI, hospitals that participated in the MJRLE episode had more beds and a higher resident-to-bed ratio on average (Exhibits 115a & 115b). Participating hospitals were also more likely to be in urban areas and non-profit. Prior to joining BPCI, participating hospitals had more MJRLE discharges. Some interviewees noted that one reason for choosing this episode was its relatively high volume, which they expected to be sufficient to allow them to see results.[95]

Matched comparison hospitals were identified for 279 of the BPCI-participating hospitals in the sample.[96] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals.

[95] See Methods section for more information on the BPCI participant site visits and interviews.

[96] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were ten participants that were not included in the analysis because either they were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the MJRLE episode or had fewer than five relevant discharges in 2011 or 2012. There were three other participants for which there was not a match within the caliper.

Exhibits 115a & 115b: Characteristics of BPCI-participating Hospitals with Major Joint Replacement of the Lower Extremity Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All MJRLE BPCI Hospitals (N=292)**		Non-participating Hospitals (N=2,774)		Matched MJRLE BPCI hospitals (N=279)		Comparison Hospitals (N=3,771+)	
		N	%	N	%	N	%	N	%
Ownership	Non-Profit	219	75%	1,594	57%*	209	75%	2,800	74%
	Government	18	6%	542	20%*	18	6%	271	7%
	For-Profit	55	19%	638	23%	52	19%	700	19%
Urban/Rural	Urban	274	94%	1,902	69%*	262	94%	3,533	94%
Part of Chain	Yes	143	49%	1,469	53%	138	49%	1,854	49%

Characteristic	All MJRLE BPCI Hospitals (N=292)**	Non-participating Hospitals (N=2,774)	Matched MJRLE BPCI hospitals (N=279)	Comparison Hospitals (N=3,771+)
	Mean	Mean	Mean	Mean
Bed Count	304	175*	302	313
Medicare Days Percent	40%	42%	40%	40%
Resident-to-bed Ratio	0.11	0.05*	0.11	0.11
Disproportionate Share Percent	26%	29%	26%	27%
Number of MJRLE Discharges, 2011	202	105*	194	205
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, MJRLE, 2011	**$23,862**	**$23,836**	**$23,917**	**$23,991**

PDP = post-discharge period

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" hospitals are compared to "All MJRLE BPCI" hospitals. "Comparison" hospitals are compared to "Matched MJRLE BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers were matched to multiple BPCI hospitals. The unique number of matched comparison providers is 1,052.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. The potential for patient selection is of particular concern for elective procedures, such as non-fracture MJRLEs. Hospitals and physicians reported that because non-fracture MJRLEs are typically scheduled surgeries, patients often receive pre-surgical education and decide upon the post-hospital discharge site of

care in consultation with their clinicians before the surgery. Additionally, beneficiaries receiving a scheduled lower joint replacement typically are healthier than the average Medicare beneficiary. In contrast, hip replacements due to a fracture are emergent procedures and typically do not allow for patient education or planning for the episode of care. Further, fractures more often involve frailer individuals who have more comorbidities. Because of these differences, hospitals have a greater ability to select non-fracture patients than fracture patients. Some BPCI participants that we interviewed acknowledged the ability to be more selective in admitting non-fracture patients. Interviewees discussed attempting to reduce patients' risk levels prior to surgery through various interventions, such as postponing surgery until the patient stopped smoking or lost weight, or requiring "pre-habilitation" classes, which could influence the patient mix.

A comparison of patient characteristics available in the claims suggests that BPCI participants may have selected healthier non-fracture patients after joining BPCI relative to the comparison group (Exhibits 116a & 116b). There was a statistically significant decrease in the proportion of BPCI non-fracture patients who had a hospital stay (-0.6 percentage points, p=0.09), a SNF stay (-0.6 percentage points, p<0.01), or an institutional nursing facility stay (-0.4 percentage points, p=0.03) during the six months prior to their index hospitalization.[97] There was also a statistically significant increase in the proportion with no institutional care (0.6 percentage points, p=0.06) in the six months prior to their index hospitalization (Exhibits 116c & 116d). Although these four reductions in the proportion of patients with prior health care use are statistically significant, they may not be clinically meaningful given the relatively small changes in each proportion. The consistent declines still raise concerns, however, that BPCI hospitals may be selecting healthier non-fracture patients.

There is no indication that BPCI participants selected healthier fracture patients (Exhibit 116c & 116d). In fact, for outcomes in which there was a statistically significant change, the BPCI patients were relatively less healthy, as evidenced by an increase in the percentage of BPCI fracture patients who used an inpatient rehabilitation facility (IRF) (p=0.01), and a decrease in the percentage who received no institutional care (p=0.09) in the six months prior to their index hospitalization, relative to comparison patients.

[97] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

Exhibits 116a & 116b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Major Joint Replacement of the Lower Extremity, Non-Fracture, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	3,702	10%	3,154	9%	3,612	9%	3,241	9%	-0.5	-1.1	0.1	0.11
	65-79	26,044	68%	2,5393	72%	26,076	68%	25,230	71%	0.5	-0.4	1.5	0.26
	80+	8,349	22%	6,921	20%	8,407	22%	6,997	20%	-0.1	-0.9	0.8	0.88
Gender	Female	24,783	65%	22,808	64%	24,450	64%	22,648	64%	-0.4	-1.4	0.6	0.40
Medicaid and Disability	% Eligible for Medicaid	4,644	12%	3,650	10%	4,864	13%	4,031	11%	-0.5	-1.2	0.2	0.14
	% Disability, no ESRD	34,501	12%	3,595	10%	4,435	12%	3,680	10%	-0.4	-1.1	0.2	0.21
Utilization Six Months Prior to Index Hospitalization*	Inpatient Acute Care Hospital	4,371	11%	3,762	11%	4,263	11%	3,857	11%	-0.6	-1.2	0.1	0.09
	Emergency Room Admission	5,675	15%	5,276	15%	5,757	15%	5,472	15%	-0.3	-1.1	0.4	0.37
	Home Health	3,372	9%	2,811	8%	3,506	9%	3,073	9%	-0.4	-1.0	0.2	0.19
	Inpatient Rehabilitation Facility	346	1%	284	1%	291	1%	226	1%	0.0	-0.2	0.2	0.83
	Skilled Nursing Facility	1,401	4%	1,066	3%	1,381	4%	1,243	4%	-0.6	-0.9	-0.2	0.00
	Psychiatric Hospital	95	0%	54	0%	103	0%	73	0%	0.0	-0.1	0.1	0.51
	Long-term Care Hospital	16	0%	17	0%	17	0%	15	0%	0.0	0.0	0.1	0.71
	Institutional Nursing Facility	1,472	4%	1,078	3%	1,471	4%	1,226	3%	-0.4	-0.8	0.0	0.03
	No Institutional Care	33,392	88%	31,468	89%	33,497	88%	31,341	88%	0.6	0.0	1.3	0.06
	No Post-acute Care	28,202	74%	26,570	75%	28,086	74%	26,283	74%	0.5	-0.4	1.4	0.27

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	0.94	1.20	0.91	1.19	0.92	1.19	0.90	1.20	-0.01	-0.04	0.01	0.37

DiD = difference-in-differences; ESRD = end stage renal disease.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. "Count" represents the numerator for the given characteristic. The '%' is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Exhibits 116c & 116d: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Major Joint Replacement of the Lower Extremity due to a Fracture, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	193	3%	154	3%	197	3%	180	3%	-0.4	-1.3	0.4	0.33
	65-79	1,848	27%	1,647	30%	1,900	27%	1,660	30%	0.5	-1.7	2.8	0.65
	80+	4,887	71%	3,705	67%	4,831	70%	3,666	67%	-0.1	-2.4	2.2	0.93
Gender	Female	5,141	74%	3,949	72%	5,075	73%	3,865	70%	0.6	-1.7	2.8	0.62
Medicaid and Disability	% Eligible for Medicaid	1,118	16%	774	14%	1,079	16%	801	15%	-1.1	-2.8	0.7	0.25
	% Disability, no ESRD	237	3%	181	3%	250	4%	205	4%	-0.2	-1.2	0.7	0.60
Utilization Six Months Prior to Index Hospitalization*	Inpatient Acute Care Hospital	1,578	23%	1,246	23%	1,533	22%	1,158	21%	0.9	-1.1	3.0	0.37
	Emergency Room Admission	1,885	27%	1,601	29%	1,920	28%	1,580	29%	0.9	-1.4	3.1	0.44
	Home Health	1,635	24%	1,223	22%	1,631	24%	1,239	23%	-0.3	-2.5	1.8	0.75
	Inpatient Rehabilitation Facility	149	2%	155	3%	170	2%	118	2%	1.0	0.2	1.7	0.01
	Skilled Nursing Facility	869	13%	674	12%	882	13%	623	11%	1.1	-0.5	2.8	0.18
	Psychiatric Hospital	76	1%	75	1%	88	1%	63	1%	0.4	-0.2	0.9	0.16
	Long-term Care Hospital	13	0%	17	0%	24	0%	24	0%	0.0	-0.2	0.3	0.82
	Institutional Nursing Facility	1,245	18%	946	17%	1,256	18%	874	16%	1.5	-0.4	3.4	0.13
	No Institutional Care	5,145	74%	4,081	74%	5,142	74%	4,182	76%	-1.9	-4.1	0.3	0.09
	No Post-acute Care	3,566	51%	2,854	52%	3,553	51%	2,842	52%	0.0	-2.5	2.5	0.99

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.24	1.85	2.54	1.93	2.55	1.92	2.47	1.87	0.08	-0.02	0.17	0.11

DiD = difference-in-differences; ESRD = end stage renal disease.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. "Count" represents the numerator for the given characteristic. The '%' is the numerator divided by the total number of episodes with non-missing values.

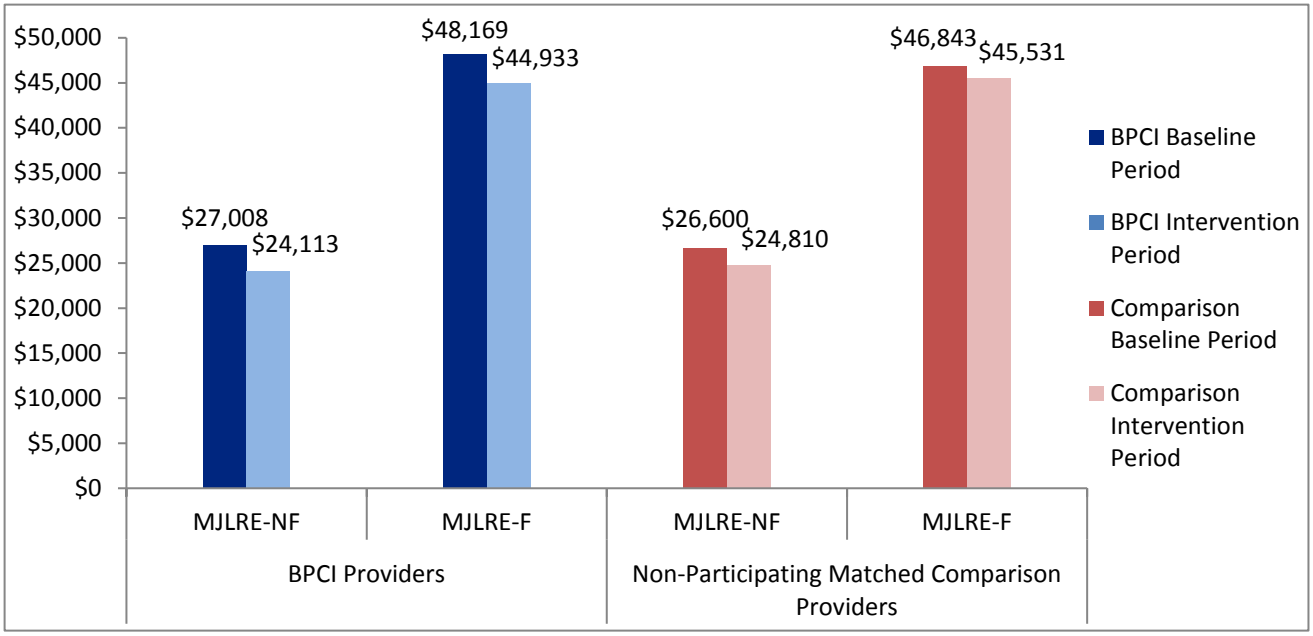
* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare payments and utilization

Total Medicare payments ($2015) for the inpatient stay and all care delivered during the 90-day post-discharge period (PDP) were $24,113 for MJRLE non-fracture BPCI episodes during the intervention, compared with $44,933 for fracture episodes, indicating the differences in the severity and care needs of patients with a non-fracture and fracture joint replacement.[99]

Total Medicare payments for the inpatient stay and all care delivered during the 90-day PDP were statistically significantly lower during the intervention than the baseline for both non-fracture and fracture BPCI episodes compared to episodes from a matched comparison group of providers (Exhibit 117). Based on the 35,126 non-fracture episodes in the intervention period, total payments ($2015) declined by $1,105 more for BPCI episodes than episodes from matched comparison hospitals ($p<0.01$). Total payments for BPCI non-fracture episodes declined by 11% from $27,008 in the baseline to $24,113 in the intervention period. Total payments for comparison episodes were $26,000 during the baseline, and declined by 7% to $24,810 in the intervention period. For fracture episodes, total payments declined by $1,924 more for BPCI episodes than comparison episodes ($p<0.01$). Total payments for BPCI fracture episodes were $48,169 during the baseline period and declined by 7% to $44,933 in the intervention period. For comparison fracture episodes, total payments declined by 3%, from $46,843 in the baseline to $45,531 in the intervention period.

Exhibit 117: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day PDP, Non-Fracture and Fracture MJRLE, Model 2, Q4 2013 - Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison hospitals.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

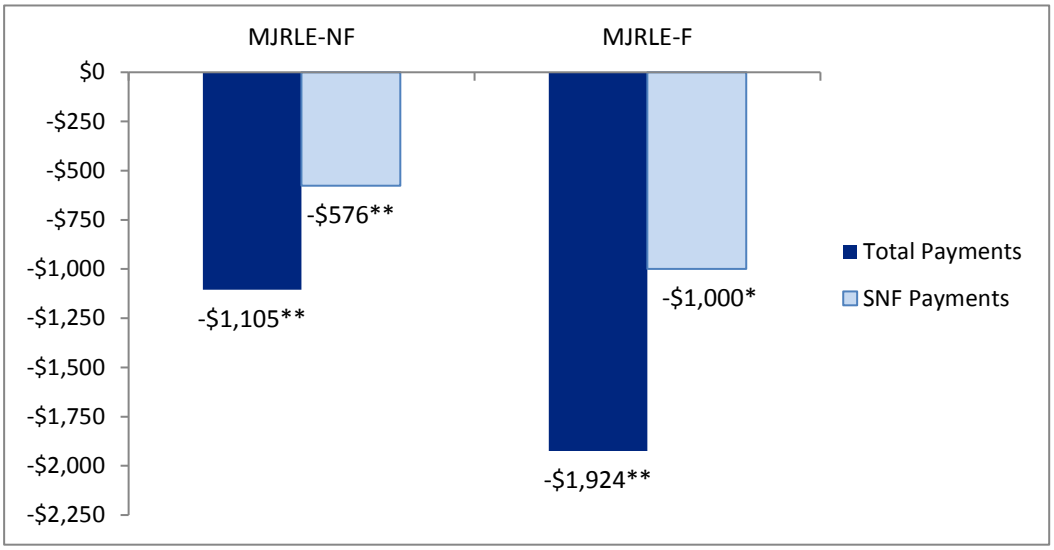
MJRLE-NF= Major Joint Replacement of the Lower Extremity (Non-Fracture); MJRLE-F= Major Joint Replacement of the Lower Extremity (Fracture).

[99] As of Q3 2016, Awardees can sign an amendment electing to use different target prices for fracture and non-fracture episodes.

The greater reductions in total Medicare payments for BPCI MJRLE episodes were driven by reductions in SNF payments for both non-fracture and fracture episodes. Among non-fracture and fracture episodes, reductions in SNF payments explain more than 50 percent of the decline in total payments (see Exhibit 118). For non-fracture episodes, SNF payments decreased $576 more from the baseline to the intervention period for BPCI episodes than comparison episodes ($p<0.01$). As displayed in Exhibit 119, prior to BPCI, SNF payments for non-fracture episodes initiated by BPCI hospitals were $4,276 and they declined to $3,233 during the intervention period. SNF payments for non-fracture episodes decreased less among comparison hospitals, from $4,263 in the baseline to $3,797 in the intervention period. SNF payments for fracture episodes decreased by $1,000 more for BPCI episodes than comparison episodes ($p=0.02$, see Exhibit 119). SNF payments decreased from $16,382 to $15,385 among BPCI fracture episodes, while payments remained relatively stable for comparison episodes ($15,914 in the baseline and $15,917 in the intervention).

Among non-fracture episodes, there was also a significant reduction in IRF payments for BPCI episodes relative to comparison episodes. The reduction in IRF payments was $407 greater for BPCI non-fracture episodes than comparison episodes ($p<0.01$). Payments for IRF services declined by 42% for BPCI non-fracture episodes, from $1,357 in the baseline to $781 in the intervention. IRF payments decreased less for comparison episodes, from $1,130 in the baseline to $961 in the intervention period (15%). Payments for evaluation and management (E&M) services for BPCI non-fracture episodes also declined significantly relative to the comparison group, but payments for E&M services comprise a small share of total episode payments (approximately 3%).

Exhibit 118: Relative Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day PDP and SNF Payments ($2015) during the 90-day PDP, Non-Fracture and Fracture MJRLE, Model 2, Q4 2013 - Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison hospitals.

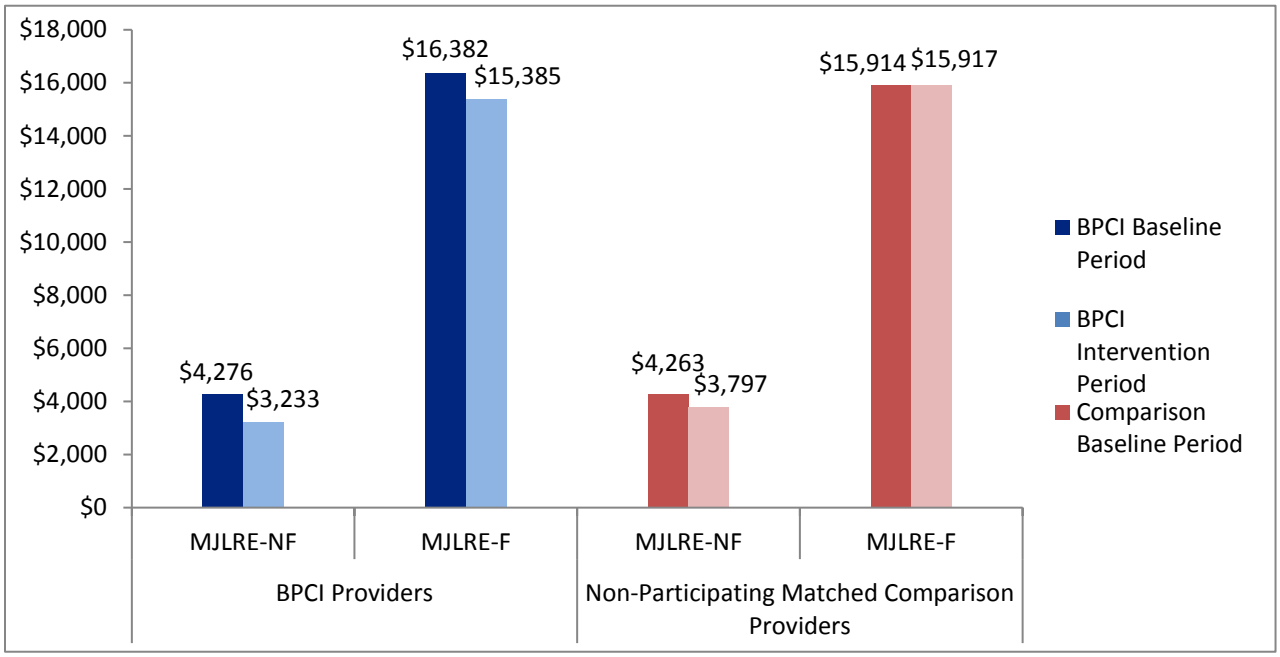
Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

* Indicates $p < 0.05$

** Indicates $p < 0.01$

MJRLE-NF= Major Joint Replacement of the Lower Extremity (Non-Fracture); MJRLE-F= Major Joint Replacement of the Lower Extremity (Fracture).

Exhibit 119: Impact of BPCI on SNF Payments ($2015) during the 90-day PDP, Non-Fracture and Fracture MJRLE, Model 2, Q4 2013 - Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison hospitals.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

* Indicates $p < 0.05$

** Indicates $p < 0.01$

MJRLE-NF= Major Joint Replacement of the Lower Extremity (Non-Fracture); MJRLE-F= Major Joint Replacement of the Lower Extremity (Fracture).

The reductions in SNF payments align with reductions in the number of days spent in a SNF. The relative decline in SNF payments was greater for fracture patients than non-fracture patients due to a greater reduction in the number of SNF days among fracture patients. For non-fracture episodes, the number of SNF days declined from 18.8 to 17.8 for BPCI episodes and remained stable at 18.3 for comparison episodes, resulting in a relative decrease of 1.0 day ($p<0.01$). For fracture episodes, there was a relative decline of 2.4 SNF days ($p<0.01$). The number of SNF days decreased from 40.0 to 36.9 for BPCI episodes, and decreased less, from 39.4 to 38.8 for comparison episodes.

There were also changes in the share of patients who received PAC services. Among non-fracture episodes, the proportion of patients discharged to an institutional PAC facility among those who received any PAC decreased by 7.0 percentage points from the baseline to the intervention period for BPCI episodes relative to the comparison group ($p<0.01$, see Exhibit 121). That is to say, among patients discharged with PAC, there was an increase in the proportion of BPCI non-fracture patients who were discharged to home health. Among fracture episodes, there was a relative decrease in the proportion of patients that were discharged with any PAC (-1.1 percentage points, $p=0.08$).

The decrease in the proportion of institutional PAC users (among non-fracture episodes) and the decrease in the proportion of patients using any PAC (among fracture episodes) align with

information we received through quarterly interviews and site visits with BPCI participants. Interviewees described various strategies they used to reduce the proportion of patients discharged to institutional PAC, including involving family or caregivers in pre-operative education classes to build a post-operative support network that could increase the success of home discharge, providing patient education on PAC use and length of stay to set expectations, beginning discharge planning earlier, and using discharge planning tools to determine post-discharge location, additional care needs, and supports required in the home environment. The Technical Expert Panel indicated that home is the best site of care, suggesting the shift away from institutional PAC could be an improvement in the quality of care.

There was also a decrease of 0.1 days in the average inpatient hospital length of stay for non-fracture BPCI episodes relative to comparison episodes ($p<0.01$, see Exhibit 120). This was driven by a larger decrease in the length of stay for BPCI hospitals (from 4.1 days to 3.6 days) compared to comparison hospitals (from 4.1 to 3.8 days) from the baseline to the intervention period.

Exhibit 120: Number of SNF days for Major Joint Replacement of the Lower Extremity MS-DRG, Non-Fracture and Fracture, Model 2, Q4 2013 - Q3 2015

Inpatient and post-acute care utilization		Number of intervention episodes		BPCI		Comparison		DiD estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Non-fracture	SNF Days	11,171	12,927	18.8	17.8	18.3	18.3	-1.0	-1.7	-0.3	<0.01
Fracture	SNF Days	4,045	4,027	40.0	36.9	39.4	38.8	-2.4	-3.9	-1.0	<0.01

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; IP = inpatient; SNF = skilled nursing facility; HH = home health; PAC = post-acute care. This outcome is dependent on having at least one day in a SNF.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Exhibit 121: Proportion of Patients Discharged to PAC (Institutional or Any) for Major Joint Replacement of the Lower Extremity MS-DRG, Non-Fracture and Fracture, Model 2, Q4 2013 - Q3 2015

Proportion of patients discharged		Number of intervention episodes		BPCI		Comparison		DiD estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Non-fracture	Institutional PAC vs. HH	28,211	29,355	56.6%	44.5%	55.0%	49.9%	-7.0	-10.4	-3.6	<0.01
	PAC vs No PAC	35,465	35,463	84.9%	79.5%	87.3%	83.7%	-1.8	-5.4	1.8	0.33
Fracture	Institutional PAC vs. HH	5,226	5,223	93.8%	93.1%	93.5%	93.6%	-0.7	-2.1	0.6	0.28
	PAC vs No PAC	5,505	5,504	94.8%	94.7%	93.9%	95.0%	-1.1	-2.4	0.1	0.08

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; HH = home health; PAC = post-acute care. The institutional PAC vs. HH outcome is dependent on being discharged to PAC.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

d. Quality of Care

There were no significant changes in claim-based quality outcomes (i.e., mortality, emergency department use, or unplanned readmission rates) between the baseline and intervention periods for BPCI relative to the comparison group for either non-fracture or fracture episodes. We did observe, however, relative declines in functional improvement as measured by the change between two patient assessments among non-fracture patients who were discharged to a SNF or IRF. There were no statistically significant changes in functional status for fracture episodes, with the exception of a relative decline in one of five activity of daily living (ADL) measures among patients discharged to HHA.

For non-fracture episodes discharged to a SNF who remained in the SNF long enough to have two patient assessments, BPCI was associated with a significant relative decrease in the proportion of patients with improvement in each of three ADL measures ($p \leq 0.06$, see Exhibit 122). For all three measures – improved overall function, improved self-care function, and improved mobility function – the proportion of BPCI patients with improvement declined from the baseline to the intervention period, while the proportion of comparison patients with improvement increased. Given the reduction in the use of institutional PAC during the intervention, it is likely that BPCI non-fracture patients discharged to SNFs were more complex than those discharged prior to BPCI. However, these impact estimates take into account patient health status upon admission to the SNF. It should be noted that BPCI non-fracture patients had fewer SNF days than comparison patients under BPCI, that is, less time to recover function in the SNF. The shorter time spent in the SNF may help explain the relative declines in the ADL measures. It is not possible to know if patients with shorter stays achieved the same level of function after leaving the SNF as they would have had they remained in the SNF longer.

There was also a significant relative decline in one of two ADL measures (the self-care score) among non-fracture episodes discharged to an IRF ($p=0.04$). While there was an increase in the average self-care score for BPCI non-fracture patients from the baseline to the intervention period (12.0 to 12.4), there was a larger increase for comparison patients (11.8 to 12.8), resulting in a relative decline (Exhibit 123).

Exhibit 122: Quality Measures in SNF settings for Major Joint Replacement of the Lower Extremity MS-DRG, Non-fracture and Fracture, Model 2, Q4 2013 - Q3 2015

	SNF Measure	Number of intervention episodes		BPCI		Comparison		DiD estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Non-fracture	Improved overall function	8,620	9,675	74.6%	73.7%	74.3%	76.3%	-2.9	-5.9	0.1	0.06
	Improved self-care function	8,623	9,682	59.7%	55.6%	59.0%	60.1%	-5.2	-8.5	-1.9	<0.01
	Improved mobility function	8,615	9,677	69.2%	67.1%	69.7%	71.2%	-3.6	-6.8	-0.5	0.02
Fracture	Improved overall function	3,087	3,020	66.4%	65.5%	65.8%	65.2%	-0.4	-3.6	2.9	0.82
	Improved self-care function	3,088	3,023	42.8%	41.0%	44.4%	42.8%	-0.2	-3.5	3.2	0.92
	Improved mobility function	3,085	3,016	61.5%	61.3%	62.7%	63.2%	-0.7	-4.1	2.7	0.70

DiD = difference-in-differences; 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 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers

Note: "Improvement" means improvement or staying in the best possible status.

Exhibit 123: Quality Measures in IRF settings for Major Joint Replacement of the Lower Extremity MS-DRG, Non-fracture and Fracture, Model 2, Q4 2013 - Q3 2015

	IRF Measure	Number of intervention episodes		BPCI		Comparison		DiD estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Non-fracture	Average change in mobility score	1,727	1,697	10.7	11.1	10.5	11.2	-0.3	-0.8	0.1	0.13
	Average change in self-care score	1,727	1,697	12.0	12.4	11.8	12.8	-0.7	-1.3	-0.0	0.04
Fracture	Average change in mobility score	1,031	1,051	9.0	9.9	9.0	9.6	0.3	-0.2	0.9	0.27
	Average change in self-care score	1,031	1,051	11.0	11.9	11.1	11.7	0.3	-0.4	1.1	0.39

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Source: Lewin analysis of Medicare claims, IRF-PAI assessment data, and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Note: Changes in mobility and self-care for IRF patients are measured as changes in a mobility related index score and a self-care related index score between IRF admission and IRF discharge. The mobility index ranges from a score of 4 (total assistance) to 28 (fully independent) while the self-care index ranges from a score of 6 (total assistance) to 42 (fully independent).

e. Functional Status and Care Experiences

Beneficiary survey respondents with MJRLE episodes initiated in BPCI-participating hospitals consistently reported similar changes or greater improvements in functional status from before to after their episode of care and generally similar care experiences as matched comparison respondents.[100]

Functional status measures regarding patient mobility, including walking without rest and using stairs, are particularly relevant for patients who undergo a MJRLE. In all four survey waves, BPCI respondents were more likely to report improvement in walking without rest from before to after their surgical episode than comparison respondents, although the difference was only statistically significant in Wave 2 (8.2 percentage points; p=0.02, see Exhibit 125). Additionally, in all four survey waves BPCI respondents were more likely to report improvement in using stairs from before to after their surgical episode than comparison respondents, although the improvement was significant in only two waves (Wave 2: 7.5 percentage points, p=0.03; Wave 4: 7.9 percentage points, p=0.02, see Exhibits 125a & 125b).

In addition to improved mobility, BPCI respondents were more likely to report improvements in the extent to which pain limited their regular activities and physical and emotional problems limited their social activities than comparison respondents. The higher rate of improvement in pain among BPCI respondents was present in all four waves, although the difference was statistically significant in Wave 4 only (6.6 percentage points, p=0.03). Similarly, the higher rate of improvement in physical and emotional problems among BPCI respondents was present in three of the four waves, and was significant in Waves 4 (7.1 percentage points, p=0.03) and 5 (6.5 percentage points, p=0.03).

Across the four waves, there was only one functional status survey measure in which BPCI respondents were more likely to report a less favorable change than comparison respondents.[101] In wave 2, BPCI respondents were more likely to report a decline in planning regular tasks (4.2 percentage points; p=0.06, see Exhibit 124).

BPCI respondents generally reported having similar care experiences as matched comparison respondents. Across all four waves, BPCI and comparison respondents reported similar levels of overall satisfaction with recovery since leaving the hospital. BPCI respondents reported significantly better experiences than comparison respondents on two out of ten measures in at least one wave.

---

[100] Unlike the analysis of the assessment-based quality measures, which quantifies the impact of BPCI by comparing changes in functional status from the baseline to the intervention period for BPCI episodes relative to comparison episodes, the analysis of the survey data compares changes in functional status from before to after an episode of care for BPCI episodes relative to comparison episodes. Furthermore, the beneficiary survey includes a sample of BPCI episodes not limited to patients who received PAC or were in a particular PAC setting.

[101] Subsequent results from the Wave 6 (Spring 2016) and Wave 7 (Summer 2017) surveys indicated a reversal of the positive changes in functional status among Model 2 BPCI respondents with MJRLE episodes. In Wave 6, BPCI respondents did not report any significantly higher rates of improvement or lower rates of decline on any measure relative to comparison respondents. Rather they indicated significantly lower rates of improvement or higher rates of decline for 4 of 7 functional status measures, including regular activities limited by pain (p<0.01), and use of a mobility device (p=0.03) planning regular tasks, use of a mobility device, using stairs, and the degree to which pain limited their regular activities. In Wave 7, BPCI respondents did report significantly higher rates of improvement in pain limiting everyday activities (p=0.03), but still reported lower rates of improvement or higher rates of decline for 3 functional status measures, including planning regular tasks, walking without rest (p=0.07) and using stairs (p=0.03).

BPCI respondents were more likely to report that medical staff always spoke to them in their preferred language (Wave 4: 1.9 percentage points, p=0.09; Wave 5: 2.4 percentage points, p=0.09) and they never received conflicting advice (Wave 4: 6.5 percentage points, p=0.04). There were also two measures in which BPCI respondents reported significantly worse care experiences than comparison respondents. In wave 2, BPCI respondents were less likely to report that medical staff clearly explained follow-up care (-2.8 percentage points, p<0.01) and they felt they were discharged at the right time (-3.4 percentage points, p=0.10).

Exhibit 124: Select Outcomes from the Beneficiary Survey, MJRLE Episodes Model 2, October/November 2014, February/March 2015, May/June 2015, and October/November 2015 (Waves 2-5)

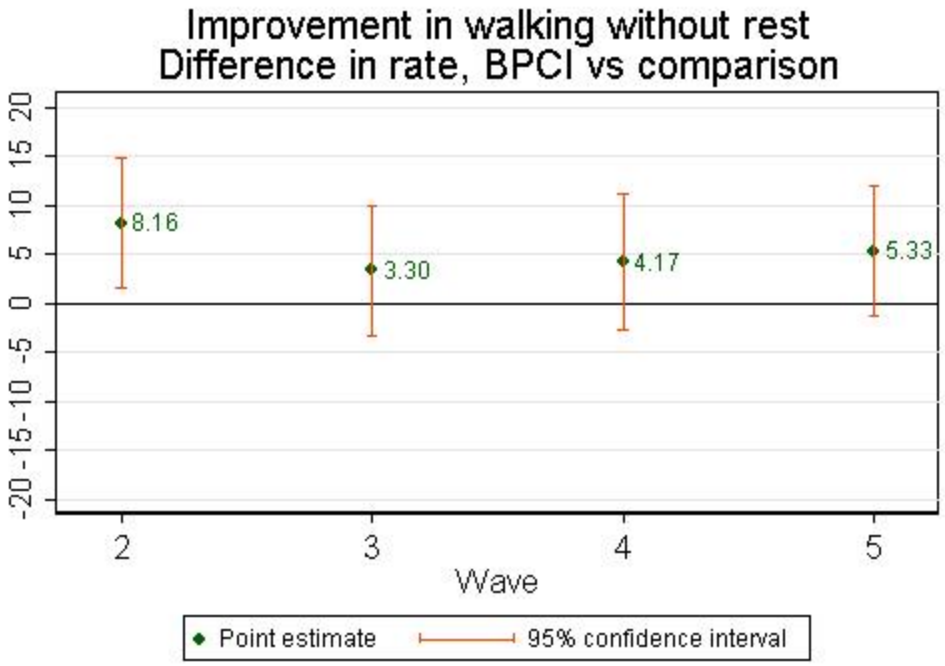
	Measure	Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
			BPCI	Comparison	BPCI	Comparison	Point Estimate	95% LCI	95% UCI	p-value
Functional Status	Improvement in bathing, dressing, using the toilet, or eating	2	353	373	85.0	84.5	0.5	-4.3	5.2	0.84
		3	353	346	85.5	82.9	2.7	-2.3	7.6	0.30
		4	360	346	84.3	85.6	-1.3	-6.3	3.7	0.60
		5	360	377	92.4	84.0	8.4	3.9	12.9	<0.01
	Improvement in walking without rest	2	351	374	65.7	57.5	8.2	1.6	14.8	0.02
		3	357	351	70.1	66.8	3.3	-3.4	10.0	0.34
		4	358	344	61.7	57.5	4.2	-2.7	11.1	0.24
		5	363	379	67.9	62.6	5.3	-1.3	12.0	0.12
	Improvement in using stairs	2	353	366	65.4	57.9	7.5	0.9	14.1	0.03
		3	354	339	64.1	60.0	4.0	-2.6	10.7	0.23
		4	358	344	65.2	57.3	7.9	1.2	14.7	0.02
		5	358	365	63.6	62.6	1.1	-5.6	7.8	0.75
	Improvement in physical/emotional problems limiting social activities	2	348	372	75.8	72.0	3.8	-2.4	10.0	0.23
		3	359	351	74.7	74.7	-0.0	-6.3	6.3	0.99
		4	361	345	76.5	69.4	7.1	0.7	13.5	0.03
		5	360	377	79.9	73.3	6.5	0.6	12.5	0.03
	Improvement in pain limiting regular activities	2	355	378	81.9	77.9	4.0	-1.6	9.7	0.16
		3	360	353	78.6	74.9	3.7	-2.1	9.5	0.21
		4	365	347	80.6	74.0	6.6	0.6	12.5	0.03
		5	364	381	81.5	78.8	2.8	-2.8	8.4	0.33
	Decline in bathing, dressing, using the toilet, or eating	2	353	373	8.2	8.4	-0.1	-3.9	3.7	0.95
		3	353	346	6.9	9.3	-2.4	-6.4	1.6	0.23
		4	360	346	7.8	8.8	-1.0	-5.3	3.3	0.65
		5	360	377	3.7	9.5	-5.8	-9.5	-2.2	<0.01
	Decline in planning regular tasks	2	355	377	13.7	9.5	4.2	-0.1	8.5	0.06
		3	359	351	9.6	11.6	-2.0	-6.5	2.5	0.38
		4	362	349	11.0	12.0	-1.0	-5.7	3.7	0.68
		5	364	377	7.3	11.7	-4.4	-8.4	-0.5	0.03

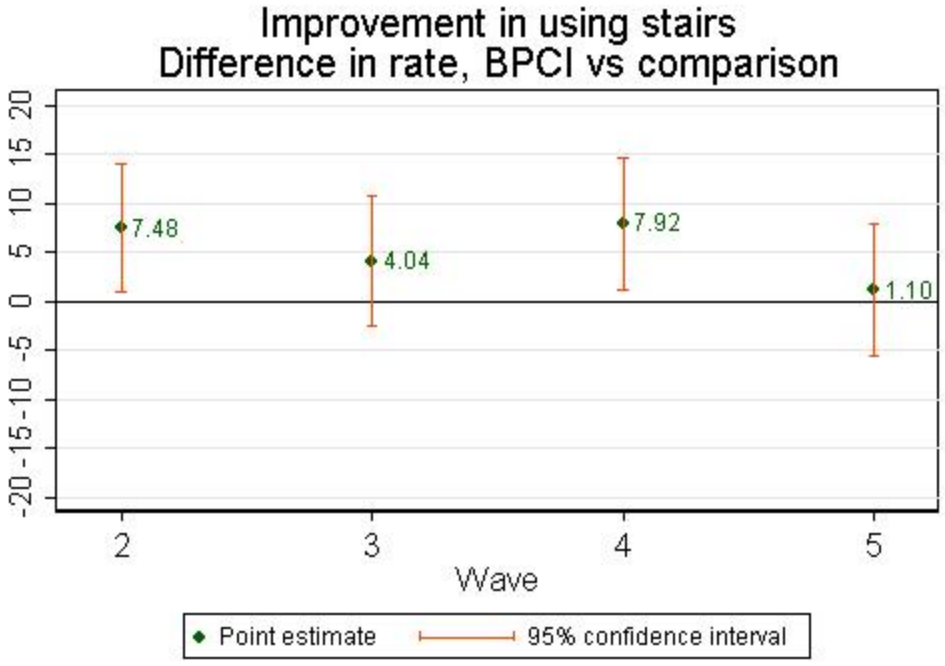
	Measure	Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
			BPCI	Comparison	BPCI	Comparison	Point Estimate	95% LCI	95% UCI	p-value
Functional Status (cont'd)	Decline in walking without rest	2	351	374	14.4	15.7	-1.3	-6.2	3.5	0.59
		3	357	351	14.5	16.0	-1.5	-7.1	4.1	0.60
		4	358	344	18.2	19.4	-1.2	-6.8	4.5	0.69
		5	363	379	11.4	14.3	-2.9	-7.7	1.9	0.23
	Decline in pain limiting regular activities	2	355	378	7.0	8.5	-1.5	-5.4	2.5	0.47
		3	360	353	7.5	10.9	-3.4	-7.3	0.5	0.08
		4	365	347	7.8	14.1	-6.3	-11.2	-1.3	0.01
		5	364	381	9.8	7.7	2.2	-1.8	6.1	0.29
Patient Care Experience	Never received conflicting advice from the medical team	2	357	379	81.2	83.1	-1.9	-7.5	3.7	0.51
		3	358	352	82.5	81.3	1.3	-4.7	7.2	0.67
		4	357	345	85.1	78.6	6.5	0.4	12.6	0.04
		5	367	377	83.8	82.4	1.5	-3.9	6.8	0.60
	Staff always used patient's preferred language	2	355	379	94.3	96.6	-2.3	-5.4	0.7	0.14
		3	357	351	96.5	96.8	-0.4	-2.5	1.7	0.72
		4	359	346	97.0	95.1	1.9	-0.3	4.1	0.09
		5	371	382	98.0	95.7	2.4	-0.4	5.1	0.09
	Patient was discharged at right time	2	358	377	89.9	93.3	-3.4	-7.5	0.7	0.10
		3	357	353	91.9	93.6	-1.7	-5.7	2.3	0.40
		4	358	343	94.0	95.2	-1.2	-4.7	2.3	0.50
		5	367	383	93.4	94.8	-1.5	-5.2	2.2	0.43
	Medical staff clearly explained follow-up care	2	332	341	96.9	99.7	-2.8	-4.7	-0.8	<0.01
		3	331	326	95.1	96.0	-0.9	-4.2	2.5	0.61
		4	327	317	97.7	97.5	0.2	-2.1	2.6	0.86
		5	344	358	95.4	97.6	-2.2	-5.1	0.6	0.12

Note: This table reports all functional status and care experience measures that have a statistically significant difference between the intervention and comparison groups at the 10% or 5% level in at least one wave. The improvement indicator takes a value of 1 if a patient moved to a better functional status level after the episode or if the patient recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response. The decline indicator takes a value of 1 if the patient moved to a worse functional status group after the episode or if the patient recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey. The indicators are assigned a value of 0 otherwise.

Source: Lewin analysis of beneficiary survey data for episodes that began October to November 2014 (Wave 2), February to March 2015 (Wave 3), May to June 2015 (Wave 4), and October to November 2015 (Wave 5) for BPCI EIs and the matched comparison providers.

Exhibits 125a & 125b: Improvement in Mobility Measures, Model 2 MJRLE Episode, Matched BPCI and Comparison Respondents





Source: Lewin analysis of beneficiary survey data for episodes that began October to November 2014 (Wave 2), February to March 2015 (Wave 3), May to June 2015 (Wave 4), and October to November 2015 (Wave 5) for BPCI EIs and the matched comparison providers.

f. Variation in NPRA among BPCI Providers

There was variation in financial performance among BPCI hospitals for MJRLE episodes, as measured by average NPRA per episode, as shown in Exhibit 126.[102] Top performers were defined as hospitals with standardized NPRA for MJRLE episodes above the 75th percentile of the standardized NPRA distribution, average performers were defined as hospitals with standardized NPRA for MJRLE episodes between the 25th and 75th percentiles of the standardized NPRA distribution, and bottom performers were those with standardized NPRA below the 25th percentile of the distribution.[103] Top performers had a mean average NPRA of $2,490 per episode, whereas bottom performers had a mean average NPRA of -$1,647.[104]

Exhibit 126: Average NPRA/Episode, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2013 - Q2 2015

Performance Group	N	Mean	Median	Min	Max
Top Performers (Above 75th Percentile)	24	$2,490	$1,939	$1,423	$7,117
Average Performers (Between 25th-75th Percentile)	48	$471	$329	-$377	$2,113
Bottom Performers (At or below the 25th Percentile)	25	-$1,647	-$1,222	-$6,053	-$412

Source: Lewin analysis of Net Payment Reconciliation Amount data for the Q4 2013- Q2 2015 period.

There were differences in the change in total payments from the baseline to the intervention period for a MJRLE episode among the three performance groups.[105] Although on average, the three performance groups decreased total payments from the baseline to the intervention period, top performers experienced the biggest decrease. The average total payments for a MJRLE episode among top performers decreased by $4,298 from baseline to intervention, while the average total payments for a MJRLE episode among bottom performers decreased by $1,667 (Exhibit 127).

Top performers reduced the proportion of patients discharged to institutional PAC and the number of SNF days to a greater extent than other hospitals participating in MJRLE. Both represent opportunities for savings given that institutional PAC is an expensive site for care. At the same time, top performers increased the proportion of patients discharged to home health, which is a less expensive form of PAC, to a greater extent than other hospitals. Changes in the use of institutional PAC and home health, as well as changes in the characteristics of the

[102] Unlike the DiD analysis that uses total standardized payments per episode to measure Medicare costs, this analysis focuses on providers' average NPRA per episode, which measures the difference between the "target price" for services provided during the episode of care and the total dollar amount of Medicare fee-for serves expenditures for that episode. When providers' episode payments are below the target price (i.e., positive NPRA) they receive this amount from CMS. However, when the providers' episode payments exceed the target price (i.e., negative NPRA) they may need to return amounts to CMS. For more detail on the NPRA performance measure, please see the methods section.

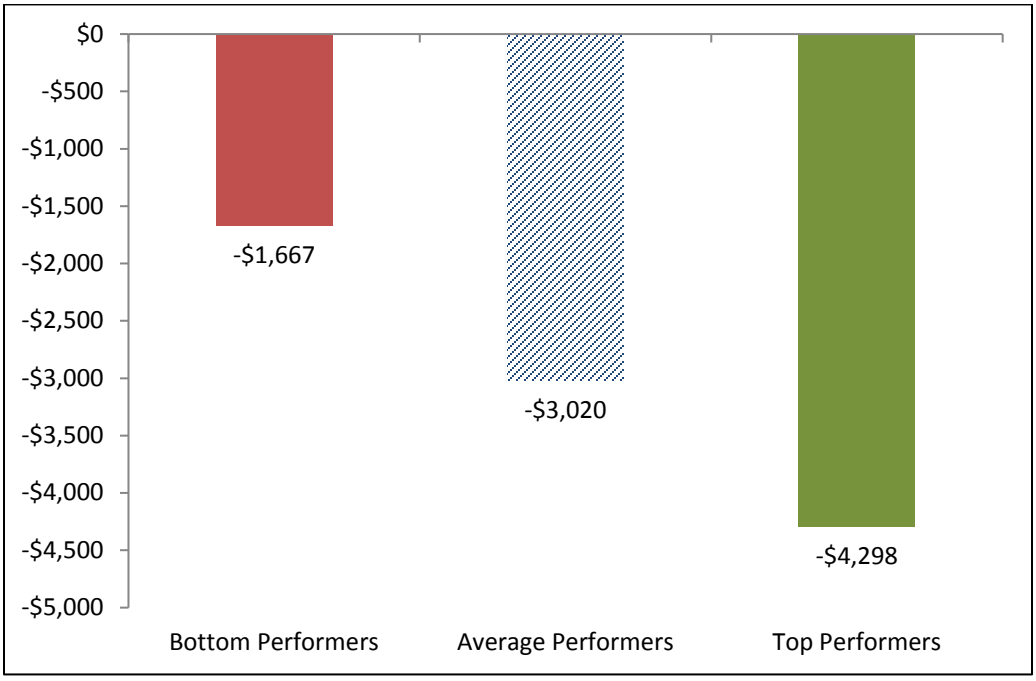
[103] To adjust for differences in baseline payments and thus target prices among BPCI providers, we standardize NPRA by dividing it by the provider's average target price.

[104] BPCI providers were required to meet certain criteria to be included in this analysis, such as having at least 50 episodes in the baseline and intervention periods. For more detail on the inclusion and exclusion criteria, please see the Methods section.

[105] The analyses presented in this section encompass the first seven quarters of the initiative (from October 2013 through June 2015) and a baseline period spanning from October 2011 through September 2012.

beneficiaries treated may account for the bigger decreases in average total payments for a MJRLE episode observed among top performers.

Exhibit 127: Mean Change in Total Standardized Allowed Payment ($2015), Inpatient Hospitalization plus 90-day Post-discharge Period, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

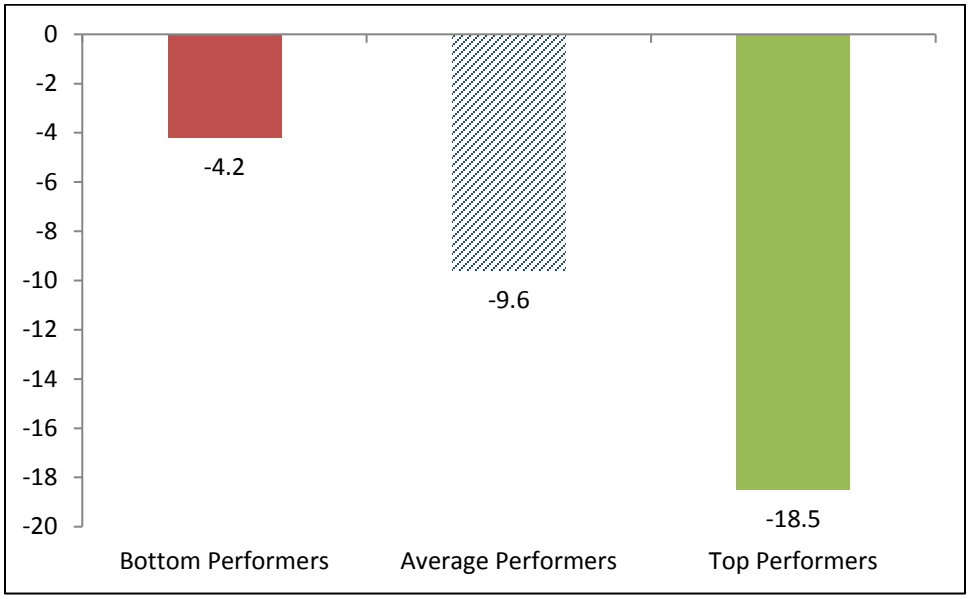
All three performance groups decreased the percent of beneficiaries discharged to institutional PAC from the baseline to the intervention period, but top performers experienced the largest decrease. Top performers reduced the percent of discharges to institutional PAC by 18.5 percentage points whereas bottom performers decreased the percent of discharges to institutional PAC by 4.2 percentage points (see Exhibit 128). This result runs parallel to anecdotal evidence gathered from site visits and quarterly interviews in which top performing hospital representatives discussed their approaches for reducing the use of SNF and substituting HHA care.

The average number of SNF days also decreased more among top performers than bottom performers (-1.96 days vs. -0.09 days, see Exhibit 129). This result is also consistent with anecdotal evidence from site visits and quarterly interviews, in which several top performing interviewees noted decreases in SNF length of stay (LOS). One interviewee attributed the decrease in SNF LOS to their patients going to partnering SNFs, and another attributed it to meeting continuously with SNF providers (long before the BPCI initiative started) and educating them about payment reform and the need to reduce SNF transfers and LOS.

Consistent with the declines in the proportion of patients discharged to institutional PAC and the number of days spent in a SNF, the three performance groups experienced increases in the percent of beneficiaries discharged to HHAs from the baseline to the intervention period. As shown in

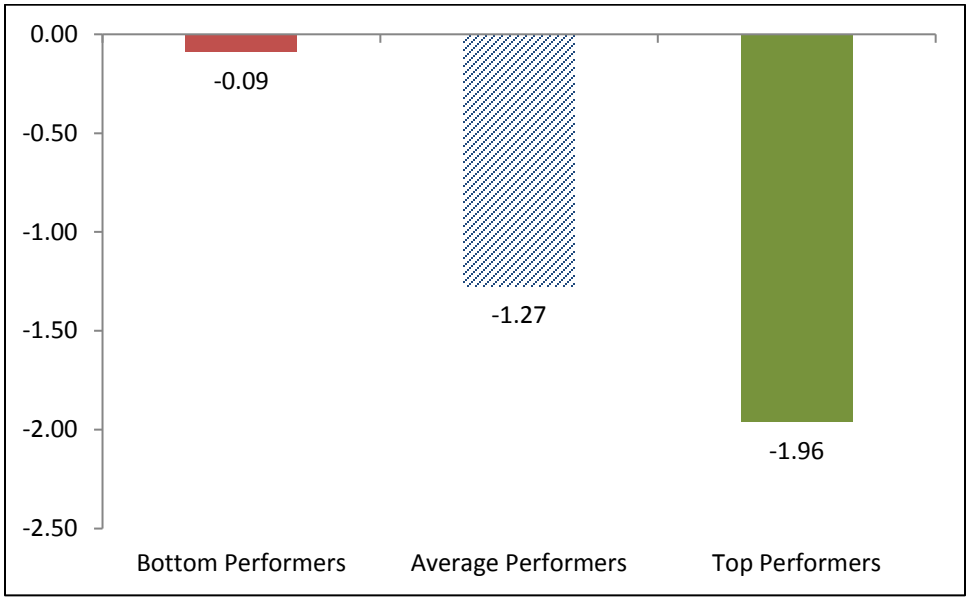
Exhibit 130, top performers experienced the largest increase (9.5 percentage points for top performers vs. 1.7 percentage points for bottom performers).

Exhibit 128: Mean Change in Percent Discharged to Institutional PAC, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



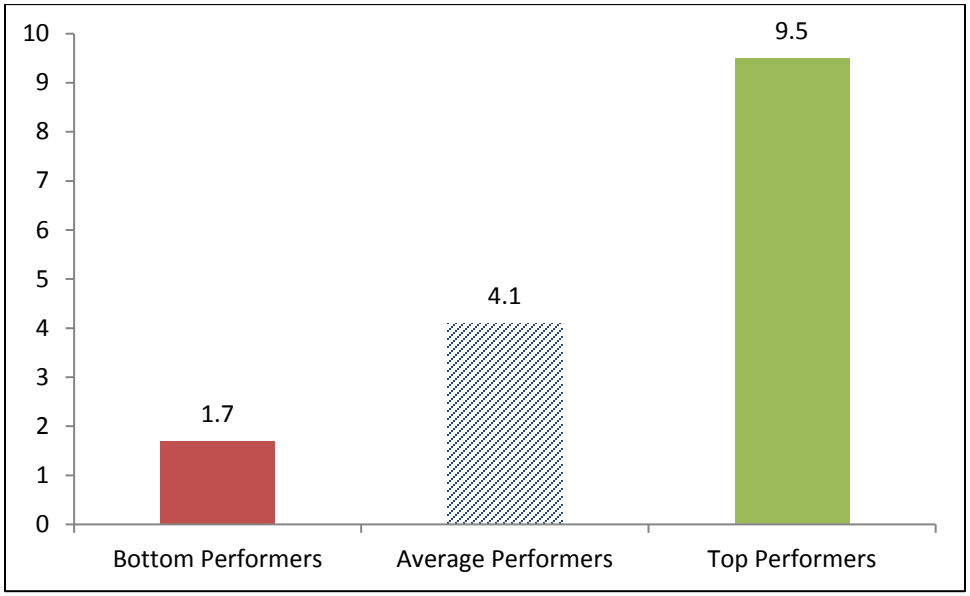
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 129: Mean Change in Number of SNF Days, 90-day PDP, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 130: Mean Change in Percent Discharged to Home Health, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



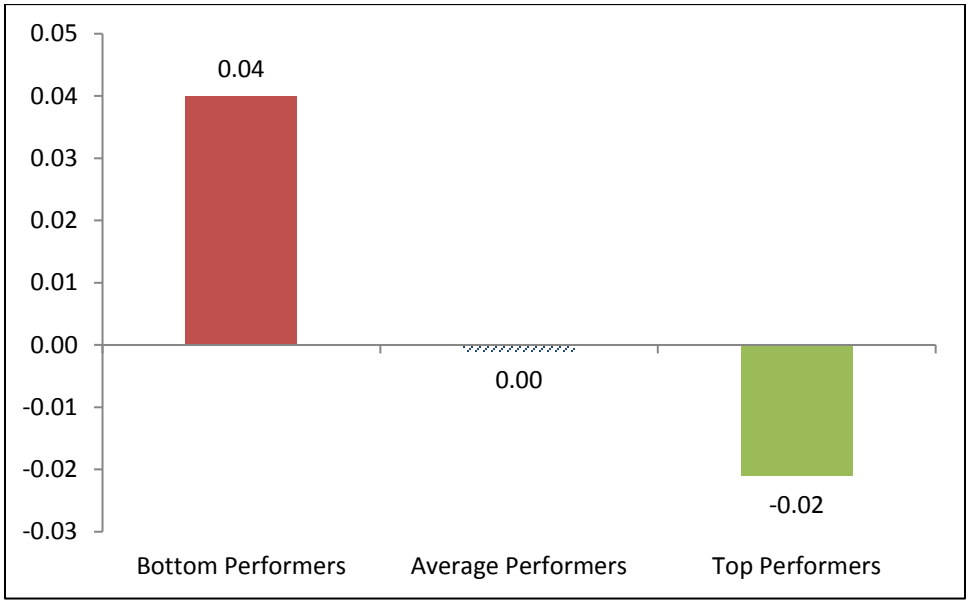
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Finally, the characteristics of the beneficiaries treated at both bottom and top performing hospitals changed from the baseline to the intervention period. A Hierarchical Condition Category (HCC) score measures the expected cost of a beneficiary's health care requirements based on diagnoses observed during a baseline period. Higher scores suggest greater future spending. Medicare uses HCC scores to adjust payments to health plans that enroll beneficiaries under the Medicare Advantage program. Top performers experienced a decrease in the average HCC overall score of their patients, whereas bottom performers experienced an increase. The average HCC overall score for top performers' patients decreased by 0.02 points whereas the average score for patients treated by bottom performers increased by 0.04 points (see Exhibit 131).

Consistent with the change in beneficiaries' average HCC score observed among top and bottom performers, the average age of beneficiaries changed as well. The average age of beneficiaries treated by top performers decreased by 0.64 years while the average age of beneficiaries treated by bottom performers increased by 0.13 years. This may be because top performers are avoiding high risk patients or they are increasing their non-fracture patient volume.

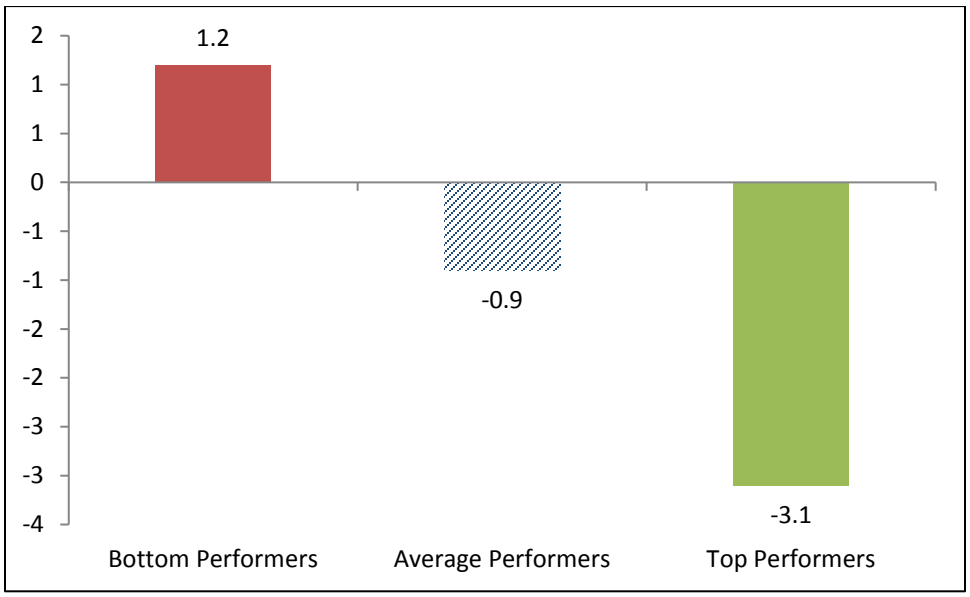
Top performers experienced decreases in the proportion of MJRLE episodes with fracture, while bottom performers experienced increases in the proportion of MJRLE episodes with fracture. The proportion of MJRLE episodes with fracture among top performers decreased by 3.1 percentage points while the proportion of episodes with fracture among bottom performers increased by 1.2 percentage points, as shown in Exhibit 132. These results are consistent with anecdotal evidence gathered from site visits and quarterly interviews. An interviewee from one top performing hospital indicated that fracture and more complex patients were operated on at a different facility, while an interviewee from another top performing hospital said that their practice selected healthier patients and used a risk stratification tool to decide whether to perform surgery on patients.

Exhibit 131: Mean Change in Average HCC Overall Score, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

Exhibit 132: Mean Change in Percent of MJRLE Episodes with Fracture, by Performance Group, Hospital-initiated MJRLE Episodes, Model 2, Q4 2011 - Q2 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q2 2015 for BPCI providers.

The above analysis suggests that there was variation in the ability of hospital EIs to reduce the payments for MJRLE episodes relative to their target prices. The most notable successful strategy was to reduce the use of institutional PAC, often by substituting less expensive HHA care. The finding that the hospitals that were more successful in reducing episode payments also changed their mix of patients to a healthier, younger, and less complex population raises the question of

whether it was actually increased efficiencies in care delivery or changes in patient needs that led to the lower episode payments.

g. Conclusions

Under BPCI, hospitals treating patients in a MJRLE episode significantly reduced total payments for both non-fracture and fracture MJRLE episodes relative to a comparison group. The reduction in total payments was driven by reductions in PAC payments, particularly SNF payments. There were corresponding changes in utilization, including a decrease in the number of days spent in a SNF for non-fracture and fracture patients. There was also a decrease in the proportion of institutional PAC users among non-fracture episodes and a decrease in the proportion of patients using any PAC among fracture episodes. Top financial performers were more likely to substitute HHA care for institutional PAC than other hospitals participating in MJRLE, supporting the hypothesis that reducing institutional PAC utilization is a major driver of reductions in total payments.

There were no significant changes in mortality, emergency department use, or unplanned readmission rates between the baseline and intervention periods for BPCI patients relative to the comparison group for either fracture or non-fracture episodes. Among the non-fracture patients discharged to a SNF who remained long enough to receive a second patient assessment, BPCI was associated with a significant relative decrease in the proportion of patients improving on each of three ADL measures. BPCI non-fracture patients spent fewer days in a SNF than they did prior to BPCI relative to comparison patients though, which may help explain the relative declines in the ADL measures. Survey results indicate that BPCI MJRLE patients consistently reported similar changes or greater improvements in functional status from before to after their surgical episode, relative to matched comparison respondents, across four survey waves.

There is some evidence that patient selection may be occurring for non-fracture episodes, based on changes favorable to BPCI participants on four measures of care utilization in the six months prior to the episode. However, this does not explain the statistically significant BPCI impact estimates on payment, utilization, and quality because the risk adjustment models control for these changes in patient mix.

9. Percutaneous Coronary Intervention

Medicare beneficiaries received 1,889 percutaneous coronary intervention (PCI) episodes of care in 37 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015). BPCI did not result in a statistically significant change in total Medicare payments during the initial hospital stay at participating providers and on services during the subsequent 90-day post-discharge period.[106] Relatively few PCI episodes include post-acute care (PAC), which limits BPCI hospitals' ability to reduce total episode payments by decreasing the intensity of PAC use. In fact, providers that we interviewed as part of this evaluation indicated that their care redesign for PCI focuses on reducing internal costs

[106] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

during the hospital admission.[107] Such strategies may be successful in increasing hospital efficiency and potentially the profitability of PCI admissions, but they would not reduce Medicare payments to help the hospital stay under the episode benchmark. Importantly, quality of care was maintained.

a. Participants

During the first eight quarters of the BPCI initiative, 37 hospitals (9% of all Model 2 BPCI hospitals) participated in the PCI episode, all of which chose 90-day episodes. There were 1,889 PCI episodes started by hospitals during the BPCI initiative (approximately 1% of episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the PCI results are based on an average of two quarters of experience. Approximately 41% of the hospitals began participating in PCI in Q3 2015; 76% joined in either Q2 or Q3 2015. None of the providers we interviewed indicated that they joined BPCI specifically to participate in the PCI episode. One hospital representative said that the hospital selected the PCI episode because of its Heart Center, which was an area of strength for quality and performance improvement. Four hospitals (11%) stopped participating in PCI within the first eight quarters of the initiative. During an interview, a BPCI Convener shared that one reason one of their hospital episode initiators dropped the PCI episode was because the bundle did not account for the number of stents used in the procedure.

Compared with hospitals that did *not* participate in BPCI, hospitals that participated in the PCI episode were larger and more likely to be for-profit and in urban locations (Exhibits 133a & 133b). Hospitals with the PCI episode were also less likely to be part of a chain. Further, prior to joining BPCI, participating hospitals had more PCI discharges and had higher average 2011 standardized Part A payments for PCI discharges during the anchor stay plus the 90-day post-discharge period (PDP) than non-participants. Matched comparison hospitals were identified for 32 of the 37 BPCI-participating hospitals in the sample.[108] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals.

[107] See Methods section for more information on the BPCI participant site visits and interviews.

[108] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. Three hospitals were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the PCI episode or had fewer than five relevant discharges in 2011 or 2012. Two additional participants were not included because there was not a match within the caliper.

Exhibits 133a & 133b: Characteristics of BPCI-participating Hospitals with Percutaneous Coronary Intervention Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals Model 2, Q4 2013 - Q3 2015

Characteristic		All PCI BPCI Hospitals (N=37)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched PCI BPCI Hospitals (N=32)		Comparison Hospitals (N=409+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	17	46%	1,594	57%*	321	77%*	14	44%	196	48%
	Government	1	3%	542	20%*	32	8%*	1	3%	11	3%
	For-Profit	19	51%	638	23%*	66	16%*	17	53%	202	49%
Urban/Rural	Urban	37	100%	1,902	69%*	387	92%*	32	100%	409	100%
Part of Chain	Yes	13	35%	1,469	53%*	218	52%*	13	41%	147	36%

Characteristic	All PCI BPCI Hospitals (N=37)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched PCI BPCI Hospitals (N=32)	Comparison Hospitals (N=409+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	323	175*	311	332	302
Medicare Days Percent	38%	42%*	39%	38%	38%
Resident-to-bed ratio	0.08	0.05*	0.12*	0.09	0.08
Disproportionate Share Percent	29%	29%	27%	29%	29%
Number of PCI Discharges, 2011	118	45*	141	120	115
Standardized Part A Allowed Payment inpatient stay plus 90-day post discharge period, PCI, 2011	**$19,283**	**$17,827***	**$17,629***	**$19,399**	**$19,184**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI during the first year.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All PCI BPCI" hospitals. "Comparison" hospitals are compared to "Matched PCI BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 185.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. A comparison of the PCI patient population of BPCI participants from baseline to intervention relative to the same time periods for the PCI patient population of the comparison providers does not suggest that BPCI providers treated a healthier patient population. (Exhibits 134a & 134b).

Exhibits 134a & 134b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Percutaneous Coronary Intervention, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	395	13%	220	14%	325	10%	199	12%	-0.9	-3.7	1.8	0.50
	65-79	1,840	59%	994	61%	1,819	58%	957	59%	1.6	-2.6	5.8	0.45
	80+	899	29%	415	25%	990	32%	473	29%	-0.7	-4.5	3.2	0.74
Gender	Female	1,298	41%	634	39%	1,333	43%	623	38%	1.8	-2.4	6.0	0.40
Medicaid and Disability	% Eligible for Medicaid	684	22%	315	19%	604	19%	295	18%	-1.3	-4.7	2.1	0.44
	% Disability, no ESRD	469	15%	252	15%	401	13%	227	14%	-0.6	-3.6	2.3	0.67
Utilization Six Months Prior to Anchor Stay	Inpatient Acute Care Hospital	954	30%	461	28%	778	25%	371	23%	-0.1	-3.8	3.7	0.96
	Emergency Room Admission	841	27%	507	31%	845	27%	476	29%	2.0	-1.8	5.8	0.30
	Home Health	456	15%	187	11%	378	12%	206	13%	-3.7	-6.5	-0.8	0.01
	Inpatient Rehabilitation Facility	43	1%	13	1%	25	1%	23	1%	-1.2	-2.1	-0.3	0.01
	Skilled Nursing Facility	133	4%	61	4%	108	3%	64	4%	-1.0	-2.6	0.6	0.24
	Psychiatric Hospital	9	0%	5	0%	11	0%	7	0%	-0.1	-0.5	0.4	0.81
	Long-term Care Hospital	11	0%	2	0%	4	0%	5	0%	-0.4	-0.8	0.0	0.05
	Institutional Nursing Facility**	157	5%	69	4%	113	4%	67	4%	-1.3	-3.0	0.4	0.14
	No Institutional Care	2,156	69%	1,154	71%	2,318	74%	1,244	76%	-0.4	-4.1	3.4	0.85
	No Post-acute Care	1,591	51%	811	50%	1,693	54%	885	54%	-1.3	-5.5	2.9	0.55

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.49	1.98	2.43	1.95	2.34	1.86	2.36	2.04	-0.08	-0.24	0.08	0.34

DiD=difference-in-differences.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

c. Medicare Payments and Utilization

During the first eight quarters of the BPCI initiative, there was no statistically significant change in total Medicare payments for the inpatient hospitalization and all care delivered during the 90-day PDP for BPCI relative to the comparison group (see Exhibit 135). However, this might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome, which is required for an unbiased estimate.

Exhibit 135: Impact of BPCI on Total Standardized Allowed Payment ($2015) Inpatient Hospitalization and 90-day Post-discharge Period for Percutaneous Coronary Intervention Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=1,577)		Comparison (N=1,592)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total Payment, Inpatient stay plus 90-day post-discharge period	$26,549	$26,315	$25,471	$25,894	-$658*	-$2,018	$703	0.34

* This might be a biased estimate because We rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate. We conclude BPCI did not have an adverse effect on total payments. The BPCI baseline change declined at a faster rate than the comparison group, leading to a negative bias on the DiD estimate. However, even with this bias, the DiD is not statistically significant.

DiD = Difference-in-Differences estimate; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

These estimates are developed using a Difference-in-Differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison hospitals.

Relative to other clinical episodes, a smaller proportion of PCI patients are discharged to PAC. Interviewees discussed this low PAC utilization. They indicated that PCI patients are typically discharged home unless they were in a PAC setting prior to the procedure, so they did not focus BPCI care redesign efforts on reducing PAC use. For this reason, it is not surprising that there were no significant differences between BPCI and comparison episodes in payments for PAC, the proportion of patients discharged to PAC, or the percent of patients discharged to institutional PAC (skilled nursing facilities, inpatient rehabilitation facilities, and long-term care hospitals) among PAC users (i.e., versus home health care). There was, however, a significant reduction in home health visits among patients who had any home health use by 4.1 days ($p<0.01$) (see Exhibit 136).[109] Home health visits declined from 18.5 days to 15.6 for BPCI PCI patients, while they increased from 16.9 days to 18.0 for the comparison group.

[109] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

Exhibit 136: Impact of BPCI on Home Health Visits among Home Health Users for Percutaneous Coronary Intervention Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=310)		Comparison (N=304)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Home Health visits, among home health users	18.5	15.6	16.9	18.0	-4.1	-6.1	-2.0	<0.01

DiD = Difference-in-Differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Another indicator of less intensive HHA use, among the small fraction of patients discharged to HHA, is the relative decline in HHA case-mix indices. The BPCI patients discharged to HHA had a greater reduction in HHA case-mix indices between the baseline and intervention period than the comparison PCI patients discharged to HHA (significant at the 10% level). The HHA case-mix index decreased from 1.33 to 0.95 for BPCI compared to a decrease from 1.30 to 1.04 for comparison group patients.[110]

In addition to a significant reduction in home health visits, the Part B payments for imaging/lab and evaluation and management visits decreased more in BPCI episodes than in matched comparison episodes (See Exhibit 137). Imaging/lab payments declined by $81 for BPCI relative to comparison episodes (p=0.01), and payments for evaluation and management visits declined by $173 (p=0.02). However, because these two services accounted for approximately 7% of total episode spending, these relatively small reductions did not appreciably affect total payments for the episode.

Exhibit 137: Impact of BPCI on Part B Payments (2015$) for Imaging/Lab and Evaluation and Management Visits for Percutaneous Coronary Intervention Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=1,577)		Comparison (N=1,592)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Part B imaging and lab payments, 90-day post-discharge period	$568	$490	$548	$550	-$81	-$142	-$19	0.01
Part B E&M visit payments, 90-day post-discharge period	$1,221	$1,118	$1,056	$1,127	-$173	-$318	-$29	0.02

DiD = Difference-in-Differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison providers.

[110] The Medicare HHA case-mix indices were revised in 2014 which resulted in lower case-mix indices than prior to 2014.

It is not surprising that PCI hospitals did not achieve significant reductions in total payments. As mentioned above, relatively few episodes involve PAC, resulting in a large portion of the total cost coming from the inpatient hospitalization (52%). To lower hospital internal costs during the inpatient stay, we would hypothesize that participants may have focused on shortening the hospital length of stay or care redesign activities that reduced internal costs during the inpatient stay. There was no statistically significant change in the inpatient length of stay though; it averaged 4 days for the BPCI participants and the comparison group during both the baseline and intervention periods.

Interviewees discussed several care redesign activities for PCI episodes, some of which focused on reducing internal costs. One hospital lowered internal costs by reducing the use of the ICU for patients with less intensive procedures. The hospital also began a new PCI approach (radial entry rather than femoral) that was shown to lower the length of stay and result in fewer complications. Several interviewees discussed other activities such as telemetry monitoring, educating long-term care staff to read monitor strips to know when they should call with concerns, and working within a multidisciplinary team to understand data to find opportunities for care improvement.

Interviewees also discussed challenges related to the PCI bundle. One interviewee noted they could not reduce supply costs through better purchasing techniques because the cost of supplies used for PCI, such as stents or catheters, do not vary much across vendors. Another interviewee noted that the number of stents varies across PCI procedures, which makes it difficult to predict and control hospital costs. Other interviewees noted they were responsible for readmissions for medically appropriate care following PCI that was not directly related to the initial PCI episode. For example, patients often need additional procedures, such as more stenting or open heart surgery, to fix concurrent issues. These additional procedures are counted as readmissions even though they are not related to a problem with the initial PCI itself, and cannot be appropriately postponed until after the BPCI episode.

d. Quality of Care

There was no difference in the quality of care for Medicare patients receiving PCI surgery at BPCI-participating hospitals and patients at comparison hospitals as measured by readmission rates, emergency department visits, and mortality from claims data, and there were no statistically significant changes in measures of functional status from patient assessments for those receiving PAC.

e. Conclusions

During the first eight quarters, the BPCI initiative resulted in few statistically significant changes in Medicare payments, utilization of health care services, and quality of care for the PCI episodes in Model 2. Representatives of BPCI-participating hospitals we interviewed indicated that care redesign for PCI episodes focuses on sending lower intensity patients to more appropriate recovery locations within the hospital after surgery and by using new, better approaches to PCI procedures. While these approaches may reduce internal hospital costs, they do not reduce Medicare payments.

10. Renal Failure

Medicare beneficiaries received 3,105 renal failure episodes of care in 55 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015).[111] Based on the subset of episodes for which we found comparison providers and patients, average Medicare payments for the inpatient hospitalization and care provided for the following 90 days did not change for episodes initiated at BPCI hospitals, compared with a matched sample of episodes from hospitals that did not participate in BPCI.[112] There is no indication that BPCI-participating hospitals altered discharge patterns or post-acute care utilization for renal failure patients, which may account for the providers' inability to generate savings relative to the comparison group. Additionally, quality of care was maintained throughout the 90-day post-discharge period (PDP). During this period, mortality rates declined for BPCI participants from baseline to intervention. Future analyses will confirm whether this improvement in mortality can be sustained.

a. Participants

During the first eight quarters of the BPCI initiative, 55 hospitals participated in the renal failure episode (13% of all Model 2 BPCI hospitals). Of those participants, 54 chose 90-day bundles and one chose 30-day bundles. There were 3,105 renal failure episodes initiated during the BPCI initiative (approximately 2% of episodes initiated across the 48 BPCI clinical episodes). Because participants were allowed to join the initiative over an extended period, providers with renal failure episodes had an average of two quarters of BPCI experience. Approximately three-quarters of the BPCI hospitals began participating in renal failure episodes in either Q2 or Q3 2015. Furthermore, 7 of the 55 hospitals stopped participating in the renal failure episode within the first eight quarters of the initiative.

Compared to hospitals *not* participating in BPCI, those that participated in the renal failure episode were more likely to be large, non-profit, and located in urban markets. BPCI hospitals were also less likely to be government-owned. Prior to BPCI, participating hospitals had a greater number of discharges for renal failure and a lower share of Medicare days across all inpatient admissions. BPCI hospitals also had higher 2011 standardized Part A payments during the anchor hospitalization and 90 days post-discharge for patients discharged with renal failure than non-participants (Exhibits 138a & 138b).

In comparison to all Model 2 BPCI hospitals, the hospitals participating in the renal failure episode were more likely to be urban and for-profit, with lower resident-to-bed ratios. BPCI hospitals participating in the renal failure episode also had a lower percentage of inpatient days for Medicare and a higher disproportionate share percentage. The hospitals participating in the

[111] As defined under the BPCI initiative, Medicare patients with end-stage renal disease are not eligible for inclusion in the renal failure bundle.

[112] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do not take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

renal failure episode also had higher standardized Medicare Part A payments for renal failure episodes in 2011 than all Model 2 BPCI hospitals (Exhibits 138a & 138b).

Matched comparison hospitals were identified for 51 of the 55 BPCI-participating hospitals in the sample.[113] The matched BPCI-participating hospitals were not statistically different from the comparison hospitals among key provider and market metrics.

Exhibit 138a & 138b: Characteristics of BPCI-Participating Hospitals with Renal Failure Episodes compared with Non-participating Hospitals, All BPCI-participating Hospitals and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Renal Failure BPCI Hospitals (N=55)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched Renal Failure BPCI hospitals (N=51)		Comparison Hospitals (N=727)+	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	37	67%	1,594	57%*	321	77%*	34	67%	485	67%
	Government	2	4%	542	20%*	32	8%	2	4%	30	4%
	For-Profit	16	29%	638	23%	66	16%*	15	29%	212	29%
Urban/Rural	Urban	54	98%	1,902	69%*	387	92%*	50	98%	714	98%
Part of Chain	Yes	30	55%	1,469	53%	218	52%	27	53%	382	53%

Characteristic	All Renal Failure BPCI Hospitals (N=55)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched Renal Failure BPCI hospitals (N=51)	Comparison Hospitals (N=727)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	312	175*	311	318	307
Medicare Days Percent	37%	42%*	39%*	37%	37%
Resident-to-bed ratio	0.07	0.05	0.12*	0.07	0.08
Disproportionate Share Percent	31%	29%	27%*	31%	33%
Number of Renal Failure Discharges, 2011	119	59*	114	122	117
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, Renal Failure, 2011	**$20,671**	**$18,686***	**$19,169***	**$20,759**	**$20,989**

PDP=post-discharge period

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 through Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals.

*Indicates differences from BPCI hospitals greater than absolute value of 0.2 standard difference when comparing all renal failure BPCI hospitals to non-participating hospitals and all Model 2 BPCI hospitals. The matched comparison hospitals are compared to the matched renal failure BPCI hospitals.

** BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers accounting for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of comparison providers is 379.

[113] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. Three participants were not included in the analysis because they were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the renal failure episode. There was one other participant for which a match could not be found within the caliper (33% of the standard deviation of the log odds propensity score).

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. A comparison of the renal failure patient population of BPCI and comparison participants from baseline to intervention suggests that BPCI providers may have treated a healthier patient population after joining BPCI (Exhibits 139a & 139b). In the six months prior to the index hospitalization, there was a statistically significant increase in the percentage of BPCI patients that had no institutional care (3.0 percentage points, $p=0.09$) and no post-acute care (7.3 percentage points, $p<0.01$) relative to the change for the comparison group. Similarly, there were statistically significant decreases in the percentage of BPCI renal failure patients who used the emergency room (-4.1 percentage points, $p=0.02$) and home health care (-4.5 percentage points, $p=0.01$) in the same period. There was also a significant reduction in the number of HCC indicators (-0.3, $p<0.01$) and the proportion of patients that were disabled among BPCI providers relative to the comparison group (-2.9 percentage points, $p=0.02$). However, BPCI providers treated older patients since joining BPCI; the share of renal failure patients aged 20-64 decreased 2.9 percentage points ($p=0.01$) for BPCI participants relative to the comparison group.

Exhibits 139a & 139b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Renal Failure, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Renal Failure Hospitals (N=51)				Comparison Hospitals (N=727)+				DiD	Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	N	%		LCI	UCI	
Age	20-64	394	12%	317	11%	410	12%	415	14%	-2.9	-5.2	-0.6	0.01
	65-79	1,223	36%	1,147	39%	1,217	36%	1,152	39%	-0.3	-3.7	3.0	0.84
	80+	1,735	52%	1,463	50%	1,725	51%	1,360	46%	3.2	-0.3	6.7	0.07
Gender	Female	1,869	56%	1,593	55%	1,838	55%	1,540	53%	0.9	-2.6	4.4	0.62
Medicaid and Disability	% Eligible for Medicaid	1,002	30%	746	25%	1,060	32%	837	29%	-1.4	-4.6	1.8	0.40
	% Disability, no ESRD	438	13%	346	12%	446	13%	439	15%	-2.9	-5.3	-0.6	0.02
Utilization Six Months Prior to Anchor Stay	Inpatient Acute Care Hospital	1,353	40%	1,109	38%	1,231	37%	1,077	37%	-2.5	-5.9	0.9	0.14
	Emergency Room Admission	1,061	32%	993	34%	1,058	32%	1,109	38%	-4.1	-7.4	-0.7	0.02
	Home Health	1,076	32%	844	29%	968	29%	880	30%	-4.5	-7.7	-1.2	0.01
	Inpatient Rehabilitation Facility	96	3%	99	3%	87	3%	72	2%	0.7	-0.5	1.8	0.27
	Skilled Nursing Facility	594	18%	499	17%	522	16%	436	15%	0.0	-2.6	2.6	1.00
	Psychiatric Hospital	65	2%	55	2%	67	2%	49	2%	0.3	-0.7	1.2	0.59
	Long-term Care Hospital	67	2%	37	1%	38	1%	27	1%	-0.5	-1.3	0.3	0.20
	Institutional Nursing Facility**	837	25%	631	22%	733	22%	557	19%	-0.6	-3.5	2.3	0.70
	No Institutional Care	1,890	56%	1,730	59%	2,015	60%	1,752	60%	3.0	-0.5	6.4	0.09
	No Post-acute Care	1,161	35%	1,105	38%	1,313	39%	1,025	35%	7.3	3.9	10.6	<0.01

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	3.39	2.57	3.25	2.54	3.14	2.46	3.27	2.52	-0.27	-0.44	-0.11	<0.01

DiD = difference-in-differences, LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

These estimates are developed using a difference-in-differences (DiD) framework, which quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. Changes are observed before and after the intervention began.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention) for BPCI EIs and matched comparison providers.

* "Count" represents the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment. All other measures are based on Medicare claims.

+This represents the weighted number of matched comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of comparison providers is 379.

c. Medicare Payments

The change in total payments from baseline to intervention for BPCI renal failure episodes is not statistically significant relative to matched comparison episodes (Exhibit 140). There is also no statistically significant change in utilization for renal failure patients in BPCI episodes relative to comparison episodes, which is consistent with the lack of payment-related findings.

Exhibit 140: Impact of BPCI on Total Payments for Inpatient Hospitalization and 90-day PDP ($2015), Renal Failure Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=2,873)		Comparison (N=2,871)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total allowed payment amount, IP and 90-day PDP	$26,949	$24,532	$27,595	$25,210	-$32	-$1,351	$1,287	0.96

DiD = difference-in-differences, LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level. PDP=post-discharge period

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors. Medicare payments are expressed in 2015 dollars, the result of adjusting actual dollar amounts based on the change in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison hospitals.

There were relative increases in two Part B payment categories during the 90-day PDP among BPCI renal failure episodes, though these services represent a relatively small share of total episode costs (Exhibit 141). From the baseline to intervention period, payments for imaging and laboratory services increased by $37 more for BPCI episodes than for comparison episodes (p=0.04), and payments for procedures increased $49 more for BPCI episodes relative to the comparison episodes (p=0.04).

Exhibit 141: Impact of BPCI on Medicare Part B Payments during the 90-day PDP ($2015), Renal Failure Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=2,873)		Comparison (N=2,871)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Imaging and laboratory services, 90-day PDP	$460	$438	$469	$410	$37	$2	$72	0.04
Procedures, 90-day PDP	$349	$318	$386	$306	$49	$2	$95	0.04

DiD = difference-in-differences, LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level. PDP=post-discharge period

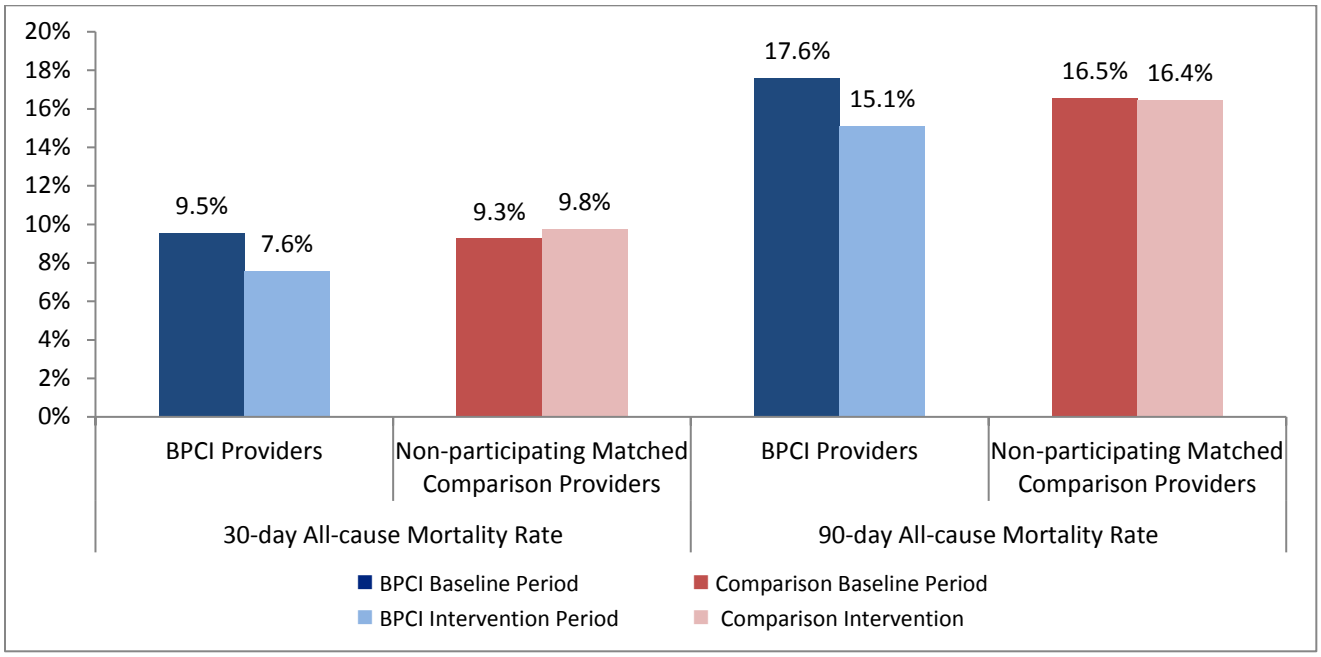
Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors. Medicare payments are expressed in 2015 dollars, the result of adjusting actual dollar amounts based on the change in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison hospitals.

d. Quality of Care

The change in hospital readmission rates and emergency department visits from baseline to intervention was not statistically different for BPCI episodes relative to the comparison group. As shown in Exhibit 142, mortality rates declined among BPCI renal failure episodes during the 30- and 90-day PDP. Additional time and provider experience with renal failure episodes will indicate whether the improvements with regards to decreased mortality rates can be sustained. Finally, there were no statistically significant changes in the proportion of patients with improved functional status among those discharged to post-acute care settings relative to the comparison group.

Exhibit 142: All-cause Mortality Rates among Renal Failure Episodes, Matched BPCI and Comparison Providers



Source: Lewin analysis of Medicare payment and enrollment data for BPCI Model 2 hospital participants and a matched comparison group.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015.

e. Conclusion

Results from the first two years of BPCI indicate that the initiative did not have a statistically significant impact on total Medicare payments for renal failure episodes, despite indications that BPCI providers treated healthier patients in the intervention period compared to the change in patient mix for the comparison group. The cost of procedures and imaging and laboratory services increased more for BPCI providers during the 90-day PDP, but this did not affect overall episode spending. Mortality rates improved among BPCI providers during the 30 and 90 days following the anchor hospitalization relative to the comparison group. Given more time and initiative experience, future analyses will reveal whether providers can generate savings while maintaining quality of care.

11. Sepsis

Sepsis is one of the leading causes of inpatient admissions for Medicare beneficiaries. As such, sepsis care has been a focus of national quality improvement efforts for over a decade. Aside from the BPCI sepsis bundle, CMS also developed a sepsis-related quality measure, which will begin to affect Medicare hospital payments in 2017.[114]

Medicare beneficiaries received 11,509 sepsis episodes of care with 103 BPCI-participating acute care hospitals (ACHs) through the first eight quarters of the BPCI initiative (from October 2013 through September 2015). This issue brief describes the sepsis episode findings from Model 2 during this period.

When compared to a matched comparison group, we observed no statistically significant change in total Medicare payments for sepsis episodes. There were also limited effects on spending by type of service in the 90-day post-discharge period (PDP).

There were no statistically significant changes in the quality of care for sepsis episodes as measured by emergency department visits, unplanned readmissions, mortality rates, and measures of functional status from patient assessments, with one exception. Only one of ten assessment measures for patients treated by post-acute care (PAC) providers revealed a statistically significant BPCI impact on activities of daily living (ADL). There was a relative decrease in the proportion of BPCI patients treated by skilled nursing facilities (SNFs) who improved in their ability to provide self-care.

Survey data from two cohorts of Medicare beneficiaries treated by BPCI and comparison providers revealed potential indications of negative effects of BPCI on quality of care. The results from the earlier cohort indicated that several self-reported changes in functional limitations declined from before the hospitalization for sepsis to after the episode for BPCI patients relative to comparison respondents. Moreover, BPCI patients reported worse perceptions of care experience. However, these differences did not persist with the most recent wave of surveyed beneficiaries. While not statistically significant, BPCI respondents in both survey waves were also less likely to be "extremely" or "quite a bit" satisfied with their recovery. It will be important to continue to monitor these changes as additional data becomes available.

a. Background

Sepsis is a leading cause of hospital admissions for Medicare beneficiaries and can lead to death. In fiscal year (FY) 2014, there were over 600,000 Medicare admissions in one of the three Medicare Severity-Diagnosis Related Groups (MS-DRGs) that fall within the sepsis bundle for

[114] For more information on the Hospital IQR program: https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1138115987129

BPCI.[115, 116] These inpatient hospital admissions amounted to approximately $8 billion in Medicare payments.

Given the high volume and costs, sepsis has been a focus for care improvement for years. In 2001, an international Surviving Sepsis Campaign developed clinical guidelines to improve care and reduce mortality rates for sepsis patients.[117] The guidelines were the basis for a sepsis measure required by CMS to be collected from hospitals beginning October 2015 as part of the Hospital Inpatient Quality Reporting (IQR) Program.[118] Starting October 2017, performance on that measure will affect Medicare hospital payment under the inpatient prospective payment system.

b. Participants

During the first eight quarters of the BPCI initiative, 103 hospitals (24% of all Model 2 BPCI hospitals) participated in the sepsis episode and almost exclusively chose 90-day episodes; 30-day and 60-day episodes were chosen by just one provider each. Anecdotal evidence from site visits and interviews indicated that high volume and national attention from the campaign and measure development efforts were reasons to participate in this episode type.[119]

The BPCI hospitals had 11,509 sepsis episodes during the initiative (approximately 8% of all episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the sepsis results are based on an average of two quarters of experience. Approximately 32% of the hospitals began participating in sepsis episodes in Q3 2015; 81% joined in either Q2 or Q3 2015. Nine hospitals stopped participating in sepsis in the first eight quarters of the initiative.

Compared with hospitals *not* participating in BPCI, hospitals that participated in the sepsis episode were less likely to be government-owned, more likely to be urban and had a higher average resident-to-bed ratio (Exhibits 143a & 143b). Participating hospitals also tended to have more beds, but fewer Medicare days as a proportion of all inpatient days. Prior to joining BPCI, participating hospitals had more sepsis discharges on average. Furthermore, participating hospitals had higher average 2011 standardized Part A payments during the anchor hospitalization plus 90-days PDP than non-participants ($28,714 versus $25,214).

Compared with all Model 2 BPCI hospitals (regardless of whether they participated in sepsis), the hospitals participating in sepsis were more likely to be in urban locations and for-profit.

---

[115] Centers for Medicare and Medicaid Services. (2014). Inpatient Charge Data FY 2014. Retrieved from https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Inpatient2014.html

[116] The three MS-DRGs are 870 - septicemia or severe sepsis with MV 96+ hours, 871 – septicemia or severe sepsis without MV 96+ hours with MCC, and 872 - septicemia or severe sepsis without MV 96+ hours without MCC.

[117] For more information on the campaign see http://www.survivingsepsis.org/Pages/default.aspx.

[118] For more information on the Hospital IQR program see https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1138115987129

[119] See Methods section for more information on the BPCI participant site visits and interviews.

Participating hospitals also had higher average standardized Part A payments during the anchor hospitalization plus 90-day PDP than other Model 2 participants ($28,714 versus $26,282).

Matched comparison hospitals were identified for 97 of the 103 BPCI-participating hospitals in the sample.[120] The characteristics of hospitals participating in the sepsis episode and the matched comparison hospitals were very similar.

Exhibits 143a & 143b: Characteristics of BPCI-participating Hospitals with Sepsis Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospital EIs and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Sepsis BPCI Hospitals (N=102)**		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)		Matched Sepsis BPCI Hospitals (N=97)		Comparison Hospitals (N=1,231+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	68	67%	1,594	57%	321	77%*	65	67%	763	62%
	Government	4	4%	542	20%*	32	8%	4	4%	49	4%
	For-Profit	30	29%	638	23%	66	16%*	28	29%	419	34%
Urban/Rural	Urban	99	97%	1,902	69%*	387	92%*	94	97%	1,169	95%
Part of Chain	Yes	50	49%	1,469	53%*	218	52%	48	49%	640	52%

Characteristic	All Sepsis BPCI Hospitals (N=102)**	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)	Matched Sepsis BPCI Hospitals (N=97)	Comparison Hospitals (N=1,231+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	283	175*	311	281	283
Medicare Days Percent	38%	42%*	39%	38%	40%
Resident-to-bed ratio	0.09	0.05*	0.12	0.09	0.09
Disproportionate Share Percent	30%	29%	27%	29%	30%
Number of Sepsis Discharges, 2011	167	93*	173	166	186
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, sepsis episodes, 2011	**28,714**	**25,214***	**26,282***	**28,644**	**28,920**

PDP = post-discharge period

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All Sepsis BPCI" hospitals. "Comparison" hospitals are compared to "Matched Sepsis BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers were matched to multiple BPCI hospitals. The unique number of matched comparison providers is 586.

[120] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were five participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the sepsis episode or had fewer than five relevant discharges in 2011 and 2012. There was one other participant for which there was not a match within the caliper.

c. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. We observed no statistically significant changes that suggest BPCI-participating hospitals saw a healthier mix of sepsis patients relative to comparison hospitals.

There are several indications, however, of changes to a less healthy sepsis patient population in the BPCI-participating hospitals relative to the comparison group (Exhibits 144a & 144b). There was a statistically significant relative decrease in the percentage of BPCI patients aged 65 to 79 (-1.6 percentage points, p=0.09).[123] In addition, there was a relative increase in the proportion that was disabled (1.2 percentage points, p=0.08). In both cases the difference is driven by small changes in the proportion for the comparison group. There was also a relative increase in the number and intensity of comorbidities for BPCI patients (represented by Hierarchical Condition Category (HCC) scores) from the baseline to the intervention period in comparison to the matched comparison group (0.1 percentage points, p=0.07). With respect to prior health care utilization, there was a statistically significant relative increase in the proportion of BPCI patients who used an institutional nursing facility (2.6 percentage points, p<0.01) and a SNF (1.4 percentage points, p=0.06) in the six months prior to the index hospitalization. There were no statistically significant changes in other measures of prior care use.

[123] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Exhibits 144a & 144b: Characteristics of BPCI and Matched Comparison Beneficiaries with a Hospitalization for Sepsis, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Sepsis Hospital				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	1,758	14%	1,558	15%	1,866	15%	1,540	15%	1.1	-0.3	2.4	0.11
	65-79	4,714	39%	4,101	39%	4,692	39%	4,245	40%	-1.6	-3.4	0.2	0.09
	80+	5,686	47%	4,825	46%	5,600	46%	4,699	45%	0.5	-1.3	2.3	0.60
Gender	Female	6,613	54%	5,644	54%	6,669	55%	5,648	54%	0.4	-1.4	2.3	0.65
Medicaid and Disability	% Eligible for Medicaid	4,447	37%	3,362	32%	4,254	35%	3,068	29%	1.2	-0.5	3.0	0.17
	% Disability, no ESRD	1,948	16%	1,675	16%	2,070	17%	1,652	16%	1.2	-0.1	2.6	0.08
Utilization Six Months Prior to Anchor Hospitalization	Inpatient Acute Care Hospital	5,119	42%	4,119	39%	4,889	40%	3,902	37%	0.2	-1.6	2.0	0.85
	Emergency Room Admission	3,865	32%	3,568	34%	3,954	33%	3,530	34%	1.1	-0.6	2.8	0.22
	Home Health	3,256	27%	2,720	26%	3,347	28%	2,811	27%	-0.1	-1.8	1.5	0.89
	Inpatient Rehabilitation Facility	325	3%	261	2%	308	3%	252	2%	-0.1	-0.6	0.5	0.86
	Skilled Nursing Facility	2,939	24%	2,287	22%	2,666	22%	1,900	18%	1.4	-0.1	3.0	0.06
	Psychiatric Hospital	189	2%	170	2%	183	2%	136	1%	0.3	-0.2	0.7	0.23
	Long-term Care Hospital	431	4%	272	3%	417	3%	272	3%	-0.1	-0.8	0.5	0.72
	Institutional Nursing Facility*	4,489	37%	3,393	32%	4,052	33%	2,746	26%	2.6	0.9	4.3	0.00
	No Institutional Care	6,554	54%	6,015	57%	6,787	56%	6,268	60%	-0.5	-2.3	1.3	0.59
	No Post-acute Care	4,373	36%	3,933	38%	4,366	36%	3,999	38%	-0.7	-2.5	1.1	0.45

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	3.73	2.98	3.63	2.90	3.67	3.01	3.46	2.88	0.10	-0.01	0.21	0.07

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

d. Medicare Payments and Utilization

During the baseline period, Medicare payments for services provided during the sepsis anchor hospitalization plus 90-day PDP averaged $34,590 ($2015) for episodes initiated by BPCI providers (Exhibit 145). Payments decreased by 8.3% from the baseline to intervention period for BPCI episodes (to $31,733). A similar change occurred for comparison hospitals during this time (a decrease of 7.8% from $35,073 to $32,330). Thus, BPCI was not associated with a statistically significant impact on overall Medicare episode spending.

Exhibit 145: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day PDP, Sepsis, Model 2, Q4 2011-Q3 2015

	BPCI (N=10,209)		Comparison (N=10,222)		DiD estimate			
Measure	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total allowed payment amount, IP and 90-day PDP	$34,590	$31,733	$35,073	$32,330	-$114	-$1,250	$1,023	0.84

DiD = difference-in-differences; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; IP = anchor inpatient stay; PDP = post-discharge period.

Note: Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome. Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison providers.

There were shifts in payments by type of service within the bundle (Exhibit 146). However, these shifts were relatively small with the only statistically significant changes occurring for home health agency (HHA) payments and for imaging and lab services, both increasing relative to the respective changes from the baseline to intervention period in the comparison group ($96.70, $p = 0.01$ for HHA and $18.52, $p = 0.10$ for imaging). Combined, these services accounted for less than five percent of the total episode payments.

Exhibit 146: Difference-in-Differences Results for Medicare Payments ($2015) by Type Service within the 90-day Post-Discharge Period (PDP), Sepsis, Model 2, Q4 2011-Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Home Health Agency Payment, 90-day PDP	10,276	10,322	$1,081	$1,128	$1,130	$1,080	$97	$22	$171	0.01
Imaging and Lab Payment Amount, 90-day PDP	10,209	10,223	$394	$380	$426	$393	$19	-$3	$40	0.10

DiD = difference-in-differences; PDP = post-discharge period; LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

Note: Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome. Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison providers.

As with Medicare payments, BPCI had limited effects on utilization during the 90-day PDP. There was a statistically significant relative increase in the hospital length of stay for BPCI episodes (0.18 day, p=0.06). The length of stay decreased by a lower amount under BPCI, 4.2% (from 7.2 days to 6.9) compared to 6.9% for the comparison group (7.1 to 6.6) from the baseline to intervention period. There was also a slightly larger relative increase in home health visits for BPCI episodes (though not statistically significant), which was consistent with the relative increase in spending.

e. *Quality of Care*

To understand the impact of BPCI on the quality of care for patients with sepsis, we examined Medicare claims and enrollment data, patient assessments, and beneficiary survey data. The survey data provided evidence of negative changes in quality from the baseline to intervention period; however, the results were not consistent across the two survey waves.

From the baseline to the intervention period, changes in the rates of unplanned readmissions, emergency department visits, and mortality following the anchor stay for sepsis episodes were not statistically different for episodes initiated in BPCI-participating hospitals and matched comparison hospitals.

We used patient assessment instruments to compare the functional status of sepsis patients treated in HHAs, SNFs and inpatient rehabilitation facilities (IRFs).[124] There were no differences between BPCI and comparison patients treated in HHAs and IRFs who received care long enough to have two assessments in functional assessment measures. Only one measure for

[124] There were five measures for HHA patients relating to improved bathing, ambulation, upper body dressing, lower body dressing, and bed transferring. There were three measures for SNF patients relating to improved overall function, self-care and mobility. There were two measures for inpatient rehabilitation facility patients relating to improved mobility and self-care.

patients treated in SNFs exhibited a statistically significant difference in changes between BPCI and matched comparison patients from the baseline to intervention periods. The proportion of BPCI patients that received SNF care who improved their self-care ability decreased from 28.7% to 27.8%, while the matched comparison group patients increased from 29.6% to 32.4% (-3.77 percentage points, p=0.04).[125]

We surveyed a sample of BPCI and comparison patients to provide additional insights regarding impacts on functional status, as well as self-reported health and care experience indictors. Beneficiaries with a sepsis episode initiated in Model 2 ACHs were surveyed in Waves 4 and 5, along with a matched comparison group.[126] Wave 4 BPCI respondents reported several worse functional changes from before to after their episode than did comparison respondents, and also reported worse perceptions of care experience (Exhibit 147). However, neither of these differences persisted into Wave 5.

In Wave 4, BPCI respondents reported significantly lower rates of improvement relative to comparison respondents from before to after the episode in bathing, dressing, using the toilet or eating (52.2 % compared to 61.3% for a difference of -9.0 percentage points, p<0.01) and using stairs (-7.8 percentage points, p=0.01), relative to the matched comparison group. They also reported significantly higher rates of functional decline in bathing, dressing, using the toilet or eating (5.3 percentage points, p=0.09) and ability to walk without rest (6.7 percentage points, p=0.05), and increased dependence on a mobility-assistance device (5.0 percentage points, p=0.09). However, differences in these functional outcomes were not significant in Wave 5, and in some cases changed direction (indicating better outcomes for BPCI respondents). Although BPCI respondents in Wave 5 were less likely to report improvement in pain limiting their everyday activities (-6.2 percentage points, p=0.07), the differences between the two groups on other functional measures were generally small and insignificant.

BPCI respondents in Wave 4 also reported significantly worse care experiences for four of ten measures, relative to comparison respondents: receiving conflicting advice (-8.0 percentage points, p=0.03), services appropriate for level of care needed (-8.4 percentage points, p=0.03), staff speaking in the respondent's preferred language (-4.1 percentage points, p=0.04), and staff considering respondent preferences for post-acute care (-3.8 percentage points, p=0.10). However, all four measures of care experience that were significantly worse among BPCI respondents in Wave 4 were not statistically significant among respondents in Wave 5. Additionally, although the differences were not statistically significant in either wave, BPCI respondents were approximately 4 percentage points less likely to be "extremely" or "quite a bit" satisfied with their recovery in both waves. It will be important to continue to monitor this change as additional data becomes available.

[125] "Improvement" means improvement or staying in the best possible status.

[126] Episodes of sepsis were sampled twice in each survey wave over two consecutive months. Each of the two samples was drawn from hospital discharges that occurred within a single month: Wave 4, May/June 2015; Wave 5, October/November 2015.

Exhibit 147: Select Beneficiary Survey Outcomes, Model 2 ACH, Sepsis Episodes, May/June 2015 and October/November 2015

Measure		Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
			BPCI	Comparison	BPCI	Comparison	Point Estimate	LCI	UCI	p-value
Functional Status	Improvement in bathing, dressing, using the toilet, or eating	4	322	341	52.2	61.3	-9	-15.5	-2.5	<0.01
		5	370	393	61	57.6	3.4	-3.2	9.9	0.31
	Improvement in using stairs	4	315	340	25.7	33.5	-7.8	-13.9	-1.6	0.01
		5	367	380	31.4	32	-0.6	-6.8	5.7	0.86
	Improvement in pain limiting regular activities	4	324	342	42.6	44	-1.4	-8.5	5.7	0.7
		5	369	398	40.9	47.2	-6.2	-13.3	0.8	0.08
	Decline in bathing, dressing, using the toilet, or eating	4	322	341	30.6	25.3	5.3	-0.8	11.4	0.09
		5	370	393	21.6	26.1	-4.6	-10.3	1.2	0.12
	Decline in use of mobility device (more likely to use mobility device)	4	331	345	48.9	43.9	5	-0.8	10.8	0.09
		5	373	400	45.7	46.5	-0.8	-6.3	4.6	0.77
	Decline in walking without rest	4	320	336	41.5	34.9	6.7	0	13.3	0.05
		5	365	396	38.8	37.6	1.2	-5	7.4	0.7
Patient Care Experience	Never received conflicting advice from medical staff	4	327	351	59.7	67.7	-8	-15.1	-0.9	0.03
		5	382	411	63.8	61.3	2.5	-4.8	9.9	0.5
	Patient always received appropriate level of care	4	329	353	51.4	59.8	-8.4	-15	-0.9	0.03
		5	388	415	51.9	54.4	-2.5	-10.1	5.1	0.52
	Staff always used patient's preferred language	4	332	357	87.1	91.2	-4.1	-8	-0.1	0.04
		5	392	421	86.9	90.5	-3.5	-8.2	1.1	0.13
	Patient preferences considered for services after discharge	4	282	313	89.9	93.7	-3.8	-8.2	0.7	0.1
		5	325	360	90.6	90.2	0.4	-4.2	5	0.86
	Satisfaction with overall recovery	4	310	342	60.6	64.9	-4.3	-11.5	2.9	0.24
		5	377	393	62.5	66.7	-4.2	-11.5	3.2	0.27

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

Note: This table reports all functional status and care experience measures that have a statistically significant difference between the intervention and comparison groups at the 10% or 5% level in at least one wave.

The improvement indicator takes a value of 1 if a patient moved to a better functional status level after the episode or if the patient recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response. The decline indicator takes a value of 1 if the patient moved to a worse functional status group after the episode or if the patient recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey. These indicators are assigned a value of 0 otherwise.

Source: Lewin analysis of beneficiary survey data for episodes that began May, June, October, and November 2015 for BPCI EIs and the matched comparison provider

f. Conclusion

We observed no statistically significant change in total Medicare payments for BPCI sepsis episodes relative to comparison episodes, and small changes in HHA payments and imaging and lab payments. There were indications that the sepsis patient population treated in BPCI hospitals may have shifted to a less healthy patient mix under the initiative. However, the estimated BPCI impact on the outcomes above take into account any changes in patient mix available in claims data. The interpretation of the quality of care indicators analyzed is less clear. The claims and assessment-based measures provided little evidence that BPCI affected quality of care for patients with sepsis. There were no significant changes in emergency department visits, unplanned readmissions, or mortality rates. Only one of ten assessment measures on patients treated by PAC providers revealed a statistically significant BPCI impact on activities for daily living (ADL). There was a relative decrease in the proportion of BPCI patients treated in SNFs who improved their ability to self-care.

Self-reported quality measures from the beneficiary survey, however, indicate that BPCI may be negatively affecting the quality of care for sepsis patients. The initial wave of respondents indicated deterioration in several self-reported changes in functional limitations from before to after the episode for BPCI patients relative to comparison respondents. Moreover, BPCI patients reported worse perceptions of care experience. However, these differences did not persist with the most recent wave of surveyed sepsis patients. While not statistically significant, BPCI respondents were also less likely to be "extremely" or "quite a bit" satisfied with their recovery, in both survey waves analyzed for this report. The characteristics of sepsis patients and these quality indicators warrant continued monitoring to detect changes in quality of care as well as potential changes in patient characteristics that may explain differences in patient outcomes.

12. Simple Pneumonia and Respiratory Infections

a. Summary

There were 10,134 simple pneumonia and respiratory infection (SPRI) episodes of care for Medicare beneficiaries initiated in 111 BPCI-participating acute care hospitals in the first eight quarters of the BPCI initiative (from October 2013 through September 2015). BPCI had no statistically significant impact on Medicare payments during the initial hospital stay at participating providers and on services during the subsequent 90-day post-discharge period (PDP).[127] BPCI also had no statistically significant impact on quality of care as measured by readmission rates, emergency department visits, mortality rates, and by measures of functional status from patient assessments, except for a relative increase in the share of BPCI beneficiaries treated by a home health agency (HHA) who experienced an improvement in bathing. Finally, there were mixed results with respect to survey data on functional status and patient experience. The results were inconsistent across measures and across patient survey waves.

[127] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and trended to 2015. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see http://innovation.cms.gov/Files/x/BPCI_Model2Background.pdf.

b. Participants

During the first eight quarters of the BPCI initiative, 111 hospitals (26% of all Model 2 BPCI hospitals) participated in the SPRI episode, four of which chose 30-day episodes and 107 of which chose 90-day episodes. There were 10,134 SPRI episodes initiated by these hospitals during the first eight quarters of the BPCI initiative (approximately 7% of episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the SPRI results are based on an average of three quarters of experience. Approximately 28% of the hospitals began participating in SPRI episodes in Q3 2015, and 73% joined in either Q2 or Q3 2015. Fourteen of the 111 hospitals stopped participating in the SPRI episode in the first eight quarters of the initiative.

Compared with hospitals that did *not* participate in BPCI, hospitals that participated in the SPRI episode had more beds and a lower proportion of Medicare days on average (Exhibits 148a & 148b). They were also more likely to be non-profit and in urban locations. Prior to joining BPCI, participating hospitals had more SPRI discharges. Further, compared with non-participating hospitals, participating hospitals had higher average 2011 standardized Part A payments for services during the impatient hospital stay plus 90-day PDP ($19,128 compared to $17,601).

Hospitals that participated in SPRI were similar to all BPCI hospitals, with some exceptions. SPRI hospitals were less likely to be government-owned and had lower average resident-to-bed ratios. While SPRI and all BPCI hospitals had a similar average number of SPRI discharges in 2011, hospitals participating in SPRI had higher average Medicare Part A payments for services during the inpatient hospitalization plus 90-day PDP than all Model 2 BPCI hospitals ($19,128 compared to $18,363).

Matched comparison hospitals were identified for 103 of the SPRI BPCI-participating hospitals in the sample.[128] The characteristics of SPRI hospitals and the matched comparison hospitals were similar.

[128] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were five participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the SPRI episode or had fewer than five relevant discharges in 2011. There were three other participants for which there was not a match within the caliper.

Exhibits 148a & 148b: Characteristics of BPCI-participating hospitals with Simple Pneumonia and Respiratory Infection Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals Model 2, Q4 2013 - Q3 2015

Characteristic		All SPRI BPCI Hospitals (N=110**)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419**)		Matched SPRI BPCI hospitals (N=103)		Comparison Hospitals (N=1,419+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	86	78%	1,594	57%*	321	76%	81	79%	1,107	78%
	Government	2	2%	542	20%*	32	8%*	2	2%	28	2%
	For-Profit	22	20%	638	23%	66	16%	20	19%	284	20%
Urban/Rural	Urban	103	94%	1,902	69%*	387	92%	96	93%	1,305	92%
Part of Chain	Yes	54	49%	1,469	53%	218	52%	52	50%	710	50%

Characteristic	All SPRI BPCI Hospitals (N=110**)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419**)	Matched SPRI BPCI hospitals (N=103)	Comparison Hospitals (N=1,419+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	274	175*	311	277	270
Medicare Days Percent	39%	42%*	39%	39%	39%
Resident-to-bed ratio	0.07	0.05	0.12*	0.08	0.08
Disproportionate Share Percent	28%	29%	27%	28%	29%
Number of SPRI Discharges, 2011	208	130*	214	209	204
Standardized Part A Allowed Payment inpatient stay plus 90-day post-discharge period, SPRI, 2011	**19,128**	**17,601***	**18,363***	**19,049**	**19,204**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

* Indicates the standardized mean difference between the two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All SPRI BPCI" hospitals. "Comparison" hospitals are compared to "Matched SPRI BPCI" hospitals.

** BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers were matched to multiple BPCI hospitals. The unique number of matched comparison providers is 723.

c. *Change in Patient Mix*

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. We observed no statistically significant changes that suggest BPCI-participating hospitals had a change in SPRI patient mix relative to comparison hospitals (see Exhibits 149a & 149b).

Exhibits 149a & 149b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Simple Pneumonia and Respiratory Infection, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	1,814	12%	1,103	12%	1,816	12%	1,139	12%	-0.4	-1.5	0.8	0.54
	65-79	5,310	34%	3,386	36%	5,388	35%	3,358	35%	0.8	-0.9	2.5	0.36
	80+	8,490	54%	5,034	53%	8,410	54%	5,026	53%	-0.4	-2.2	1.4	0.64
Gender	Female	8,351	53%	5,077	53%	8,383	54%	5,103	54%	-0.1	-1.9	1.7	0.94
Medicaid and Disability	% Eligible for Medicaid	4,406	28%	2,331	24%	4,436	28%	2,419	25%	-0.7	-2.3	0.9	0.37
	% Disability, no ESRD	2,015	13%	1,200	13%	2,025	13%	1,230	13%	-0.3	-1.5	1.0	0.68
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	5,470	35%	3,389	36%	5,018	32%	3,068	32%	0.5	-1.2	2.2	0.58
	Emergency Room Admission	5,083	33%	3,333	35%	5,099	33%	3,332	35%	0.1	-1.6	1.8	0.90
	Home Health	4,043	26%	2,503	26%	3,911	25%	2,547	27%	-1.3	-2.9	0.3	0.10
	Inpatient Rehabilitation Facility	376	2%	229	2%	380	2%	258	3%	-0.3	-0.8	0.3	0.33
	Skilled Nursing Facility	2,627	17%	1,635	17%	2,279	15%	1,425	15%	0.0	-1.3	1.3	0.97
	Psychiatric Hospital	239	2%	138	1%	238	2%	114	1%	0.2	-0.2	0.7	0.26
	Long-term Care Hospital	209	1%	75	1%	186	1%	111	1%	-0.5	-0.9	-0.1	0.01
	Institutional Nursing Facility*	3,941	25%	2,243	24%	3,519	23%	1,995	21%	-0.1	-1.6	1.4	0.90
	No Institutional Care	9,650	62%	5,890	62%	10,100	65%	6,173	65%	-0.1	-1.8	1.6	0.92
	No Post-acute Care	6,345	41%	3,818	40%	6,577	42%	3,852	40%	1.1	-0.6	2.9	0.21

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.89	2.53	2.98	2.56	2.76	2.48	2.81	2.52	0.05	-0.04	0.14	0.27

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers. DiD = difference-in-differences; ESRD = end-stage renal disease.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

d. Medicare Payments and Utilization

During the first eight quarters of the BPCI initiative, the change in total Medicare payments ($2015) for the inpatient hospitalization and all care delivered during the 90-day PDP for BPCI episodes was not statistically different from the change for episodes from a matched comparison group of providers not participating in BPCI (see Exhibit 150).

Exhibit 150: Impact of BPCI on Total Standardized Allowed Payment ($2015) Inpatient Hospitalization and 90-day PDP for SPRI Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=9,374)		Comparison (N=9,369)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total Payment, Inpatient stay plus 90-day post-discharge period	$24,628	$22,664	$24,836	$23,096	-$224	-$1,005	$557	0.57

DiD = difference-in-differences, LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison providers.

These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Anecdotal evidence from site visits and interviews provides some insight into the lack of statistically significant impacts on payments.[129] One interviewee cited the complexity of SPRI patients, who can have multiple diverse co-morbidities, including congestive heart failure and chronic obstructive pulmonary disease. Thus, specific clinical strategies that only focus on SPRI needs may not work well. Another interviewee cited the lack of opportunities for reducing device or implant costs (compared to certain surgical episodes, such as those involving joint replacements) as reasons why it was difficult to reduce total payments.

There were no statistically significant shifts in payments by type of service within the 90-day PDP, with the exception of outpatient therapy services. Spending for BPCI patients for outpatient therapy services decreased 23.4% from $149 in the baseline period to $114 in the intervention period, while the comparison group decreased by 8.3 percent from $149 to $137. The result is that BPCI was associated with a relative decrease of $22 (p=0.05).[130] Because spending for outpatient therapy makes up a small proportion (less than 1%) of total SPRI episode spending, this relatively small reduction did not appreciably affect total payments for the episode.

There was a statistically significant decrease of 2.1 days (p=0.03) in skilled nursing facility (SNF) days among SNF users for beneficiaries treated at BPCI hospitals relative to the matched comparison group. The number of SNF days decreased from 32.0 in the baseline period to 30.5

[129] See Methods section for more information on the BPCI participant site visits and interviews.

[130] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

in the intervention period for BPCI participants, compared to an increase of from 33.1 to 33.6 for the matched comparison group. Although it was not statistically significant, there was a $235 decline in SNF spending for BPCI episodes relative to the matched comparison group.

There was also an increase of 0.1 days (p=0.07) in the average inpatient hospital length of stay for SPRI beneficiaries treated by BPCI hospitals relative to those treated by matched comparison hospitals. This was driven by a smaller decrease in the length of stay for BPCI providers (from 5.7 to 5.4 days) compared to comparison providers (from 5.8 to 5.4) from the baseline to intervention periods.

e. Quality of Care

There was no difference in the quality of care for Medicare patients with SPRI treated by BPCI-participating hospitals and patients at comparison hospitals, as measured by Medicare claims and patient assessments, with one exception. While BPCI and comparison group beneficiaries reported similar levels of satisfaction with their overall recovery in the beneficiary survey, there were inconsistent results in other outcomes.

Changes from the baseline to intervention period in unplanned readmission rates, emergency department visits and mortality rates were explored using claims and enrollment data. From the baseline to the intervention period, the change in rates of unplanned readmissions, emergency department visits, and mortality following the anchor stay for SPRI episodes was not statistically different for episodes initiated by BPCI-participating hospitals and matched comparison hospitals.

The functional status of patients treated by HHAs, SNFs and inpatient rehabilitation facilities (IRFs) were assessed using the respective patient assessment instruments.[131] BPCI was not associated with significant differences in functional status outcomes from patient assessments relative to the comparison group, with the exception of a 4.6 percentage point (p=0.05) improvement in bathing for patients discharged to an HHA. The share of BPCI patients discharged to an HHA who experienced an improvement in bathing increased from 59.5% to 63.3%, while the share in the comparison group decreased from 59.7% to 58.8%.[132] There were no other statistically significant results, although the relative differences for BPCI patients were in a positive direction for four of the five HHA measures.

Beneficiaries treated by BPCI and comparison hospitals were surveyed in two waves (4 and 5) to provide additional insights regarding impacts on functional status and self-reported health and care experiences.[133] Relative to a matched comparison group, there were few consistent results (Exhibit 151).

With respect to functional status, in Wave 4 BPCI respondents were significantly more likely than comparison respondents to report a decline from before to after their episode in bathing,

[131] There were five measures for home health patients relating to improved bathing, ambulation, upper body dressing, lower body dressing, and bed transferring. There were three measures for skilled nursing facility patients relating to improved overall function, self-care and mobility. There were two measures for inpatient rehabilitation facility patients relating to improved mobility and self-care.

[132] "Improvement" means improvement or staying in the best possible status.

[133] SPRI patients were sampled twice each survey wave over two consecutive months: Wave 4 was in May/June 2015, and Wave 5 was in October/November 2015.

dressing, using the toilet, or eating (6.1 percentage points, p=0.01). At the same time, BPCI respondents reported significantly higher rates of improvement from before to after the episode in the degree to which pain limited their regular activities (5.9 percentage points, p=0.10). In Wave 5, BPCI respondents were significantly more likely to report improvement in planning regular tasks (6.5 percentage points, p=0.04) and significantly more likely to report that pain limited their activities (7.1 percentage points, p=0.04). The decline in pain limiting regular activities is inconsistent with the improvement reported by Wave 4 respondents. There were no significant differences in the rate of improvement or decline in the other five functional status measures in either Wave 4 or 5.

There were also inconsistent results with regard to patient experience measures across waves. BPCI respondents reported worse care experiences in two of ten measures relative to comparison respondents in Wave 4. BPCI respondents were 4.6 percentage points (p=0.06) less likely to indicate that their preferences for services after discharge were taken into account than comparison group respondents, and they were 4.2 percentage points (p=0.02) less likely to report that they had a good understanding of how to care for themselves prior to discharge. However, in Wave 5, these two differences were no longer statistically significant. The only significant difference in care experience in Wave 5 was that BPCI respondents were 4.7 percentage points (p=0.07) more likely to agree that they were discharged at the right time than comparison respondents.

Exhibit 151: Select Outcomes from the Beneficiary Survey, Simple Pneumonia and Respiratory Infections, Model 2, May/June 2015 and October/November 2015

Outcomes		Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
			BPCI	Comparison	BPCI	Comparison	Point Estimate	LCI	UCI	p-value
Functional Status	Improvement in planning regular tasks	4	351	319	49.5	50.4	-0.9	-7.3	5.6	0.79
		5	306	340	53. 6	47	6.5	0.3	12.8	0.04
	Improvement in pain limiting regular activities	4	353	308	53.9	48	5.9	-1	12.8	0.1
		5	306	343	44.7	46.6	-1.8	-9.4	5.7	0.63
	Decline in bathing, dressing, using the toilet, or eating	4	350	316	20.2	14.1	6.1	1.4	10.7	0.01
		5	309	345	17.7	21.6	-3.9	-9.3	1.5	0.16
	Decline in pain limiting regular activities	4	353	308	18.6	23.5	-5	-11	1	0.1
		5	306	343	27.4	20.3	7.1	0.2	14	0.04
Patient Care Experience	Patient was discharged at right time	4	356	325	85.5	89.2	-3.7	-9	1.5	0.16
		5	324	349	89.4	84.7	4.7	-0.5	9.8	0.07
	Patient's preferences considered for services after discharge	4	307	265	88.6	93.2	-4.6	-9.5	0.2	0.06
		5	284	304	93.2	90.9	2.3	-2	6.7	0.3
	Prior to discharge patient understood how to care for self	4	302	263	93.3	97.6	-4.2	-7.7	-0.8	0.02
		5	280	304	95.1	95.6	-0.6	-4.2	3.1	0.77

Source: Lewin analysis of beneficiary survey data for episodes that began May, June, October, and November 2015 for BPCI EIs and the matched comparison providers.

Note: This table reports all functional status and care experience measures that have a statistically significant difference between BPCI and the comparison group at the 10% level in at least one wave. The improvement indicator takes a value of 1 if a patient moved to a better functional status level after the episode of care or if the patient recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response. The decline indicator takes a value of 1 if the patient moved to a worse functional status group after the episode of care or if the patient recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey. These indicators are assigned a value of 0 otherwise.

f. Conclusion

During its first eight quarters, the BPCI initiative resulted in few statistically significant changes in Medicare payments, utilization of health care services, and quality of care for the SPRI episodes in Model 2. Anecdotal evidence from BPCI interviewees offered some explanation for the difficulties in reducing payments, such as patients' multiple comorbidities and lack of opportunities to lower device or implant costs. While there was evidence BPCI had a positive impact on one out of 10 functional status outcomes, the results of the beneficiary survey were inconclusive.

13. Spinal Fusion (Non-Cervical)

Medicare beneficiaries received 1,314 non-cervical spinal fusion episodes of care in 29 BPCI-participating acute care hospitals in the first eight quarters of the initiative (from October 2013 through September 2015). BPCI had no statistically significant impact on total Medicare payments, post-acute care (PAC) utilization, or hospital length of stay. However, BPCI did appear to increase emergency department (ED) use rates. Declines in improvements in activities of daily living (ADLs) were also observed among patients discharged to PAC who received care long enough to have two patient assessments.

a. Participants

During the first eight quarters of the BPCI initiative, 29 hospitals (7% of all Model 2 BPCI hospitals) participated in the spinal fusion episode, all of which chose 90-day episodes. Because participants were allowed to join BPCI over an extended period, the spinal fusion results are based on an average of three quarters of experience. Eight of the hospitals began participating in spinal fusion in Q3 2015; 12 (41%) joined in either Q2 or Q3 2015. Seven hospitals stopped participating in spinal fusion within the first eight quarters of the initiative. The BPCI hospitals had 1,314 spinal fusion episodes during the initiative (approximately 1% of all episodes across the 48 BPCI clinical episodes).

Compared to hospitals that did *not* participate in BPCI, hospitals that participated in the spinal fusion episode were larger and more often urban, non-profit facilities, with greater teaching intensity as measured by the resident-to-bed ratios. Hospitals participating in the spinal fusion episode on average had a lower share of Medicare days and were also less likely to be part of a chain. These hospitals also had more spinal fusion episodes and higher 2011 standardized Part A payments for spinal fusion discharges in 2011 than the hospitals not participating in BPCI (Exhibits 152a & 152b).

Matched comparison hospitals were identified for 26 of the 29 BPCI-participating hospitals.[134] The comparison hospitals were more likely to be non-profit than the BPCI-participating hospitals, but were not statistically different across other key characteristics.

[134] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. There were three BPCI hospitals that were not included in the analysis because they had fewer than five relevant discharges in 2011 or 2012.

Exhibits 152a &152b: Characteristics of BPCI-participating Hospitals with Spinal Fusion (Non-Cervical) Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Spinal Fusion BPCI Hospitals (N=29)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)		Matched Spinal Fusion BPCI Hospitals (N=26)		Comparison Hospitals (N=341)+	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	22	76%	1,594	57%*	321	77%	20	77%	290	85%*
	Government	0	0%	542	20%*	32	8%*	0	0%	0	0%
	For-Profit	7	24%	638	23%	66	16%	6	23%	51	15%*
Urban/Rural	Urban	29	100%	1,902	69%*	387	92%*	26	100%	341	100%
Part of Chain	Yes	10	34%	1,469	53%*	218	52%*	10	38%	143	42%

Characteristic	All Spinal Fusion BPCI Hospitals (N=29)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)	Matched Spinal Fusion BPCI Hospitals (N=26)	Comparison Hospitals (N=341)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	338	175*	311	364	348
Medicare Days Percent	39%	42%*	39%	38%	38%
Resident-to-bed ratio	0.15	0.05*	0.12	0.16	0.13
Disproportionate Share Percent	25%	29%*	27%	26%	27%
Number of Spinal Fusion Discharges, 2011	57	16*	47	63	61
Standardized Part A Allowed Payment inpatient stay plus 90-day post-discharge period, Spinal fusion discharges, 2011	**$31,833**	**$30,450***	**$30,963**	**$31,983**	**$32,069**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 EIs, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI during the first year.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All Spinal fusion BPCI" hospitals. "Comparison" hospitals are compared to "Matched Spinal fusion BPCI" hospitals.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 207.

b. Medicare Payments and Utilization

BPCI did not appear to affect Medicare payments. The change in total Medicare payments ($2015) for the anchor hospitalization plus all care delivered during the 90-day post-discharge period was not significantly different between BPCI and comparison hospitals (Exhibit 153). Similarly, none of the changes in payments by type of service differed at a 10% significance level. There were no significant differences between BPCI and matched comparison episodes in inpatient hospital length of stay or PAC use.

Exhibit 153: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day Post-discharge Period, Spinal Fusion (Non-Cervical) Episodes, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=1,229)		Comparison (N=1,223)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total allowed payment amount, inpatient stay plus 90-day post-discharge period	$42,946	$42,618	$42,384	$41,043	$1,013	-$1,999	$4,025	0.51

DiD = difference-in-differences, LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

These estimates are developed using a Difference-in-Differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. B Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison hospitals.

Anecdotal information from telephone interviews and site visits indicated that the variability of patients and their procedures that are included under the spinal fusion clinical episode contributed to the difficulty in reducing episode payments.[135] For example, one interviewee noted that "somebody with a two-stage, two-level posterior spine fusion is different than a six-level spine fusion, but they are all included." Interviewees from physician group practices mentioned that "[spinal fusion] episodes are unpredictable especially when trauma is involved, which makes it challenging to predict care plans and control costs." Others stated that "trauma spine patients can cost over $100,000 and surgeons often don't know the severity of the trauma until they are in surgery." Additional issues associated with reducing costs were also mentioned. "New surgical techniques make it almost impossible to be below the target price, even if they are getting good overall outcomes." "Complex surgeries are riskier, costlier, and have longer recoveries, but are lumped into the same DRG as less complex surgeries. Therefore, one outlier surgery can have a significant adverse impact on performance relative to target price."

c. *Quality of Care*

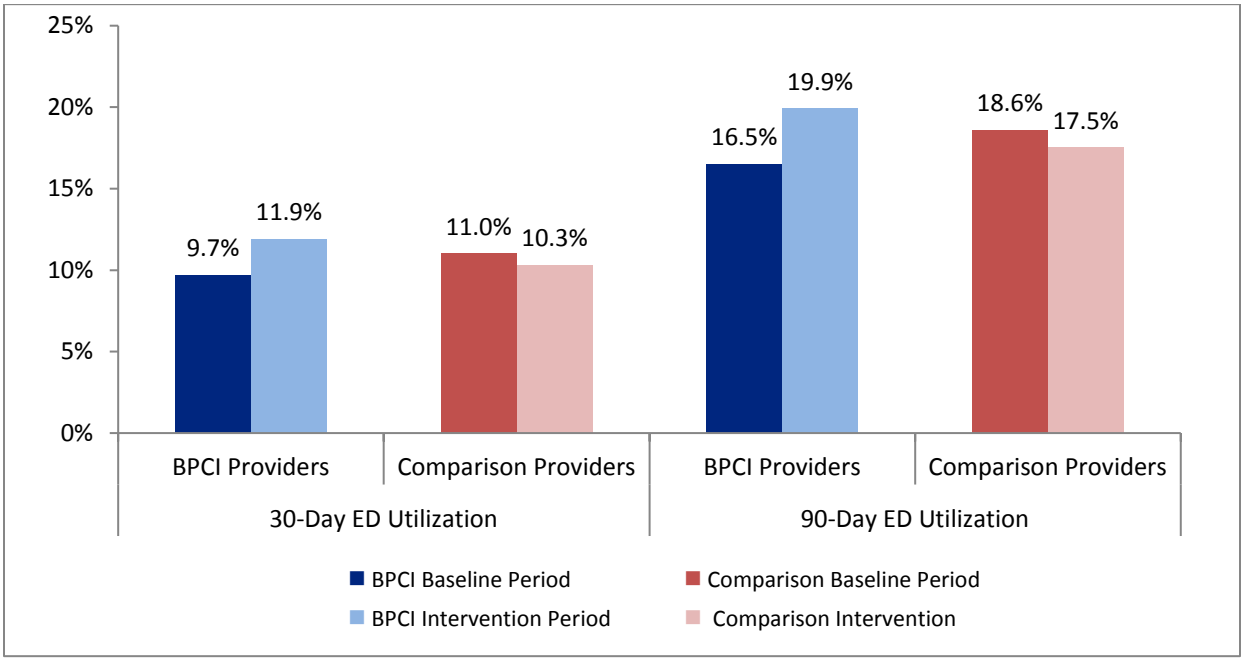
Changes in readmission rates and mortality did not differ between BPCI and comparison patients with spinal fusion surgery, however, ED use increased more for BPCI patients than the comparison patients. The average percent of BPCI patients who visited the ED during the 30 days post discharge increased from 9.7% to 11.9% while it declined from 11.0% to 10.3% among comparison group patients (Exhibit 154). The net effect was an increase of 3 percentage points among BPCI beneficiaries (p=0.04).[136] The relative increase was even larger across the 90

[135] See Methods section for more information on the BPCI participant site visits and interviews.

[136] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

days post discharge (4.5 percentage points, p=0.01). The average percent of BPCI patients who used the ED during the 90 days increased from 16.5% to 19.9%, while it declined among comparison patients from 18.6% to 17.5%.

Exhibit 154: Impact of BPCI on 30-day and 90-day Emergency Department utilization, Spinal Fusion (Non-Cervical) Episodes, Model 2, Q4 2011-Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

For patients who were discharged to PAC after their spinal fusion surgery and had two patient assessments, we examined the changes in function between the two PAC assessment instruments (Exhibit 155). BPCI patients who were discharged to SNF were less likely to have improved functional status across all three functional measures, although only two were statistically significant. BPCI patients discharged to SNF experienced a significant decline in self-care function relative to the comparison group from the baseline to intervention period (13.4 percentage points, p=0.05). During the baseline, 60.1% of BPCI beneficiaries experienced an improvement in self-care functioning, compared with 50.7% in the intervention period. By contrast, the percentage increased for the comparison group from 61.8% to 65.8%. The proportion of BPCI patients discharged to SNF that experienced an improvement in overall function also declined relative to the comparison group from the baseline to the intervention period (8.6 percentage points, p=0.10). During the baseline, 72.9% of BPCI beneficiaries experienced an improvement in overall function, compared with 68.6% in the intervention period, while the share experiencing an improvement increased for the comparison group from 77.5% to 81.9%.

Among BPCI patients discharged to IRF, both measures of functional status declined relative to the comparison group.[137] During the baseline, the average increase in mobility score among BPCI patients was 9.7, which declined to 9.3 in the intervention period. For the comparison group, the average change in the mobility score increased from 9.3 to 10.4 among comparison group patients, for a relative difference of -1.6 ($p<0.01$). The average change in self-care score among BPCI patients also declined from baseline to intervention period, while the average change increased among comparison group providers, for a difference of -2.4 ($p=0.05$).

[137] Changes in mobility and self-care for IRF patients are measured as changes in a mobility related index score and a self-care related index score between IRF admission and IRF discharge. The mobility index ranges from a score of 4 (total assistance) to 28 (fully independent) while the self-care index ranges from a score of 6 (total assistance) to 42 (fully independent).

Exhibit 155: Impact of BPCI on Select Functional Status Outcomes, Spinal Fusion (Non-cervical) Patients Discharged to SNF or IRF, Model 2, Q4 2013 – Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Share SNF patients with improved overall function[1]	195	190	72.9	68.6%	77.5%	81.9%	-8.6	-18.9	1.6	0.10
Share SNF patients with improved self-care function[1]	196	190	60.1%	50.7%	61.8%	65.8%	-13.4	-26.6	-0.2	0.05
Share SNF patients with improved mobility function[1]	195	190	67.1%	66.7%	74.0%	75.5%	-2.1	-12.1	8.0	0.69
Average change in mobility score among IRF patients	140	163	9.7	9.3	9.3	10.4	-1.6	-3.1	0.0	0.05
Average change in self-care score among IRF patients	140	163	11.7	10.9	11.5	13.0	-2.4	-4.0	-0.9	<0.01

DiD=difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; SNF = Skilled Nursing Facility; IRF= Inpatient Rehabilitation Facility.

[1] Improvement means improvement or staying in the best possible status

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Source: Lewin analysis of Medicare assessment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

d. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. There were no statistically significant changes in patient characteristics that suggest BPCI-participating hospitals saw a healthier mix of spinal fusion patients relative to the comparison group (Exhibits 156a & 156b). The average number of Hierarchical Condition Categories (HCCs—a risk measure used in Medicare's managed care program) for BPCI patients increased from 1.2 to 1.3 from the baseline to the intervention period compared with a stable 1.1 for the comparison group (p=0.03). SNF utilization in the six months prior to their index hospitalization also increased for the BPCI patients and declined for the comparison patients (p=0.08). Both of these measures may indicate worse baseline health status in the BPCI patients, which could contribute to higher costs or reductions in functional gains.

Exhibits 156a & 156b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Spinal Fusion (Non-Cervical), Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	320	22%	235	19%	261	18%	219	18%	-2.7	-6.9	1.5	0.20
	65-79	954	65%	848	68%	1,007	69%	877	71%	1.3	-3.7	6.3	0.62
	80+	193	13%	159	13%	199	14%	146	12%	1.5	-2.1	5.0	0.43
Gender	Female	879	60%	727	59%	900	61%	748	60%	-0.3	-5.5	5.0	0.92
Medicaid and Disability	% Eligible for Medicaid	237	16%	160	13%	223	15%	163	13%	-1.2	-5.0	2.6	0.53
	% Disability, no ESRD	375	26%	260	21%	318	22%	249	20%	-3.0	-7.4	1.4	0.19
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	233	16%	178	14%	188	13%	142	11%	-0.2	-3.8	3.5	0.93
	Emergency Room Admission	370	25%	291	23%	354	24%	287	23%	-0.8	-5.3	3.8	0.74
	Home Health	140	10%	116	9%	108	7%	93	7%	-0.3	-3.3	2.6	0.83
	Inpatient Rehabilitation Facility	23	2%	17	1%	15	1%	13	1%	-0.2	-1.4	1.0	0.71
	Skilled Nursing Facility	42	3%	44	4%	40	3%	23	2%	1.6	-0.2	3.3	0.08
	Psychiatric Hospital	5	0%	4	0%	6	0%	3	0%	0.1	-0.5	0.8	0.64
	Long-term Care Hospital	1	0%	0	0%	1	0%	0	0%	0	-0.2	0.2	1.00
	Institutional Nursing Facility*	44	3%	39	3%	44	3%	25	2%	1.1	-0.6	2.9	0.21
	No Institutional Care	1,226	84%	1059	85%	1,264	86%	1089	88%	0.2	-3.6	3.9	0.93
	No Post-acute Care	923	63%	818	66%	966	66%	841	68%	1.1	-4.0	6.2	0.68

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	1.20	1.42	1.32	1.44	1.14	1.36	1.10	1.33	0.16	0.01	0.31	0.03

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 for BPCI and comparison providers.

"Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set assessment data. All other measures are based on Medicare claims.

e. Conclusion

The BPCI initiative had no observable effect on Medicare payments or utilization for spinal fusion episodes. Interviews with participating providers indicated that this might be due to the heterogeneity of patients within this clinical episode, which reduces ability to plan or implement care redesign, and the likelihood that a single high cost patient could raise average episode costs. In particular, we measured a relative increase in ED use and a relative decline in the share of patients discharged to SNF or IRF with improvements in ADL measures.

14. Stroke

There were 4,227 episodes of care for stroke in 63 BPCI-participating acute care hospitals for Medicare beneficiaries in the first eight quarters of the BPCI initiative (from October 2013 through September 2015). BPCI exhibited few differences in Medicare payment and quality of care changes relative to a matched comparison group over the baseline and intervention periods.

- We observed no statistically significant change in total Medicare payments for BPCI stroke episodes relative to comparison episodes. There were also no statistically significant differences in payment changes from the baseline to intervention period 30 days prior to the bundle or 30 days after the bundle.
- There were limited effects on Medicare payments by type of service within the 90-day post-discharge period.
- There was a relative increase in the rate of emergency department visits in the 90-day post-discharge period for BPCI episodes; however the 30-day rate was not statistically significant. There were also no differences in the rate of unplanned readmissions or mortality.
- Seven of eight measures using assessment data pointed to a relative decline in functional improvement for BPCI patients who received post-acute care (PAC) in home health agencies (HHAs) and skilled nursing facilities (SNFs) long enough to have two patient assessments, although only two of the changes from the baseline to intervention period were statistically significant. In contrast, both measures assessing functional limitations for patients treated by inpatient rehabilitation facilities (IRFs) pointed to relative improvements for BPCI patients.
- There were limited BPCI impacts on self-reported measures of functional limitations, as well as health and patient experience, when comparing survey data from BPCI and matched comparison cohorts.

a. Participants

During the first eight quarters of the BPCI initiative, 63 hospitals (15% of the all hospitals participating in Model 2) participated in the stroke episode. A 90-day episode length was selected by all but one hospital, which chose a 60-day episode. These providers initiated 4,227 episodes during their participation, about 3% of all episodes under BPCI Model 2 for acute care hospitals. The average length of participation among the hospitals participating in stroke was approximately three quarters and 67% of hospitals joined in quarters 2 or 3 of 2015. Eight out of the 63 hospitals stopped participating in the stroke episode by the end of the eighth quarter.

Compared to hospitals *not* participating in BPCI, hospitals that participated in the stroke episode had more beds on average, and were more likely to be non-profit and located in urban areas (Exhibits 157a & 157b). They also had higher resident-to-bed ratios and disproportionate share percentages, but fewer inpatients days associated with Medicare patients as a proportion of total inpatient days on average. Prior to joining BPCI, participating hospitals had more stroke discharges for Medicare patients. Further, participating hospitals had higher average 2011 standardized Part A payments for services during the stroke anchor hospitalization plus 90 days post-discharge than non-participants ($26,500 compared to $24,107).

In general, hospitals participating in stroke were more similar to all BPCI hospitals, regardless of whether they participated in stroke, compared to those not participating in BPCI at all. However, there were some notable differences. Those participating in stroke were more likely to be in urban areas. They also had more beds and a higher disproportionate share percentage on average, but a lower proportion of Medicare days. Stroke and all BPCI participants had a similar number of stroke discharges in 2011 on average, but hospitals participating in the episode had higher average Medicare Part A payments for services during the anchor hospitalization plus 90-day post-discharge period ($26,500 compared to $24,626).

Matched comparison hospitals were identified for 57 of the 63 BPCI-participating hospitals in the stroke episode sample.[138] The characteristics of the BPCI and the matched comparison hospitals were similar.

[138] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics, as well as baseline outcomes. There were five participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter before they stopped their participation in the stroke episode or had fewer than five relevant discharges in 2011 and 2012. There was one other participant for which there was not a match within the caliper.

Exhibits 157a &157b: Characteristics of BPCI-participating Hospitals with Stroke Episodes, compared with Non-participating Hospitals, All BPCI-participating Hospital EIs, and Comparison Hospitals, Model 2, Q4 2013 - Q3 2015

Characteristic		All Stroke BPCI Hospitals (N=63)		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=419)**		Matched Stroke BPCI Hospitals (N=57)		Comparison Hospitals (N=732)+	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	43	68%	1,594	57%*	321	77%	41	72%	523	71%
	Government	6	10%	542	20%*	32	8%	5	9%	67	9%
	For-Profit	14	22%	638	23%	66	16%	11	19%	142	20%
Urban/Rural	Urban	63	100%	1,902	69%*	387	92%*	57	100%	732	100%
Part of Chain	Yes	34	54%	1,469	53%	218	52%	32	56%	424	58%

Characteristic	All Stroke BPCI Hospitals (N=63)	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=419)**	Matched Stroke BPCI Hospitals (N=57)	Comparison Hospitals (N=732)+
	Mean	Mean	Mean	Mean	Mean
Bed Count	369	175*	311*	357	380
Medicare Days Percent	37%	42%*	39%*	37%	36%
Resident-to-bed ratio	0.12	0.05*	0.12	0.13	0.13
Disproportionate Share Percent	33%	29%*	27%*	32%	34%
Number of Stroke Discharges, 2011	119	54*	109	116	123
Standardized Part A Allowed Payment inpatient stay plus 90-day PDP, Stroke episodes, 2011	**$26,500**	**$24,107***	**$24,626***	**$26,352**	**$26,518**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 2 hospitals in Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 4 hospitals participating in BPCI.

*Indicates differences from BPCI hospitals greater than absolute value of 0.2 standard difference when comparing all stroke BPCI hospitals to non-participating hospitals and all BPCI hospitals. The matched comparison hospitals are compared to the matched stroke BPCI hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+The values for the matched comparison providers are weighted to take into account that some are matched to multiple participants. The unique number of matched providers is 415.

b. Change in Patient Mix

We monitor changes in patient characteristics because a change in patient mix could affect a participant's ability to reduce episode payments below their target. While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine changes in patient mix. There was no consistent pattern of changes in patient characteristics from the baseline to the intervention period for stroke cases treated by BPCI hospitals relative to those treated by comparison hospitals (Exhibits 158a & 158b).There were no statistically significant differences in the proportions of patients by age group, gender, Medicaid eligibility, or disability status. Moreover, there were no differences in

the change in the number and intensity of patients' comorbidities (represented by Hierarchical Condition Category (HCC) scores).

With respect to prior health care utilization, there were statistically significant relative decreases in the percentage of BPCI participants' stroke patients who used an emergency room (p=0.06) or an inpatient rehabilitation facility (p=0.03) in the six months prior to their index hospitalization. These might indicate that the BPCI participants saw a healthier mix of patients. However, there was a statistically significant relative increase in the percentage of BPCI participants' stroke patients that used an inpatient acute care hospital (p=0.09), and a statistically significant relative decrease in the percentage of BPCI participants' patients that had no institutional care (p=0.06) in the six months prior to their index hospitalization. There were no statistically significant changes in the percentage of patients using other health care services.

Exhibits 158a & 158b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for Stroke, Model 2, Q4 2013 - Q3 2015

Characteristic		BPCI Stroke Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	433	8%	274	7%	475	9%	319	8%	-0.4	-2.0	1.2	0.60
	65-79	2,223	40%	1,598	42%	2,275	41%	1661	44%	-0.7	-3.6	2.2	0.62
	80+	2,891	52%	1,930	51%	2,797	50%	1822	48%	1.1	-1.8	4.1	0.44
Gender	Female	3,190	58%	2,114	56%	3,152	57%	2055	54%	0.9	-2.0	3.8	0.56
Medicaid and Disability	% Eligible for Medicaid	1,206	22%	730	19%	1,189	21%	745	20%	-0.7	-3.1	1.7	0.56
	% Disability, no ESRD	505	9%	301	8%	550	10%	342	9%	-0.3	-1.9	1.4	0.75
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	1,330	24%	930	24%	1,189	21%	754	20%	2.1	-0.3	4.5	0.09
	Emergency Room Admission	1,458	26%	1,052	28%	1,450	26%	1141	30%	-2.5	-5.1	0.1	0.06
	Home Health	1,036	19%	674	18%	1,004	18%	684	18%	-0.8	-3.1	1.4	0.46
	Inpatient Rehabilitation Facility	164	3%	99	3%	124	2%	111	3%	-1.0	-2.0	-0.1	0.03
	Skilled Nursing Facility	512	9%	355	9%	415	7%	282	7%	0.2	-1.4	1.8	0.84
	Psychiatric Hospital	35	1%	24	1%	36	1%	17	0%	0.2	-0.2	0.7	0.38
	Long-term Care Hospital	23	0%	11	0%	23	0%	7	0%	0.1	-0.2	0.4	0.55
	Institutional Nursing Facility*	702	13%	480	13%	591	11%	391	10%	0.3	-1.5	2.2	0.72
	No Institutional Care	4,124	74%	2,813	74%	4,243	76%	2985	79%	-2.4	-4.9	0.1	0.06
	No Post-acute Care	3,020	54%	2,016	53%	3,066	55%	2044	54%	0.1	-2.8	3.0	0.95

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.52	1.99	2.55	2.07	2.39	1.89	2.44	1.97	-0.03	-0.14	0.09	0.65

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

"Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

*Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

c. Medicare Payments and Utilization

During the baseline period (Q4 2011 through Q3 2012), Medicare payments for services provided during the stroke index hospitalization plus 90 days post-discharge was $33,589 ($2015) on average for episodes initiated by BPCI hospitals (Exhibit 159).[139] Payments decreased by 6.6% from the baseline to intervention period for BPCI hospitals (dropping to $31,378). A similar change occurred for comparison hospitals during this time (a decrease of 6.2% from $33,282 to $31,212). Thus, BPCI did not have a statistically significant impact on overall Medicare payments during the 90-day episode timeframe.

Exhibit 159: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization and 90-day PDP, Stroke, Model 2, Q4 2011-Q3 2015

Measure	BPCI (N=3,743)		Comparison (N=3,729)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total allowed payment amount, IP and 90-day PDP	$33,589	$31,378	$33,282	$31,212	-$142	-$1,565	$1,282	0.85

DiD = difference-in-differences, LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level; IP = anchor inpatient stay; PDP = Post-discharge period.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI and comparison providers.

There were no statistically significant changes in Medicare payments by type of service within the 90-day post-discharge period, with the exception of other non-institutional services.[140] Medicare payments for BPCI patients for other non-institutional services decreased 13.1% from $449 in the baseline period to $391 in the intervention period, while the comparison group decreased by 3.8% from $425 to $409. The result is that BPCI was associated with a relative decrease of $43 (p=0.08).[141] Non-institutional services comprise a very small proportion (less than 2.0%) of total stroke episode payments. There were no statistically significant changes in utilization for acute care hospital length of stay or PAC utilization, during the 90-day post-discharge period.

[139] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

[140] The other non-institutional category of services includes services such as ambulance, chiropractic, vision, hearing, and speech services.

[141] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Pre-bundle and post-bundle changes in payments from the baseline to intervention period for BPCI and comparison providers were also similar, providing no indication that hospitals shifted services outside of the bundle to reduce episode payments.

d. Quality of Care

To explore quality of care impacts of BPCI on patients with stroke, Medicare claims and enrollment data, assessment data and survey data were used. In general, there were limited and inconsistent results.

Changes from the baseline to intervention period in unplanned readmission rates, emergency department visits and mortality rates were explored using claims and enrollment data. From the baseline to the intervention period, the change in unplanned readmission rates and mortality rates following the anchor hospitalization for stroke episodes was not statistically different for episodes initiated by BPCI-participating hospitals and matched comparison hospitals.

The rate of emergency department visits in the 90-day post-discharge period for BPCI episodes increased 1.3 percentage points (from 19.9% in the baseline period to 21.1% in the intervention period). There was a larger increase for matched comparison episodes of 3.6 percentage points (from 18.8% to 22.4%). The result is that BPCI was associated with a statistically significant relative decrease of 2.3 percentage points (p=0.07) in emergency department visits. It is not clear that this is a sign of improved care as the rates were increasing for both the BPCI and comparison cohorts, and there was no significant difference in the rate of emergency department visits within the 30-day post-discharge period.

We used ten measures to assess the functional status of patients treated by PAC providers who remained in the PAC setting long enough to receive two assessments.[142] Each of the three SNF measures were associated with relative declines for BPCI patients, although only two were statistically significant; ability to self-care and overall functionality. The share of BPCI patients discharged to a SNF who experienced an improvement in self-care functioning decreased from 33.2% to 29.4%, while the comparison group share increased from 33.0% to 37.0%, resulting in a relative decrease of 7.8 percentage points (p=0.03) in self-care functioning attributable to BPCI.[143] The proportion of BPCI patients that experienced improvement in overall functioning decreased from 56.2% to 52.7% for BPCI patients, while increasing from 56.6% to 58.8% for comparison group patients (-5.7 percentage points, p=0.08). Four out of five measures for patients treated by HHAs also exhibited relative decreases (no statistical significance). In contrast, both measures involving IRF patients exhibited relative increases for BPCI patients, although only one was statistically significant. BPCI was associated with a relative increase in mobility functioning for patients who were treated by IRFs. The average self-care score among BPCI IRF patients improved from 6.8 to 7.4, while the average self-care score among comparison patients remained around 6.9 from the baseline to the intervention period (0.7, p=0.04).[144] These results were based

[142] There were five measures for HHA patients relating to improved bathing, ambulation, upper body dressing, lower body dressing, and bed transferring. There were three measures for SNF patients relating to improved overall function, self-care and mobility. There were two measures for IRF patients relating to improved mobility and self-care.

[143] "Improvement" means improvement or staying in the best possible status.

[144] Change in self-care for IRF patients was measured as change in a self-care related index score between IRF admission and IRF discharge. The self-care index ranges from a score of 6 (total assistance) to 42 (fully independent).

on the approximately three-quarters of patients who received care long enough to have two assessments and may not be representative of all patients discharged to PAC.

Beneficiaries treated by BPCI and comparison hospitals were also surveyed to provide additional insights regarding impacts on functional status, as well as self-reported health and care experience indictors. Beneficiaries with a stroke episode initiated in a Model 2 participating hospitals during May, June, October, and November 2015 and a matched comparison group were surveyed.[145] There were no differences between BPCI and comparison respondents with a stroke episode in self-reported rates of functional change (improvement or decline) from before to after their hospital episode that were significant at the 10% level.[146] There were also no statistically significant differences in self-reported overall physical or mental health. Among measures of care experience, BPCI respondents with stroke were 7.0 percentage points less likely than comparison respondents to agree that they were discharged at the right time ($p<0.01$) (Exhibit 160). However, BPCI and comparison respondents reported similar levels of satisfaction with the overall recovery.

Exhibit 160: Select Beneficiary Survey Outcomes, Model 2 ACH, Stroke Episodes, May/June 2015 and October/November 2015

Measure	Number of Episodes		Risk adjusted Rates		Estimated Difference			
	BPCI	Comparison	BPCI	Comparison	Point Estimate	LCI	UCI	p-value
Patient was discharged at right time	383	392	84.16	91.17	-7.0	-12.0	-2.0	0.01

Source: Lewin analysis of beneficiary survey data for episodes that began May, June, October, and November 2015 for BPCI EIs and the matched comparison providers.

This table reports all functional status and care experience measures that have a statistically significant difference between the intervention and comparison groups at the 10% or 5% level. LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level.

e. Conclusion

BPCI episodes exhibited few statistically significant differences in Medicare payment and quality of care changes relative to a matched comparison group over the baseline to intervention period. We observed no statistically significant impact of BPCI on total Medicare payments for stroke episodes. There were also no differences in payments 30 days prior to the bundle or 30 days after the bundle. Changes from the baseline to intervention period for Medicare payments by type of service within the 90-day episode time frame were also largely similar.

The analysis revealed few significant impacts on quality of care due to BPCI. There was a relative decrease in the rate of emergency department visits in the 90-day post-discharge period

[145] Episodes of stroke were sampled twice in each survey wave over two consecutive months. Each of the two samples was drawn from hospital discharges that occurred within a single month: Wave 4, May-June 2015; Wave 5, October-November 2015.

[146] For each functional status measure, we created binary indicators for both improvement and decline to measure change from before to after the treatment episode. The improvement indicator takes a value of 1 if a respondent moved to a better functional status group after the hospitalization or if the respondent recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response. The indicator is assigned a value of 0 otherwise. The decline indicator takes a value of 1 if the respondent moved to a worse functional status group after hospitalization or if the respondent recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey.

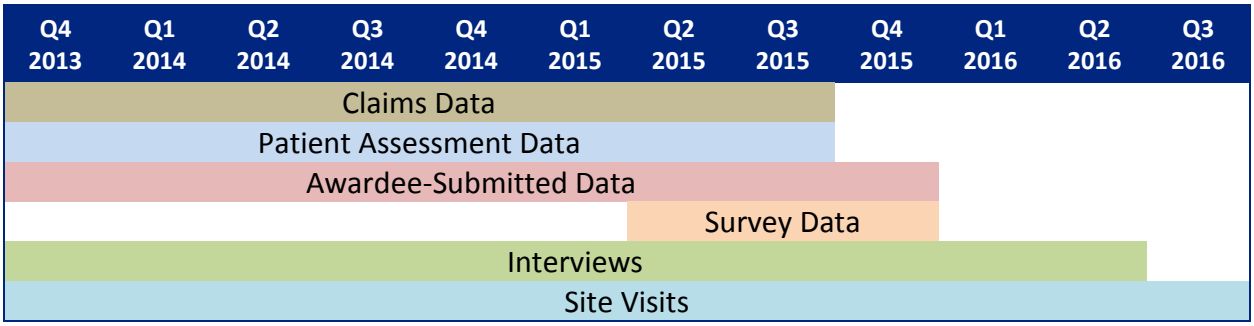
for BPCI episodes; however the 30-day rate was not statistically significant. There were no differences in the rate of unplanned readmissions or mortality.

While only three of the assessment measures on functional limitations were statistically significant, there were relative decreases for BPCI patients in 7 out of the 10 measures, including all of the measures for patients in SNFs. This may suggest that quality of care decreased for BPCI patients who remained in a SNF long enough to receive two assessments relative to the change for comparison group. These results were based on the approximately three-quarters of patients who received care long enough to have two assessments and may not be representative of all patients discharged to PAC. Patient survey data indicated limited impacts of BPCI on self-reported measures of functional limitations, as well as health status and patient experience. As BPCI participants gain more experience, impacts along these domains may become clearer.

IV. Model 3 Results and Discussion

This section presents information about the experience of Model 3 BPCI participants and their episodes of care. Exhibit 161 presents the Phase 2 BPCI time period reflected in each of the data sources used throughout the Model 3 results section. The quantitative outcomes are risk adjusted as described in Section II.D.2.

Exhibit 161: BPCI Quantitative and Qualitative Data Sources Used for Model 3 Results



Note: The risk-bearing phase (Phase 2) of BPCI began Q4 2013. Awardee-submitted data for gainsharing analysis covers Q4 2013 through Q1 2016.

A. Characteristics of the initiative and participants

This section summarizes the characteristics of the BPCI participants during the first two years of the initiative. Where relevant, we summarized the qualitative results from the last BPCI evaluation annual report in a call-out box in the beginning of a section. The narrative that follows the call-out box provides additional insights gathered through the last year's qualitative data collection and analysis.

1. Key takeaways on Model 3 characteristics of the initiative and participants

- 136 Awardees that represented 873 skilled nursing facility (SNF) episode initiators (EIs), 144 physician group practice (PGP) EIs, 116 home health agency (HHA) EIs, 9 inpatient rehabilitation facility (IRF) EIs, and one long-term care hospital (LTCH) EI joined the risk-bearing phase in Model 3 of BPCI. These participants initiated over 35,000 episodes of care during the first two years of the initiative.
- BPCI-participating SNF and HHA EIs were larger and likely had more resources (as measured by for-profit status and chain membership) than non-participating providers. They also had higher 2011 standardized Part A payments prior to joining BPCI for the episodes they chose, which may indicate that they had greater opportunities for reducing episode payments under BPCI. At this time, we have not assessed characteristics of the PGP EIs relative to non-participating providers.
- BPCI-participating SNF and HHA EIs were located in more heavily populated, urban and competitive markets than non-participants.
- Model 3 BPCI participants did not make extensive use of gainsharing and program rule waivers. Fourteen (25%) Awardees entered into gainsharing agreements from Q4 2013 through Q2 2015. Awardees distributed approximately $4.6 million in net payment reconciliation amounts (NPRA) to their gainsharing partners between Q4 2013 and

Q1 2016. No more than 5% of EIs used the telehealth or post-discharge home visits waivers or provided beneficiary incentives.

- Interviewees noted the importance of their relationships with hospitals, particularly in markets where post-acute care (PAC) providers competed over admissions. SNF EIs noted that because they had reduced length of stay under BPCI, patient turnover had increased, so they needed to rely on relationships with hospitals to boost referrals. Interviewees noted, however, that hospitals participating in BPCI or other payment reform initiatives (e.g., Accountable Care Organizations (ACOs), Comprehensive Care for Joint Replacement (CJR) Model) were focused on discharging patients directly home.
- Interviewees noted that relationships with Conveners or external contractors are important for data management and analysis.
- Interviewees reported providing early patient education and consistent messaging from facility staff to manage patient expectations after discharge and prevent readmissions. Additionally, assessing patient risks and social support was used to determine the most appropriate level of care and follow-up required after discharge.
- To reduce readmissions, some interviewees reported contacting HHAs in the post-discharge period and coordinating care with primary care providers (PCPs) (e.g., setting up follow-up appointments, communicating patient status updates directly, sharing discharge plans). If necessary, PAC providers preferred to readmit patients back into their care to avoid an emergency department (ED) readmission.
- Interviewees noted that obtaining the Medicare Severity Diagnosis Related Group (MS-DRG) upon discharge from the hospital was often difficult and took time, making it challenging to identify BPCI patients.
- During the first two years of the initiative, 44 SNF EIs and 15 HHA EIs withdrew from BPCI.

2. Participants

Interview Findings Reported Previously

- Interviewees joined BPCI for the opportunity to learn and gain experience with bundled payment models.

a. Entry decisions

Similar to Model 2 interviewees and information gained from previous years, Model 3 interviewees reported that they joined BPCI for the opportunity to learn. Interviewees sought experience using financial incentives to drive change, coordinating care beyond discharge, improving care transitions, and managing financial risk. They also sought opportunities to learn about bundled payments, which interviewees described as the future of payment reform. Being considered a leader in health care and payment reform was another reason interviewees joined the initiative, which often improved their ability to negotiate with private payers. One interviewee described a goal of being "indispensable to commercial payers" due to their experience in BPCI. Two interviewees had established contracts similar to BPCI with commercial payers.

According to PGP interviewees, the selection of Model 3 was typically made at a corporate level rather than by specific EIs. Respondents stated that they considered clinical opportunity, financial opportunity, and risk profile when selecting a Model. PGPs that selected Model 3 instead of Model 2 stated that they did so because of perceived challenges with influencing care at hospitals. Similarly, some PGP respondents stated that because their physicians work with many different hospitals, it would have been difficult to partner with all of them effectively.

b. Participant characteristics

This section describes participants that joined BPCI through Q3 2015. There were 136 Awardees with 1,143 EIs in Model 3 during the first two years of the initiative, including 873 SNFs, 116 HHAs, 9 IRFs, 1 LTCH, and 144 PGPs.

Exhibits 162a & 162b compare the BPCI-participating SNF EIs during the first two years of the initiative to non-participating SNFs. A higher proportion of BPCI-participating SNF EIs were for-profit organizations (86%) compared with non-participating SNFs (70%) and were likelier to be located in urban areas (84% vs. 70%). Participating SNFs were more likely to be part of a chain (52% vs. 22%), had higher average bed counts (122 vs. 112), and averaged more admissions for BPCI episode MS-DRGs during 2011 (136 vs. 94) than non-participating SNFs.

As shown in Exhibit 162c, in 2011, BPCI-participating SNFs had higher 2011 standardized Part A payments during the 90 days following SNF admission for patients admitted with BPCI MS-DRGs than non-participating SNFs. The difference varied by clinical episode, from 7% higher for stroke admissions to 1% higher for major joint replacement of the lower extremity (MJRLE) admissions.

Exhibit 162a & 162b: Baseline Characteristics of BPCI-participating SNF EIs and Non-participating SNFs, Model 3, Q4 2013 – Q3 2015

Characteristic		BPCI SNF EIs (N=864)		Non-participating SNFs (N=13,302)	
		N	%	N	%
Ownership	For Profit	740	86%	9,374	70%
	Government	3	0%	617	5%
	Non-Profit	121	14%	3,311	25%
Urban/Rural	Rural	140	16%	3,938	30%
	Urban	724	84%	9,364	70%
IRF in CBSA	Yes	488	56%	7,238	54%
Hospital-Based	Yes	7	1%	586	4%
Part of Chain	Yes	216	52%	2,946	22%

Characteristic	BPCI SNF EIs (N=864)	Non-participating SNFs (N=13,302)
	Mean	Mean
Bed Count	122	112
Number of Admissions for BPCI Episode MS-DRGs, 2011	136	94
SNF Market Share	6%	6%
Nursing Home Overall Score*	3.45	3.32

Source: Lewin analysis of 2013 Provider of Service (POS) and 2011 Medicare claims. BPCI participating SNFs are defined as Episode Initiators, Q4 2013 – Q3 2015. Non-Participating SNFs are all other SNFs not participating in any BPCI initiative that reported values for all measures listed in the table above. Please note that BPCI-participating SNFs that received Medicare certification after 2011 are not included in this table.

* Number of points out of 5 in overall rating and in three areas: Quality, Survey/Health Inspections, and Staffing. The closer to 5 the better the quality, inspections, and staffing.

Exhibit 162c: 2011 Standardized Allowed Part A Payments for the 90 days following SNF Admission, for BPCI-participating SNF EIs and Non-participating SNFs, by Clinical Episode[1], Model 3, Q4 2013 – Q3 2015

Clinical Episode	BPCI SNF EIs			Non-participating SNFs		
	Number of SNF EIs	Number of Admissions	Mean Payment Across SNFs	Number of SNFs	Number of Admissions	Mean Payment Across SNFs
Major joint replacement of the lower extremity	202	5,177	$19,449	5,694	120,215	$19,284
Pneumonia	192	2,752	$24,788	8,058	97,425	$24,183
Chronic obstructive pulmonary disease	112	1,231	$26,205	4,046	37,652	$24,642
Congestive heart failure	150	2,113	$25,023	6,021	66,055	$24,832
Sepsis	159	2,490	$27,169	7,163	87,553	$25,841
Urinary tract infection	131	1,713	$25,898	6,023	61,535	$25,498
Other respiratory	90	810	$27,294	3,128	24,806	$26,889
Stroke	92	943	$30,737	3,823	33,927	$28,603
Medical non-infectious	109	1,479	$28,840	4,313	46,178	$26,995
Hip and femur	110	1,349	$31,001	4,876	54,218	$30,198
Renal failure	98	1,161	$26,450	4,648	41,909	$25,796

Source: Lewin analysis of 2013 Provider of Service (POS) and 2011 Medicare claims. Standardized allowed payments are based on 2011 claims and trended to 2015 dollars. BPCI-participating SNFs are defined as EIs, Q4 2013 – Q3 2015. Non-Participating SNFs are all other SNFs not participating in any BPCI initiative, have reported values for all measures listed in Exhibits 162a & 162b, and have fewer than 5 discharges in 2011 in the episode of relevance. Please note that BPCI-participating SNFs that received Medicare certification after 2011 are not included in this table.

[1] The clinical episodes included in this table are those that had enough sample size to include in the risk adjusted difference-in-differences (DiD) analysis presented in section IV.B.1.

Exhibits 163a & 163b compare the Model 3 BPCI-participating HHAs to all non-participating HHAs. A higher proportion of participating HHAs were part of a chain (73% vs 32%) and for-profit (81% vs 76%). BPCI-participating HHAs had more employed nurses on average than did non-participating HHAs (29 vs. 9), although the BPCI average is driven by one large HHA. BPCI-participating HHAs also had a greater number of admissions for BPCI episode MS-DRGs during 2011 (374 vs. 101).

As shown in Exhibit 163c, BPCI-participating HHAs had higher 2011 standardized Part A payments during the 90 days from the start of receiving home health for patients admitted with BPCI MS-DRGs in 2011 relative to non-participating HHAs. The difference in standardized payments varied by clinical episode; standardized payments were 9% higher for pneumonia admissions in BPCI-participating HHAs and virtually the same for CHF admissions.

Exhibits 163a & 163b: Baseline Characteristics of BPCI-participating HHA EIs and Non-participating HHAs, Model 3, Q4 2013 – Q3 2015

Characteristic		BPCI HHA EIs (N=116)		Non-participating HHAs (N=9,769)	
		N	%	N	%
Ownership	For Profit	94	81%	7,458	76%
	Government	0	0%	612	6%
	Non-Profit	22	19%	1,699	17%
Urban/Rural	Rural	25	22%	1,886	19%
	Urban	91	78%	7,883	81%
Part of Chain	Yes	85	73%	3,110	32%

Characteristic	BPCI HHA EIs (N=116)	Non-participating HHAs (N=9,769)
	Mean	Mean
Number of Employed Nurses in HHA	29	9
Number of Admissions for BPCI Episode MS-DRGs, 2011	374	101

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating HHAs are defined as Episode Initiators, Q4 2013 – Q3 2015. Non-Participating HHAs are all other HHAs not participating in any BPCI initiative that reported values for all measures listed in the table above. Please note that BPCI-participating HHAs that received Medicare certification after 2011 are not included in this table.

Exhibit 163c: 2011 Standardized Allowed Part A Payments for the 90 days from the start of receiving home health, for BPCI-participating HHA EIs and Non-participating HHAs, by Clinical Episode[1], Model 3, Q4 2013 – Q3 2015

Clinical Episode	BPCI HHA EIs			Non-participating HHAs		
	Number of HHA EIs	Number of Admissions	Mean Payment Across HHAs	Number of HHAs	Number of Admissions	Mean Payment Across HHAs
Major joint replacement of the lower extremity	25	1,328	$4,604	3,263	115,529	$4,235
Simple pneumonia and respiratory infections	34	868	$10,610	3,325	62,154	$9,779
Congestive heart failure	42	2,570	$11,594	3,561	76,022	$11,578

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. Standardized allowed payments are based on 2011 claims and trended to 2015 dollars. BPCI participating HHAs are defined as Episode Initiators, Q4 2013 – Q3 2015. Non-Participating HHAs include all other HHAs not participating in any BPCI initiative that have reported values for all measures listed in Exhibits 163a & 163b and have fewer than 5 discharges in 2011 in the episode of relevance. Please note that BPCI-participating HHAs that received Medicare certification after 2011 are not included in this table.

[1] The clinical episodes included in this table are those that had enough sample size to include in the risk adjusted Difference-in-differences analysis presented in section IV.B.1.

3. Participant Readiness

Exhibit 164 summarizes the engagement of Model 3 SNFs, HHAs, and PGPs in care redesign initiatives and payment incentives prior to joining BPCI.[1] Relatively few Model 3 EIs reported experience with care redesign initiatives, and even fewer had experience with innovative payment models.

Exhibit 164: Care Redesign and Payment Incentives Experience, Model 3, Prior to BPCI Participation*

		Model 3 SNF EIs N=837		Model 3 HHA EIs N=114		Model 3 PGP EIs N=144	
		N	%	N	%	N	%
Prior experience with care redesign initiatives	Redesign of Care Pathways	128	15.3%	28	24.6%	58	40.3%
	Enhancements in Care Delivery	81	9.7%	25	21.9%	0	0.0%
	Patient Activation, Engagement, & Risk Management	87	10.4%	27	23.7%	0	0.0%
	Care Coordination	101	12.1%	29	25.4%	0	0.0%
	System Changes to Support Care	94	11.2%	24	21.1%	0	0.0%
	Other Redesign Activities	3	0.4%	0	0.0%	0	0.0%
Prior experience with payment initiatives	Bundled Payments	22	2.6%	1	0.9%	0	0.0%
	Pay-for-Performance	69	8.2%	16	14.0%	1	0.7%
	Shared Savings	35	4.2%	0	0.0%	0	0.0%
	Other Payment Incentives	7	0.8%	1	0.9%	0	0.0%

Source: Lewin analysis of Awardee-submitted data collected February 2015 through February 2016 for Model 3 episode initiators participating in BPCI between Q4 2013-Q4 2015

*Prior experience in care redesign and payment incentive initiatives was reported when episode initiators first became active in Phase 2 of BPCI. Therefore, PGP episode initiators that changed models are only included in the sample for the first model in which they participated. This table only includes information about episode initiators that submitted data. EIs that did not submit complete data are excluded from the counts in the table.

a. Episode and length selection

The average Model 3 EI participated in six clinical episodes; three EIs participated in all 48 clinical episodes. As shown in Exhibit 165, the most popular clinical episode among Model 3 participants was MJRLE, in which 32% of EIs participated. Congestive heart failure (CHF) was the second most common clinical episode, chosen by 25% of EIs. During the first eight quarters of the initiative, EI participation was less than 10% in nine clinical episodes.

[1] We also collected data from nine IRFs and one LTCH. However, due to limited sample sizes, we did not present care redesign and payment incentives for these facilities.

Exhibit 165: Participation of Episode Initiators by Clinical Episode, Model 3, Q4 2013 - Q3 2015

Clinical Episode	Episode Initiators by Participant Type (N=1,143)					
	SNF (N=873)	HHA (N=116)	IRF (N=9)	LTCH (N=1)	PGP (N=144)	%
Major joint replacement of the lower extremity	308	32	1	1	26	32%
Simple pneumonia and respiratory infections	230	40	1	1	12	25%
Congestive heart failure	205	50	1	1	6	23%
Chronic obstructive pulmonary disease, bronchitis, asthma	210	37	0	1	3	22%
Urinary tract infection	175	27	0	0	9	18%
Other respiratory	170	29	0	1	9	18%
Sepsis	181	10	1	1	6	17%
Stroke	165	23	3	0	7	17%
Acute myocardial infarction	157	23	0	0	9	17%
Hip & femur procedures except major joint	162	11	1	1	8	16%
Fractures of the femur and hip or pelvis	167	9	1	0	2	16%
Cardiac arrhythmia	145	19	0	0	10	15%
Medical non-infectious orthopedic	163	7	0	0	4	15%
Nutritional and metabolic disorders	147	11	0	0	8	15%
Renal failure	141	10	0	0	9	14%
Other vascular surgery	125	22	0	0	8	14%
Syncope & collapse	143	3	0	0	8	13%
Cellulitis	140	6	0	0	6	13%
Major bowel procedure	136	7	0	0	3	13%
Medical peripheral vascular disorders	123	20	0	0	3	13%
Cardiac valve	119	21	0	0	5	13%
Coronary artery bypass graft	115	25	0	0	4	13%
Esophagitis, gastroenteritis and other digestive disorders	133	4	0	0	4	12%
Pacemaker	130	4	0	0	5	12%
Gastrointestinal hemorrhage	126	6	0	0	5	12%
Gastrointestinal obstruction	127	3	0	0	4	12%
Percutaneous coronary intervention	106	22	0	0	4	12%
Transient ischemia	121	6	0	0	4	11%
Revision of the hip or knee	114	14	1	0	1	11%
Major joint replacement of the upper extremity	117	9	1	0	2	11%
Chest pain	107	16	0	0	4	11%
Pacemaker device replacement or revision	82	1	0	0	44	11%
Diabetes	111	10	0	0	4	11%
Red blood cell disorders	109	6	0	0	7	11%
Lower extremity and humerus procedure except hip, foot, femur	112	8	0	0	1	11%
AICD generator or lead	99	3	0	0	14	10%

Clinical Episode	Episode Initiators by Participant Type (N=1,143)					
	SNF (N=873)	HHA (N=116)	IRF (N=9)	LTCH (N=1)	PGP (N=144)	%
Other knee procedures	112	3	0	0	1	10%
Spinal fusion (non-cervical)	102	11	1	0	2	10%
Amputation	110	3	0	0	2	10%
Back & neck except spinal fusion	95	8	1	0	0	9%
Double joint replacement of the lower extremity	87	4	3	0	9	9%
Atherosclerosis	95	2	0	0	4	9%
Removal of orthopedic devices	92	2	0	0	3	8%
Cardiac defibrillator	71	19	0	0	1	8%
Major cardiovascular procedure	70	18	0	0	2	8%
Complex non-cervical spinal fusion	80	4	0	0	0	7%
Cervical spinal fusion	59	8	1	0	3	6%
Combined anterior posterior spinal fusion	42	2	0	0	2	4%

Source: Lewin analysis of CMS' BPCI database, June 2016. Note: The sum of EIs participating in each of the clinical episodes exceeds the total number of EIs because EIs can participate in more than one clinical episode.

Interview Findings Reported Previously

- Interviewees selected high volume episodes. They also selected episodes that were challenging and provided a learning opportunity and those that could be effectively managed by the PAC provider.

In terms of episode selection, interviewees noted that they compared provider-specific historical data to national or regional benchmarks to identify areas of greatest financial opportunity. Many noted that these analyses were conducted by consultants or Conveners. Similar to previously reported findings, most respondents indicated that expected patient volume was another important factor in their choice of clinical episodes. Others noted the role of their corporate office or Convener in episode selection.

Interview Findings Reported Previously

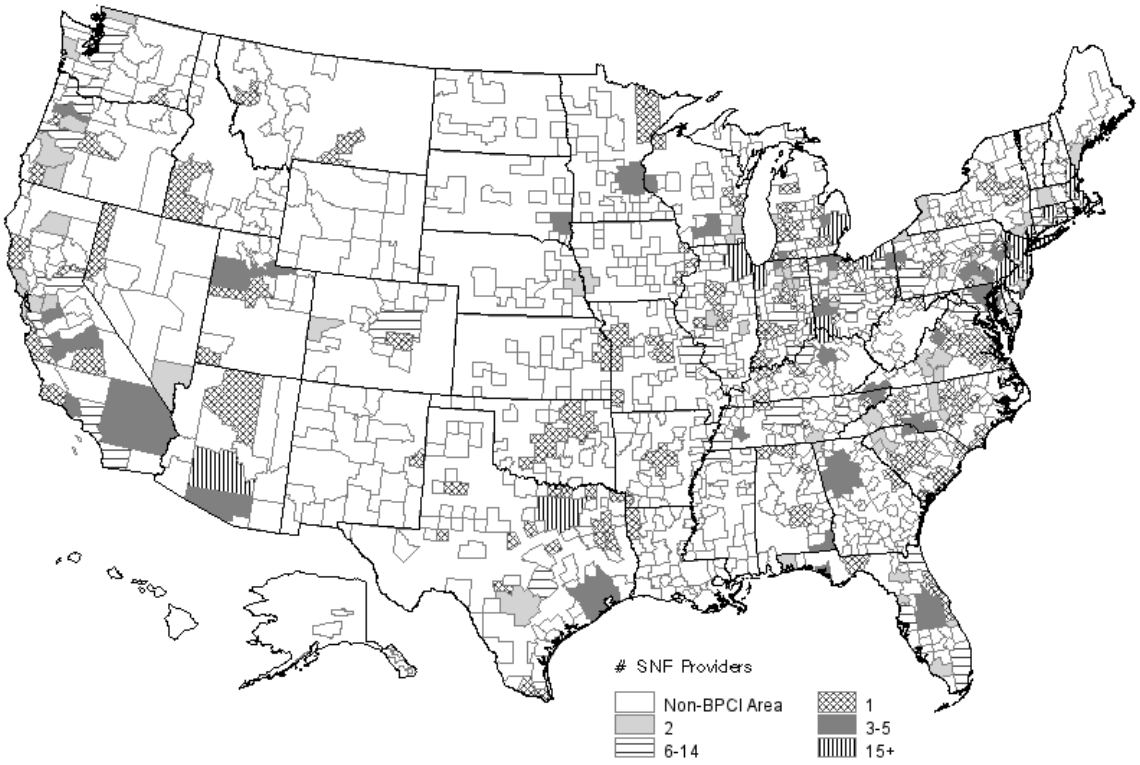
- In choosing the 90-day episode period, interviewees stated that the 90-day period was strongly recommended by CMS when sites were applying to participate in BPCI.
- Interviewees felt that 90 days allowed for their care redesign system to function as intended.

The most popular episode length among Model 3 EIs was the 90-day episode. HHAs tended to select the longest episode lengths compared to SNFs and PGPs, with 84%, 73%, and 70%, selecting the 90-day episode, respectively. Interviewees' indicated that they chose 90-day episodes because they were strongly urged to select 90 days by CMS when they were applying to participate in BPCI. Respondents indicated that they were under the impression that joining BPCI was a competitive process and selected 90 days because it would improve their chances of participation. Others selected the 90-day episode because it allowed them to better evaluate if their care redesign interventions were working or because the 90-day length was considered more appropriate for chronic condition episodes.

4. Market characteristics

The 1,143 Model 3 BPCI EIs that participated in the first eight quarters of the BPCI initiative were located throughout the country. Exhibits 166 to 170 display the geographic locations of the episode-initiating SNF, IRF, HHA, LTCH, and PGP Model 3 EIs. In the following tables and narrative, we compare characteristics of the markets where Model 3 BPCI PAC EIs were located with the areas where there were no BPCI-participating PAC EIs during the first two years of the initiative.[2]

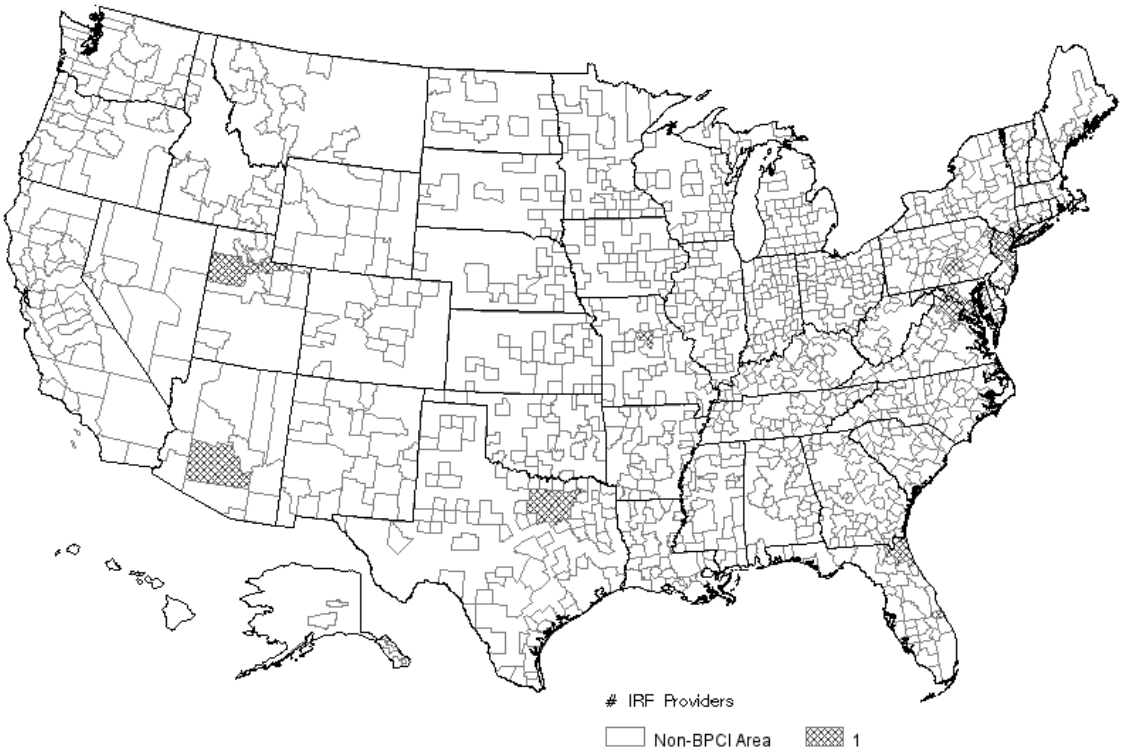
Exhibit 166: Number of BPCI Participating SNFs by CBSA, Model 3, Q4 2013 – Q3 2015



Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating SNF EIs.

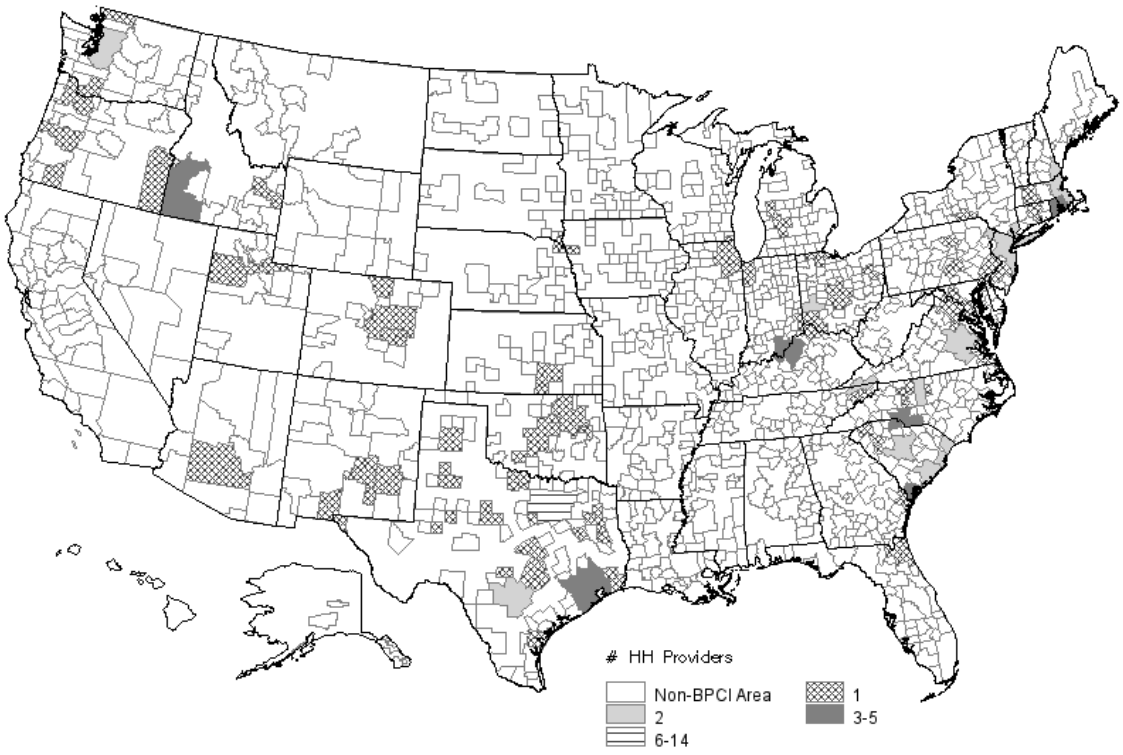
[2] The market is defined as the Core Based Statistical Area (CBSA). Providers not located within a CBSA were assigned to the largest CBSA within their Hospital Referral Region (HRR). Non-BPCI markets represent all CBSAs that do not have a Model 3 BPCI PAC participant. Areas of the country that are not in a CBSA are therefore not included in these non-BPCI markets.

Exhibit 167: Number of BPCI Participating IRFs by CBSA, Model 3, Q4 2013 – Q3 2015



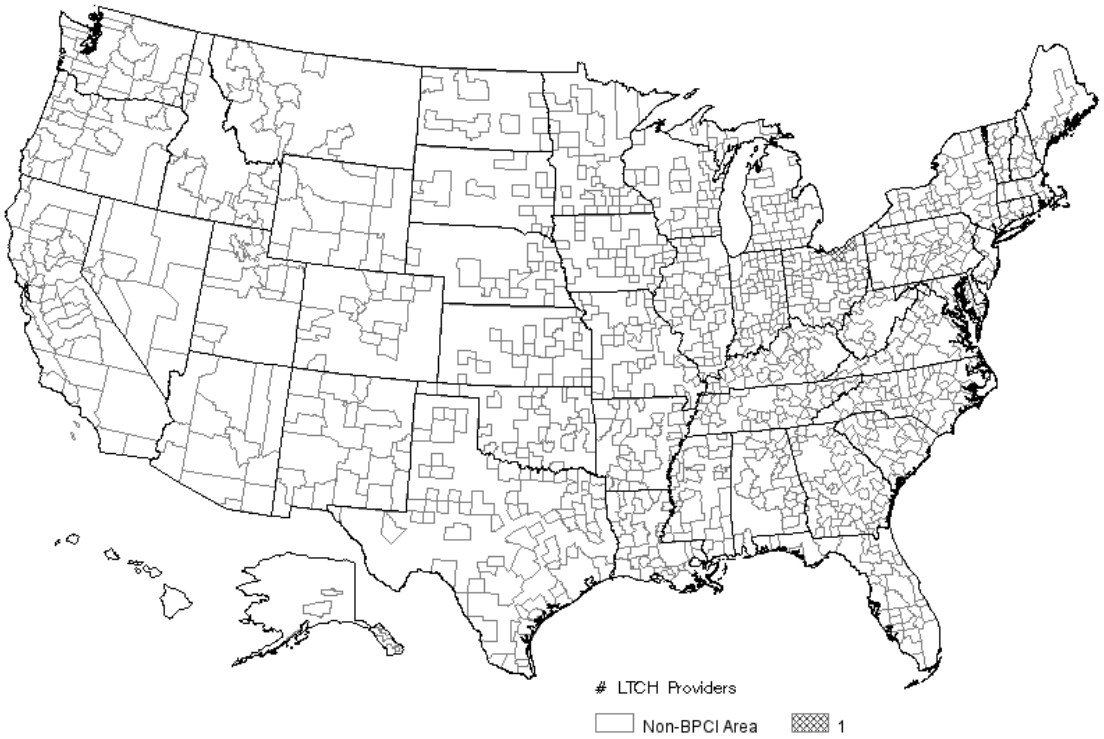
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating IRF EIs.

Exhibit 168: Number of BPCI Participating HHAs by CBSA, Model 3, Q4 2013 – Q3 2015



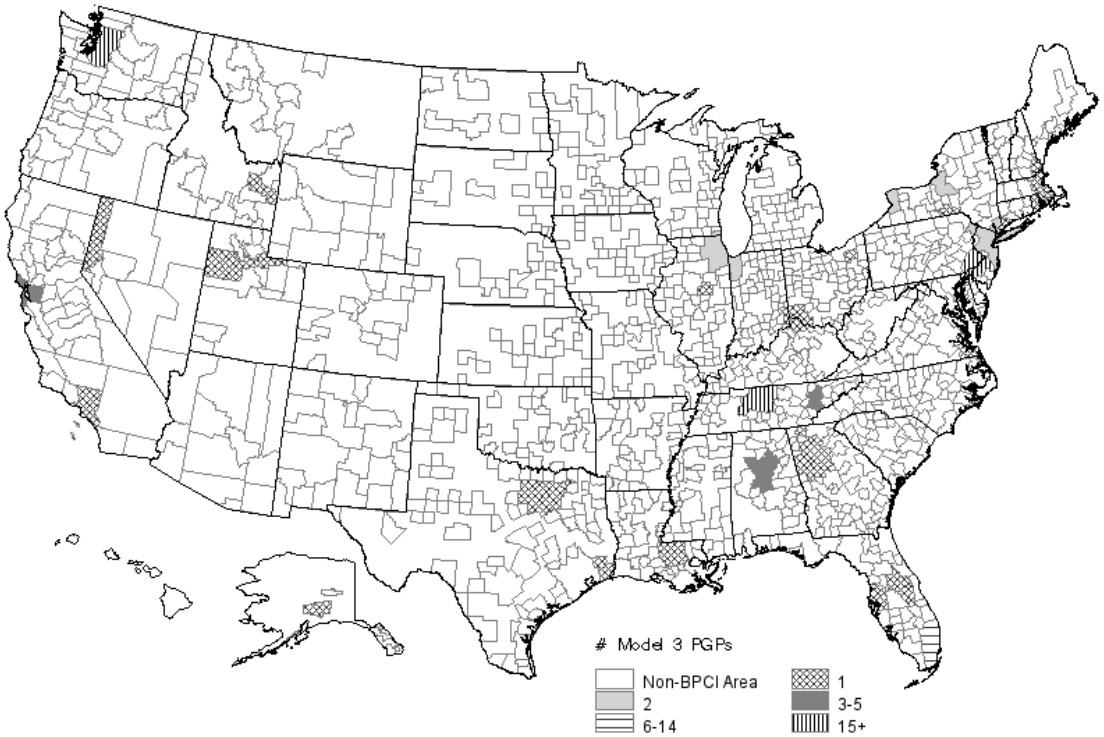
Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating HHA EIs.

Exhibit 169: Number of BPCI Participating LTCHs by CBSA, Model 3, Q4 2013 – Q3 2015



Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating LTCH EIs.

Exhibit 170: Number of BPCI Participating PGPs by CBSA, Q4 2013 – Q3 2015



Source: Lewin analysis of CMS' BPCI database for all Q4 2013 – Q3 2015 BPCI participating PGP EIs.

Markets with BPCI-participating Model 3 PAC EIs differed from markets without them, as shown in Exhibit 171, which includes key characteristics for all Model 3 EIs. The average market penetration for BPCI-participating SNFs, defined as the proportion of all SNF admissions in the CBSA related to the 48 BPCI clinical episodes to a BPCI-participating SNF, was 17.8%; the average market penetration for BPCI-participating HHAs was 6.5%. BPCI markets had high SNF competition (average Herfindahl index of 0.16), relative to non-BPCI markets (0.39). The same was true for BPCI HHA markets compared to non-BPCI markets (Herfindahl index of 0.38 vs. 0.59).

BPCI markets were more populous (average population 761,276) compared with markets without BPCI-participating SNFs or HHAs (average population 104,462). On average, BPCI markets had higher median household income than non-BPCI markets ($47,251 vs. $42,820), as well as more primary care physicians, specialists, and nurse practitioners per 10,000 residents.

Exhibit 171: Characteristics of BPCI Markets and Non-BPCI Markets, Model 3 PAC EIs, Q4 2013-Q3 2015

Market Characteristics	BPCI Markets N=295; 31.3% of Markets				Non-BPCI Markets N=647; 68.7% of Markets			
	Mean	Median	25th	75th	Mean	Median	25th	75th
BPCI Market Penetration – Hospital	19.5%	0.0%	0.0%	38.0%	5.2%	0.0%	0.0%	0.0%
Herfindahl Index – ACH	0.51	0.49	0.22	1.00	0.74	1.00	0.50	1.00
BPCI Market Penetration – SNF	17.8%	12.0%	6.0%	26.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index – SNF	0.16	0.13	0.05	0.23	0.39	0.33	0.21	0.50
BPCI Market Penetration – HHA	6.5%	0.0%	0.0%	1.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index – HHA	0.38	0.28	0.13	0.59	0.59	0.57	0.29	1.00
BPCI Market Penetration – IRF	1.9%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Herfindahl Index – IRF	0.27	0.00	0.00	0.51	0.06	0.00	0.00	0.00
Medicare Advantage Penetration	22.6%	20.7%	12.3%	32.2%	16.3%	13.8%	7.4%	21.8%
Population	761,276	203,882	89,352	627,664	104,462	53,278	35,544	107,484
Median Household Income	$47,251	$45,957	$41,558	$51,979	$42,820	$41,774	$37,042	$47,585
% Age 65+	15%	14%	12%	16%	15%	15%	13%	17%
PCP Per 10,000	6.9	6.7	5.5	8.1	6.2	5.9	4.6	7.5
Specialist Per 10,000	7.5	6.3	4.3	9.4	4.6	3.8	2.3	5.9
PA/NPs Per 10,000	6.5	6.0	4.4	7.9	5.9	5.4	3.7	7.3
SNF Beds Per 10,000	66.6	58.1	41.6	84.1	72.4	67.1	44.9	92.7
LTCH Beds Per 10,000	0.8	0.0	0.0	1.1	0.4	0.0	0.0	0.0
IP Rehab Bed Per 10,000	0.6	0.0	0.0	0.6	0.2	0.0	0.0	0.0
Critical Access Hospital Beds Per 10,000	0.8	0.0	0.0	0.4	2.1	0.0	0.0	2.2

Source: Lewin analysis of 2011 Medicare claims and 2011 Area Health Resource File (AHRF).

Note: 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 PAC participant. Areas of the country that are not in a CBSA are therefore not included in these non-BPCI markets. Variable definitions are included in **Appendix L**.

5. Model incentive structure characteristics

a. Conveners in BPCI

Interview Findings Reported Previously

- Interviewees reported they joined BPCI as a Facilitator Convener (FC) to gain access to CMS data, enhance their relationship with local hospitals, or be a leader in care redesign.
- FCs varied in the role they played; some facilitated shared learning and the exchange of best practices, others were more involved in care redesign and helped develop new care protocols.
- Awardee Conveners (ACs) and Designated Awardee Conveners (DACs) also provided a varied level of support to their Awardees/EIs. Some helped with program set-up (e.g., episode selection), assisted with data analytics, and provided administrative support, while others were involved in the design, development, and implementation of care protocols.

Interviewees that joined BPCI under a Convener often reported that they did so because this structure included benefits such as sharing risk, having administrative and analytic support for the initiative, and learning from the Convener's experience. As noted in prior evaluation reports, interviewees reported that their Conveners (both ACs and FCs) provided data analysis support, either directly or through an arrangement with a third-party contractor. Interviewees also reported that their ACs assisted with care redesign, providing a call center service to track BPCI patients, offering software to support care transitions, assisting with development of care protocols, and facilitating relationships with hospital partners. PGP interviewees noted that ACs provided care coordination software and helped to establish formal partnerships with preferred PAC partners. Interviewees also reported that ACs shared lessons learned from other BPCI participants and helped identify promising strategies to achieve savings.

Among HHAs, EIs participating under an AC or DAC (N=104) were much more likely to choose 90-day episodes (93% vs. 29%) than other HHA EIs (N=12). Generally, EIs under an AC or DAC chose to participate in more clinical episodes (7.3 vs 5.3) and were less likely to sign up for only one clinical episode (33% vs 51%). Among SNFs, there was no difference in choice of episode length between EIs participating under an AC or DAC (N=783) and other SNF EIs (N=90). However, SNF EIs under a convener, like HHAs, chose more clinical episodes (7.7 vs. 2.6) and were less likely to participate in only one clinical episode (52% vs 66%).

b. Partners

Interview Findings Reported Previously

- Interviewees cited more collaborative relationships and sharing common care strategies and protocols as benefits of partnerships.
- Interviewees continued to partner with other PACs to ensure downstream quality, track patients after discharge to reduce readmission rates, and to stratify patient risk.
- Interviewees partnered with third-party administrators and contractors to assist with data management, BPCI Model administration and oversight, and gainsharing payment calculation and distribution.

In quarterly interviews and site visits, Model 3 interviewees described collaborations with a variety of entities including hospitals, other PAC providers, external consultants, and other organizations in the community in response to BPCI. Model 3 interviewees described the

importance of strengthening relationships with hospital referral partners to differentiate themselves from others in the market. They noted that being a preferred provider was crucial, especially in competitive markets. To help develop and strengthen these partnerships, interviewees described promoting their BPCI participation, demonstrating their ability to provide high quality care by sharing outcomes data, and joining in collaborative care management efforts. PGP interviewees described partnering with hospitals to encourage alignment of goals. For instance, several PGP interviewees discussed integrating existing transitional care programs from the hospitals into their BPCI care redesign activities.

Similar to what we heard last year, Model 3 interviewees discussed partnerships with other PAC providers to track beneficiaries after discharge, ensure quality, and align goals. Interviewees indicated that they selected preferred PAC partners based on quality metrics (e.g., star ratings, readmission rates, length of stay), clinical capabilities, and willingness to follow protocols and share information. To help align goals, several PGP interviewees provided PAC partners with scorecards that included efficiency and quality data. According to interviewees, relationships with preferred PAC partners have strengthened as a result of BPCI, and the interviewees attributed this success to increased communication and collaboration.

> "We are in heavy contact with [the HHA], especially about our bundled patients, and we can print off documents and make a packet on a case-by-case basis for [the HHA] that has all the information they need to provide the care the patient needs."
>
> – Model 3 SNF

As they reported last year, PAC provider interviewees continued to partner with external contractors and consultants for help with data management and program support. Partners conducted data analysis to help Model 3 participants select episodes, evaluate performance, and identify participation opportunities in health care reform initiatives, such as BPCI. In addition to performing data analysis, external partners were also involved in supporting care redesign activities such as care coordination and telemonitoring. For Model 3 PGP interviewees, Conveners were key players in these roles. Interviewees also described strengthening relationships with organizations in the community such as "Meals on Wheels" or transportation services, noting that these organizations helped beneficiaries after discharge. Relationships with physicians varied among PAC interviewees; some interviewees mentioned having formal arrangements (such as gainsharing agreements or ACO networks) with physicians or PGPs, while others described informal partnerships. Some PAC interviewees became preferred providers for these physicians after teaching them about the BPCI initiative.

c. *Waiver use*

The BPCI initiative allows the waiver of Medicare program requirements with respect to telehealth and post-discharge home visits for Model 3 Awardees to facilitate the implementation of care redesign. In order for an Awardee to provide beneficiary incentives or to engage in gainsharing agreements under BPCI that may be protected under the BPCI fraud and abuse law waivers, the Awardee must include its plans for those activities in its Implementation Protocol (IP) and the IP must be accepted by CMS. Most (71%) Model 3 EIs could provide beneficiary incentives because their Awardee had included plans for doing so in the IP that was accepted by CMS, but only 5% of those eligible used the waiver (Exhibit 172). With regard to gainsharing,

14 of the 56 Awardees entered into gainsharing agreements from Q4 2013 through Q2 2015. Few EIs used the telehealth (1%) or post-discharge home visit waivers (0%).

Exhibit 172: Participation of Episode Initiators in BPCI Waivers, Model 3, Q4 2013 - Q3 2015*

Model 3 Waivers	Model 3 EIs (N=690)			
	EIs that signed up to use the waiver		EIs that used waiver of those signed up	
	N	%	N	%
Beneficiary Incentives	490	71.0%	25	5.1%
Telehealth**	NA	NA	7	1.0%
Post-Discharge Home Visit**	NA	NA	1	0.1%

* The number of EIs includes those who had at least one BPCI episode between Q4 2013 – Q3 2015. Data on the Awardees that signed up for the gainsharing waiver in their IPs are presented in Exhibit 174.

**The Post-Discharge Home Visit and Telehealth waivers are available to all Model 3 Awardees without specifying it in their IPs. Therefore, the denominator used to calculate the percentage of M3 EIs that used these two waivers was the number of Model 3 EIs through Q3 2015.

Sources: Lewin analysis of Awardee IPs for Q3 2015 BPCI participants, Medicare claims data for episodes initiated Q4 2013 – Q3 2015, and Awardee-submitted data Q4 2013 – Q3 2015.

Note: The 690 EIs that participated Q4 2013 - Q3 2015 are distributed among 87 Model 3 Awardees. EIs include SNF, HHA, PGP, IRF, and LTCH.

Beneficiary incentives

Although most EIs that could provide beneficiary incentives did not, interviewees indicated that they were considering beneficiary incentives, such as personal emergency response systems and transportation assistance. Equipment was the most widely reported incentives category, though only two Awardees distributed this type of incentive (Exhibit 173).

Exhibit 173: Beneficiary Incentives Distributed by EIs to Beneficiaries, Model 3, Q3 2015

Incentive Description	Awardees allowed to use incentives	EIs allowed to use incentives that had at least one BPCI episode	Awardees that provided one or more incentive	Episodes receiving one or more incentive	Average cost per incentive ($)
Transportation	16	16	1	13	$143.14
Equipment	7	22	2	182	$32.54
Home care/home visits	4	5	1	44	$286.74
Living arrangement services	2	2	1	21	$181.57
Telehealth/technology	25	57	7	57	$111.51
Wellness program/resources	3	405	0	0	NA
Medication management tools	2	2	1	1	$85.00

Source: Lewin analysis of Awardee-submitted data, Q3 2015.

Gainsharing

Interview Findings Reported Previously

- Interviewees viewed gainsharing as a tool to incentivize cooperation among upstream (e.g., hospitals) and downstream (e.g., other PAC providers) providers in an episode.
- Interviewees cited market characteristics as a reason for not gainsharing. For example, in markets with an oversupply of SNF beds, potential gainsharing partners such as hospitals had little reason to coordinate care with individual SNFs, even if provided with gainsharing funds.
- Interviewees described conducting thorough reviews of the providers in their market to narrow the pool of prospective gainsharing partners. Interviewees also described developing a list of requirements for their gainsharing partners, such as a reasonable readmission rate.

Fourteen Awardees entered into gainsharing agreements from Q4 2013 through Q2 2015, the time period from which NPRA earnings would be given to the Awardees from CMS and available for distribution through Q1 2016, and 20 Awardees reported gainsharing data from Q4 2013 through Q1 2016. Nine Awardees reported distributing approximately $4.6 million. All of the distributed gainsharing amounts were from NPRA, that is, positive differences between the target amount and actual Medicare fee-for-service (FFS) payments during the episode. Awardees did not report distribution of internal cost savings (ICS). Exhibit 174 describes gainsharing distributions received by type of gainsharing partners among Model 3 Awardees from Q4 2013 through Q1 2016. Physicians and institutional PAC providers were the most common partner to receive a distribution.

"Everyone needs to accept risk in order to improve quality of care."
– Model 3 SNF

Exhibit 174: Gainsharing Distributions Received by Partner Type, Model 3, Q4 2013 – Q1 2016

Gainsharing Partners	Awardees that reported eligible partners of this type	Number of Partners Receiving a Distribution	Number of Partners receiving NPRA	Number of Partners receiving ICS
Physicians	10	124	124	-
PGPs	7	2	2	-
Hospitals	7	13	13	-
Institutional PAC	11	44	44	-
HHAs	7	24	24	-
Other	1	3	3	-

Source: Lewin analysis of Awardee-submitted data collected February 2016 and May 2016 for Model 3 EIs participating in BPCI between Q4 2013 – Q1 2016.

Based on interviews with representatives of organizations that had developed gainsharing arrangements, interviewees indicated that they were most likely to initiate gainsharing with physicians and hospitals. Fewer mentioned gainsharing with PAC providers. The most common reason for participating was to increase physician engagement, develop collaborative relationships, and share new quality metrics. Some interviewees expected to change provider behavior, but most did not have a clear indication of whether this had happened.

The most common reason interviewees gave for not participating in gainsharing was that they felt they had sufficient influence over the other care settings, either because they were part of a

larger health system or because they used a care navigator to monitor patients across settings. Another common reason interviewees gave was that they were unsure of their financial outcomes and the amount of savings that would be available to share. A few indicated that they were not gainsharing because they either had not been asked to do so by their Convener or they had not identified appropriate gainsharing partners.

Telehealth and Post-Discharge Home Visit waivers

Telehealth and Post-Discharge Home Visit waivers were infrequently discussed during interviews with Model 3 organizations. Several Model 3 interviewees indicated that telehealth was important in their care redesign activities, but they did not indicate whether they used the telehealth waiver.

6. Care redesign and cost savings

Interview Findings Reported Previously

- Care Redesign Activities
  - Interviewees started using different risk assessment tools as a part of their care redesign strategies.
  - To reduce readmissions, interviewees focused on "front-end" education rather than education at discharge (e.g., self-care, length of stay (LOS) expectations, avoiding ED visits).
- Care Coordination
  - Interviewees indicated that the roles of case managers, care navigators, or care coordinators were crucial to their care redesign efforts. Although specific roles varied by participant, these staff helped communicate with primary care physicians, fill prescriptions, coordinate home support, and track patients after discharge.
- Cost Savings Strategies:
  - Interviewees described managing PAC utilization as a key strategy to reduce costs. Their methods focused on diverting patients to HHAs and reducing the average SNF LOS.
  - Reducing readmissions was another approach taken to reduce costs. Interviewees reported engaging in various activities to accomplish this goal, including conducting root cause analysis for all readmissions, educating staff on the common causes of readmissions, and keeping patients in the SNF for an extra day to reduce readmission risk.

"It used to be that a patient came in and we created a person dependent on us. [Now] we incorporate teaching to consider how they can be involved and learn what they need to."
- Model 3 SNF

All Model 3 site visit interviewees discussed how they incorporated patient education into their care redesign activities. As noted above, a common theme was early patient education. This was especially crucial when PAC length of stay was reduced. As one interviewee stated, "We can't wait for 14 days to have a care conference when the patient is gone in 10." Additionally, interviewees mentioned that consistent messaging is delivered by all staff members throughout the patients' stay and family members are frequently engaged in the educational processes.

Model 3 interviewees detailed a range of goals for patient education. Interviewees commonly discussed how BPCI was described to patients and their families, although some stated that education about the initiative was difficult due to a lack of publically available information and skepticism from patients and their families. Other patient education topics included information

about their medical conditions, pain management, symptom recognition, and follow-up care plans. Specific tools included journals for patients to track their symptoms and printed materials. Model 3 interviewees said that these educational efforts were helpful in managing patient expectations, improving patient adherence, reducing complications, and preventing readmissions. For some interviewees, patient education included encouraging patients to first call their PCP or return to the SNF if their condition deteriorated after discharge rather than go to the ED.

Interviewees indicated that they begin the discharge planning process earlier since joining BPCI, and they give more consideration to the appropriate discharge destination and types of services provided after discharge. Several indicated that they begin planning for discharge on the patient's first day and include family members to discuss the patient's care needs and discharge goals. To increase the feasibility of discharging patients home, interviewees reported that they helped to address social issues by connecting patients with community resources, such as "Meals on Wheels", and transportation services. PGP interviewees under one Convener described the use of the Care At the Right Location (CARL) tool to help select the most appropriate next level of care, while PGPs under another Convener reported working with a multidisciplinary team in the hospital to determine the most appropriate PAC setting. Some HHA interviewees reported that they had liaisons in the SNF setting to visit patients and gather relevant information prior to their discharge home.

> "Our big fear has been that if we discharge a patient too early, the patient would be readmitted."
> - Model 3 HHA

Interviewees discussed methods to enhance their post-discharge follow-up process, including conducting post-discharge calls and increasing their coordination with other PAC providers. One PGP Convener noted that when patients were discharged to a SNF, they worked with the SNF to ensure that the patient's LOS was based on care needs rather than payment potential. Many SNF interviewees reported that they educated their HHA partners about the care goals for their patients, and they worked to ensure that the care plans were followed. Some SNF interviewees reported that they sent their therapy notes or an informational packet on the patient's care to the HHA at the time of discharge. Several interviewees also noted that patients took their care journals with them upon discharge to smooth the transition between levels of care.

In addition to coordinating with other PAC providers, interviewees highlighted the importance of collaborating with PCPs and increasing physician engagement in the care of BPCI patients after PAC discharge. In particular, many interviewees reported that they helped their patients schedule follow-up appointments with their PCPs. Interviewees also communicated directly with PCP offices to provide updates on patient status and share discharge plans with the physicians.

PAC provider and PGP site visit interviewees used risk stratification as a component of their care redesign strategy. They reported assessing for many types of risk, including readmissions, patient complications, and increased length of stay. Interviewees described a wide range of metrics and tools, such as LACE (which includes length of stay, acuity, co-morbidities, and emergency department visits), OASIS (Outcome and Assessment Information Set), and PAM (Patient Activation Measure), while others described relying on home-grown tools that brought together numerous patient metrics.

Risk stratification was most commonly used as a tool to determine patient care needs. One interviewee stated that BPCI led to a more formalized risk assessment process that began before hospital discharge. In this example, clinicians began using a standardized risk assessment and looked at patient risk scores to "front load" visits for higher risk patients during their first two weeks at the facility. Interviewees at this site visit also noted that the risk stratification process helped them create tailored care plans for patients. Other uses for risk stratification included tailoring education and goal setting based on patient risk scores, targeting post-discharge services, and determining the frequency of follow-up after discharge.

Model 3 site visit participants noted various approaches to care standardization under BPCI. Some reported that they standardized their approaches to pain management and medication reconciliation. Interviewees specifically discussed standardized care protocols for certain episodes, such as sepsis or joint replacement. Sometimes the standardization was initiated by the participant's Convener or corporate-level leaders.

Interviewees indicated that they needed to hire additional staff or shift current staff responsibilities to meet the clinical and administrative demands of BPCI. Several interviewees reported hiring nurse practitioners or care coordinators to provide education, discharge planning, and post-discharge follow up for BPCI patients. Some also hired non-clinical staff to assist with data entry and admissions coordination under BPCI.

We asked SNF interviewees about the financial implications of reducing LOS and how that compared with the potential for positive NPRA. Most of the 25 interviewees were hesitant to discuss how their BPCI NPRA compared with the Medicare per diem payments. A few said that they had successfully reduced LOS, which led to the implementation of strategies to increase their volume to maintain revenue. Most indicated that patient turnover had increased under BPCI and that they relied on relationships with hospitals for additional referrals to make up for shorter SNF stays.

7. Implementation challenges

Interview Findings Reported Previously

- Interviewees indicated that episodes could include readmissions that were unrelated to the initial hospitalization or that the EI could not prevent.
- Interviewees encountered difficulties coordinating care with physicians and found that referring hospitals, especially those outside of their market, were often unwilling to collaborate.
- Interviewees encountered difficulties coordinating care with physicians and found that referring hospitals, especially those outside of their market, were often unwilling to collaborate.
- Interviewees noted that obtaining data from referring hospitals to facilitate patient identification was often a challenge. Specific data-related challenges included concerns related to HIPAA, interoperability issues, and inconsistency between the expected and final MS-DRG. Additionally, interviewees noted that DRGs often take months to confirm, which leads to misallocation of resources.
- In some markets, there were challenges related to the increased number of partner hospitals that joined BPCI as Model 2 participants, such as identifying which facility had precedence as the EI for a given patient.
- Interviewees also noted that partner hospitals that were also in BPCI (Model 2) might reduce SNF admissions, which lowered the SNF's occupancy rate.

- Interviewees were deterred from gainsharing by the financial investment and time required to execute gainsharing agreements.
- Interviewees discussed the difficulty in changing the mindset of patients who felt entitled to 21 days in the SNF, even if it was not medically necessary. In addition, interviewees cited that patients were often resistant to changing their perspectives that only hospitals/EDs could provide acute care as opposed to the SNF.
- Interviewees mentioned the challenge of changing the culture among staff regarding discharge procedures.
- Interviewees reported that an unexpected increase in the number of medically complex patients posed a challenge to their cost-saving strategies, as these patients typically require higher cost treatments, longer LOS, and are at higher risk for readmission.

The majority of Model 3 interviewees noted challenges in identifying BPCI patients because the MS-DRG was not always known at the time of PAC admission. Although some interviewees said that identifying BPCI patients has become easier, others indicated that patient identification was one of the biggest challenges under BPCI because of the lag in receiving the final MS-DRG assignment from hospitals. Some indicated that it can be difficult to get hospitals to share the MS-DRG assignment. Another common challenge was the discrepancy between expected and final MS-DRG, with some interviewees noting that CHF patients were particularly difficult to identify.

Interviewees used patient tracking systems and decision support tools to aid in identifying BPCI patients. They often collaborated with hospitals to obtain patient MS-DRGs. They also relied on PAC staff to identify BPCI patients through pre-admission communication with patients or through internal meetings where patient data were reviewed.

Model 3 interviewees described challenges related to their partners. Interviewees reported that hospitals participating in BPCI, ACOs, or the CJR Model, have been less cooperative. Some BPCI-participating hospitals and ACOs are pursuing cost savings by reducing SNF admissions, which in turn, affects the SNF occupancy rate. The multiple payment models have made it more challenging to determine who "owns" a given patient. One interviewee noted, however, that hospitals now "see the value (in) collaboration" due to the various payment models.

With respect to managing their patients, some interviewees indicated that effective discharge planning was difficult when patients did not want to leave the SNF or did not have a stable place to live after discharge. Several interviewees also reported challenges related to the expectations of their own staff. A common concern focused on staff training and education on BPCI. Interviewees noted that it has been a challenge to ensure consistency in staff members' understanding of BPCI and the goals of the initiative. One interviewee noted that staff turnover further complicated their training efforts. Alternatively, one interviewee did report that BPCI had enhanced communication between their multidisciplinary team members and staff.

8. Participants that terminated or withdrew from BPCI

Our analysis of Model 3 participants that have withdrawn or terminated from BPCI found that they did so due to a lack of internal capacity to support the initiative, such as not receiving support from internal staff, or for financial reasons, such as receiving an insufficient volume of patients or not realizing positive NPRA. In cases where EIs withdrew for financial reasons, Conveners were at times involved in the decision to terminate participation.

During the past year we conducted interviews with Awardees that terminated their participation in the BPCI initiative. Model 3 Awardee interviewees cited challenges with leadership engagement, physician engagement, and finances as the primary reasons for ending their participation in BPCI. One Awardee respondent indicated that only one EI would be withdrawing from BPCI and four would remain. The EI that withdrew experienced leadership changes in their Facility Administrator and Director of Nursing positions and as a result, the planned care redesign changes were not effectively adopted. Another respondent stated that they had received pushback from external physicians and hospitals, both of which did not want the EI "controlling" their patients. Finally, an Awardee with a large number of participating EIs terminated their participation in BPCI for financial reasons. They indicated that all of their facilities were in high-cost, urban areas; however, they felt that their target prices (based on a state-wide averages) were skewed lower because of facilities located in low-cost, rural areas. Due to this concern and their initial reconciliation reports, interviewees from this respondent indicated that continued participation in BPCI would not be feasible.

a. Withdrawn BPCI SNFs between Q4 2013 and Q3 2015

This section discusses characteristics of BPCI Model 3 SNFs that withdrew from BPCI in the first eight quarters (Q4 2013 through Q3 2015) of the initiative. As shown in Exhibit 175, by the third quarter of 2015, 44 SNFs (5% of all Model 3 BPCI SNFs) had stopped participating in all clinical episodes. Withdrawn SNFs were similar to participating BPCI SNFs with respect to average tenure in BPCI, but they initiated more BPCI episodes. All of the withdrawn SNFs operated under a Convener. Additionally, 51% of withdrawn SNFs had positive NPRA, compared to 62% of SNFs that remained in the initiative.

Exhibit 175: BPCI Characteristics among Active and Withdrawn BPCI SNF EIs, Model 3, Q4 2013 – Q3 2015

	BPCI SNF EIs that have not Withdrawn (N=829)	Withdrawn BPCI SNF EIs (N=44)
	Mean	Mean
Tenure in BPCI (months)	4.5	4.0
Average number of BPCI episodes initiated per quarter, within a given clinical episode	2.3	3.4*
Percentage of SNFs operating under a convener	89%	100%*
Percentage of EIs with Positive NPRA+	62%	51%

Source: Awardee-submitted data, claims data, and CMS' BPCI database.

Notes: *Withdrawn BPCI SNF EIs* include all BPCI SNFs that dropped **all** of their clinical episodes at some point during the first eight quarters of the initiative. *BPCI SNF EIs that have not withdrawn* include all SNFs that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative. *Indicates statistical significance at 5% level.

+This measure takes into account EIs that initiated at least one episode between Q4 2013 and Q3 2015. 350 EIs were excluded because they initiated zero BPCI episodes over the course of the intervention period (not withdrawn N=484, withdrawn N=39).

Exhibits 176a, 176b, & 176c compare BPCI-participating and withdrawn SNFs on a number of provider and market characteristics. SNFs that withdrew from BPCI were more likely to be for-profit (100% vs. 85%, p=0.02), had a higher number of admissions for BPCI episodes (207 vs. 132, p<0.01), but were less likely to be part of a chain (7% vs. 20%, p=0.03) than those that

remained. Withdrawn SNFs also had a lower overall nursing home score (2.39 vs. 3.51, p<0.01) based on three areas: quality, survey/health inspections, and staffing. Finally, withdrawn SNFs were more likely to be located in more populous markets (p=0.02), with lower Medicare Advantage penetration (21% vs. 27%, p=0.01) and a higher percentage of patients over the age of 65 (16% vs. 14%, p<0.01).

Exhibits 176a, 176b, & 176c: Provider and Market Characteristics among Active and Withdrawn BPCI SNF EIs, Model 3, Q4 2013 – Q3 2015

Characteristic		BPCI SNF EIs that have not Withdrawn (N=820)		Withdrawn BPCI SNF EIs (N=44)	
		N	%	N	%
Ownership	Non-Profit	121	15%	0	0%*
	Government	3	0%	0	0%*
	For-Profit	696	85%	44	100%*
Urban/Rural	Urban	689	84%	35	80%
Part of Chain	Yes	165	20%	3	7%*
Hospital Based	Yes	7	1%	0	0%

Characteristic	BPCI SNF EIs that have not Withdrawn (N=820)	Withdrawn BPCI SNF EIs (N=44)
	Mean	Mean
Bed Count	122	129
Number of Admissions for BPCI Episode MS-DRGs, 2011	132	207*
Nursing Home Overall Score[1]	3.51	2.39*

Market Characteristics	BPCI SNF EIs that have not Withdrawn (N=820)	Withdrawn BPCI SNF EIs (N=44)
	Mean	Mean
Market Share	18%	16%
Herfindahl Index	0.27	0.26
Medicare Advantage Penetration	27%	21%*
PCPs per 10,000	7.5	7.3
SNF beds per 10,000	62.3	55.5
IRF in Market	56%	64%
Population	2,835,124	4,390,493*
Median Household Income	$51,951	$51,501
% Age 65+	14%	16%*

Source: Lewin analysis of 2013 POS and 2011 Medicare claims, Model 3 SNF EIs, Q4 2013 – Q3 2015.

Note: *Withdrawn SNF EIs* include all BPCI SNFs that dropped all of their clinical episodes at some point during the first eight quarters of the initiative. *BPCI SNF EIs that have not withdrawn* include all SNFs that have continued participation in at least one clinical episode during the first eight quarters of the initiative.

[1] Number of points out of 5 in overall rating and in three areas: Quality, Survey/Health Inspections, and Staffing.

This exhibit excludes BPCI-participating SNFs that received Medicare certification after 2011.

*Indicates statistical significance at 5% level.

b. Withdrawn BPCI HHAs between Q4 2013 and Q3 2015

This section discusses characteristics of BPCI Model 3 HHAs that withdrew from BPCI in the first eight quarters (Q4 2013 through Q3 2015) of the initiative. As shown in Exhibit 177, by the third quarter of 2015, 15 BPCI HHAs (13% of all Model 3 BPCI HHAs) had stopped participating in all clinical episodes. On average, withdrawn HHAs had a significantly longer average tenure in BPCI than HHAs that did not withdraw. HHAs that withdrew were more likely to operate under a Convener and averaged fewer episodes per clinical episode and quarter. Additionally, 33% of withdrawn BPCI HHAs had positive NPRA compared to 53% of BPCI HHAs that have not withdrawn.

Exhibit 177: BPCI Characteristics among Active and Withdrawn BPCI HHA EIs, Model 3, Q4 2013 – Q3 2015

	BPCI HHA EIs that have not Withdrawn (N=101)	Withdrawn BPCI HHA EIs (N=15)
	Mean	Mean
Tenure in BPCI (months)	4.9	12.1*
Average number of BPCI episodes initiated per quarter, within a given clinical episode	8.4	3.1
Percentage of HHAs operating under a convener	88%	100%
Percentage of EIs with Positive NPRA+	53%	33%

Source: Awardee-submitted data, claims data, and CMS' BPCI database.

Note: *Withdrawn HHA EIs* include all BPCI HHAs that dropped **all** of their clinical episodes at some point during the first eight quarters of the initiative. *BPCI HHA EIs that have not withdrawn* include all HHAs that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative. *Indicates statistical significance at 5% level.

+This measure takes into account EIs that initiated at least one episode between Q4 2013 and Q3 2015. 13 EIs were excluded because they did not initiate any BPCI episodes during the intervention period (not withdrawn N=88, withdrawn N=15).

Exhibits 178a, 178b, & 178c compare BPCI HHAs and withdrawn HHAs on a number of provider and market characteristics. Compared to BPCI HHAs that have not withdrawn from the initiative, withdrawn BPCI HHAs were more likely to be for profit (100% vs. 78%, p=0.05) and part of a chain (100% vs. 69%, p=0.01). Markets with withdrawn EIs have a lower Medicare Advantage penetration (15% vs. 23%, p=0.01) and a higher percentage of patients over the age of 65 (15% vs. 13%, p=0.02) than markets where HHA EIs have not withdrawn.

Exhibits 178a, 178b & 178c: Provider and Market Characteristics among Active and Withdrawn BPCI HHA EIs, Model 3, Q4 2013 – Q3 2015

Characteristic		BPCI HHA EIs that have not Withdrawn (N=101)		Withdrawn BPCI HHA EIs (N=15)	
		N	%	N	%
Ownership	Non-Profit	22	22%	0	0%*
	Government	0	0%	0	0%*
	For-Profit	79	78%	15	100%*
Urban/Rural	Urban	81	80%	10	67%
Part of Chain	Yes	70	69%	15	100%*

Characteristic	BPCI HHA EIs that have not Withdrawn (N=101)	Withdrawn BPCI HHA EIs (N=15)
	Mean	Mean
Number of Employed Nurses in HHA	31	15
Number of Admissions for BPCI Episode MS-DRGs, 2011	380	360

Market Characteristics	BPCI HHA EIs that have not Withdrawn (N=101)	Withdrawn BPCI HHA EIs (N=15)
	Mean	Mean
Herfindahl Index	0.34	0.28
Medicare Advantage Penetration	23%	15%*
PCPs per 10,000	7.1	6.5
SNF beds per 10,000	57.9	45.1
IRF in Market	61%	60%
Population	1,895,853	611,265
Median Household Income	$48,929	$45,345
% Age 65+	13%	15%*

Source: Lewin analysis of 2013 POS and 2011 Medicare claims, Model 3 HHA EIs, Q4 2013 – Q3 2015.

Note: *Withdrawn HHA EIs* include all BPCI HHAs that dropped **all** of their clinical episodes at some point during the first eight quarters of the initiative. *BPCI HHA EIs that have not withdrawn* include all HHAs that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative. This exhibit excludes BPCI-participating HHAs that received Medicare certification after 2011.

*Indicates statistical significance at 5% level.

B. Model 3 impact of BPCI

This section presents the impact estimates of Model 3 BPCI on payments, utilization, and quality of care based on episodes that were initiated by PAC providers during the first eight quarters of

the BPCI initiative (Q4 2013 through Q3 2015). Outcomes were analyzed across clinical episodes separately by EI type (SNF and HHA).[3]

1. Payment, utilization, quality

We present results for the key outcomes across 14 Model 3 EI type/clinical episode combinations in this section. The exhibits present the estimated differential change in risk adjusted outcomes for patients receiving care from PAC providers participating in BPCI between the baseline and the intervention period relative to the same change for the patients receiving care from PAC providers in a comparison group. See Section II.D.2 for additional details on the statistical approach. All claim and assessment-based results for the 14 Model 3 clinical episodes are in **Appendix Q and Appendix R.**

Across all of the Model 3 EI type/clinical episode combinations we examined, BPCI providers reduced total costs within the bundle in two -- SNF MJRLE and HHA CHF. The BPCI-participating providers in these clinical episodes reduced the services that are most within their control; there was a statistically significant decline in the number of MJRLE SNF days and in CHF HHA visits relative to the comparison group. In BPCI SNF episodes, there are mixed findings in terms of quality outcomes: a number of clinical episodes with risk adjusted increases in mortality, emergency department use, and readmissions rates also have risk adjusted improvements in Activities of Daily Living (ADL) measures among those who had two assessments. The same pattern can be observed in HHA clinical episodes, though to a lesser degree. These apparently contradictory quality outcomes may be due to the differences in samples, because the ADL measures are only available for approximately three-quarters of patients discharged to PAC. These patients may not be representative of all Model 3 patients due to the reasons they do not have two patient assessments. Patients were not included in the assessment-based quality outcomes if the patient was readmitted to the hospital, died, had little or no cognitive function, or if the episode occurred later during our measurement period and therefore a second assessment was not yet available in the data.

a. BPCI-participating SNFs

Sample characteristics

Before discussing the impact of BPCI among participating SNFs on payments, utilization, and quality, we present some basic statistics by clinical episode in Exhibit 179 to better understand the sample included in the analysis. Among SNF EIs, the number of matched EIs ranged from 58 to 159 for a given clinical episode; these EIs initiated between 431 and 2,321 episodes over the first eight quarters of the initiative. Because providers were allowed to join BPCI over an extended period, these data represent an average of three quarters of participation among SNF EIs. In all clinical episodes analyzed, over 50% of participating EIs joined BPCI in Q2 or Q3 of 2015. As many as 26 (83%) EIs discontinued participation in a single episode over the first eight quarters.

[3] There was insufficient sample size to conduct DiD analyses for any IRF and LTCH clinical episodes. In addition, we did not conduct any impact analysis on PGP EIs. The validated list of physicians for BPCI-participating TINs was not available at the time the analysis for this report was conducted.

Exhibit 179: Characteristics of the Matched BPCI Providers included in the DiD Estimates, Model 3 SNF, Q4 2013 – Q3 2015

Clinical Episodes	Matched EIs (#)	Matched Intervention Period Episodes (#)	Average length of participation (quarters)	EIs that joined in Q3 2015 (%)	EIs that joined in Q2 or Q3 2015 (%)	EIs that stopped participating in the episode (#)	Intervention Period Episodes from EIs that stopped participation (%)
Congestive heart failure	118	1,382	4	31%	58%	3	1%
chronic obstructive pulmonary disease (COPD), Bronchitis, Asthma	69	547	4	33%	52%	0	0%
Hip & femur procedures except major joint	89	633	2	53%	92%	0	0%
Major joint replacement of lower extremity	153	2,321	2	54%	89%	3	5%
Medical non-infectious orthopedic	84	1,180	4	36%	60%	0	0%
Other respiratory	58	431	4	40%	50%	1	1%
Renal failure	72	886	4	42%	57%	0	0%
Sepsis	134	1,858	3	49%	69%	3	3%
Simple pneumonia and respiratory infections	159	1,192	3	41%	71%	3	1%
Stroke	61	659	3	31%	57%	26	83%
Urinary tract infection	103	931	4	41%	60%	0	0%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2015 for BPCI providers. This table is limited to the matched BPCI providers used to calculate the difference-in-differences (DiD) results in the remainder of this section.

Key payment, utilization, and quality of care outcomes

How have the average standardized allowed amounts (Medicare payments and coinsurance/ copayments combined) changed under BPCI?

Across all M3 SNF clinical episodes, there were few statistically significant differences in payment measures (Exhibit 180). One reason we may not have observed statistically significant results in the total payment amount included in the bundle is that although the BPCI SNF standardized allowed amount decreased in all but one clinical episode relative to the change in the comparison group, the BPCI group experienced a relative increase in payments for readmissions in 9 of the 11 clinical episodes. Another reason may be the small sample sizes and the average tenure under BPCI of three quarters, a result of the large number of EIs that joined BPCI in the last two quarters included in this report (Q2 2015 and Q3 2015).

When taking into account non-significant results across clinical episodes, some patterns emerge. The total amount included in the bundle definition for 90-day episodes decreased in 7 of the 11 (64%) clinical episodes relative to the change in the comparison group; this difference was statistically significant for MJRLE episodes ($p<0.01$). Changes in SNF payments seem to be the major driver of these declines in total payments (Exhibit 180). The SNF standardized allowed amount decreased in all but one clinical episodes relative to the change in the comparison group; only MJRLE was statistically significant ($p<0.01$). The HHA standardized allowed amount increased in 7 of 11 (64%) clinical episodes relative to the change in the comparison group; this difference was statistically significant for stroke episodes ($p=0.06$). These payment outcomes appear to corroborate findings from the qualitative data. BPCI participants claimed to be trying to reduce costs incurred by institutional PAC by shifting services to HHAs. At the same time, the increase in the standardized allowed amount for readmissions observed in 9 of 11 (82%) clinical episodes is unexpected, as reducing readmissions is another strategy to reduce costs in the BPCI initiative. The increase in readmission payments was statistically significant for stroke ($1,870, $p<0.01$).

Exhibit 180: Impact of BPCI on Medicare Allowed Payment ($2015) Outcomes, by Clinical Episode, Model 3 SNF, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of episodes initiated Q4 2013 – Q3 2015	Total Amount Included in Bundle Definition[1]	SNF Standardized Allowed Amount, 90-day PDP[2]	HHA Standardized Allowed Amount, 90-day PDP[2]	Readmissions Standardized Allowed Amount, 90-day PDP[2]
Congestive Heart Failure	1,382	-$825	-$1,110	$161	$201
COPD, Bronchitis, Asthma	547	$22	-$837	$9	$1,392
Hip & Femur Procedures Except Major Joint	633	-$2,401	-$1,843	$117	$97
Major Joint Replacement of the Lower Extremity	2,321	**-$2,877**	**-$2,255**	$0	-$235
Medical Non-Infectious Orthopedic	1,180	$407	-$561	-$103	$823
Other Respiratory	431	-$465	-$1,877	$156	$1,170
Renal Failure	886	-$1,178	-$597	$125	$221
Sepsis	1,858	-$1,069	-$7	-$29	-$463
Simple Pneumonia & Respiratory Infections	1,192	-$935	-$1,143	-$16	$327
Stroke	659	$787	-$1,471	$306	**$1,870**
Urinary Tract Infection	931	$466	$100	$55	$125

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. PDP = post-discharge period.

[1] The total amount included in bundle definition values are based on only the 90-day episodes.

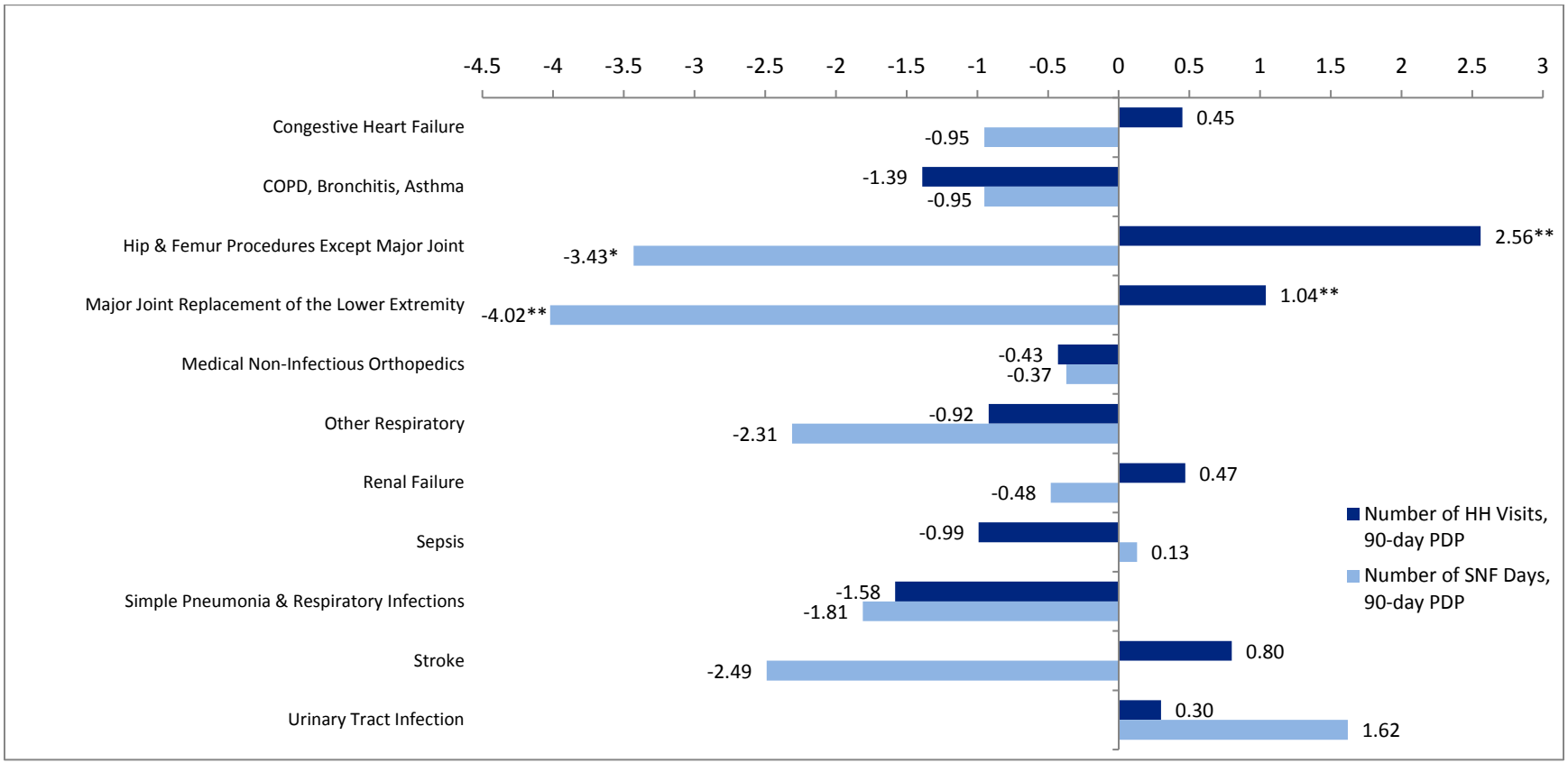
[2] These payment measures are not conditional upon use of the service.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How have the services changed under BPCI?

This section seeks to shed light on how changes in payment measures discussed in the prior section were achieved. When taking into account non-significant results across clinical episodes, the number of SNF days decreased in 9 of 11 (81%) clinical episodes relative to the change in the comparison group (Exhibit 181). This reduction in SNF days aligns with the reduction in SNF payments observed in Model 3 SNFs relative to the change in the comparison group. There was no clear change in HHA utilization from the baseline to intervention period. However, the two clinical episodes that had a significant reduction in SNF days, MJRLE and hip and femur procedures except major joint, had statistically significant increases in HH visits. This suggests that for those two episodes, BPCI-participating SNFs may have substituted HH visits for SNF days.

Exhibit 181: Impact of BPCI on Utilization Outcomes, by Clinical Episode, Model 3 SNF, Baseline to Intervention, Q4 2013 – Q3 2015



Note: The estimates in this table are the results of a difference-in-differences (DiD) model. These utilization measures are conditional upon the use of the service. Beneficiaries must have spent a minimum of one day in a SNF or HH setting during the 90-day post-discharge period (PDP) to be included in the DiD estimate for number of SNF days or HH visits, respectively.

**Indicates DiD estimates are statistically significant at the 5% level.

*Indicates DiD estimates are statistically significant at the 10% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Are participants shifting services outside of the episode period or increasing services not covered in the bundle, which may reduce overall savings to Medicare?

BPCI participants may attempt to reduce episode costs by changing the timing of services so that they are delivered before the initiating PAC stay or after the end of the episode period. Although these tactics could reduce episode payments, they would not necessarily achieve Medicare savings. Practically, however, the ability for Model 3 EIs to implement this approach may be limited. To shift services to the period before the episode is triggered by admission to PAC, the PAC EIs would need to have identified the patient in one of their clinical episodes. To shift services to the period after the end of the episode, the EI would need to track and influence the patient's health care services for an extended period. What we found was that payments in the 30-day pre-bundle period decreased in 8 of 11 clinical episodes (renal failure was significant at the 0.10 level), while payments in the 30-day post-bundle period increased in 8 of 11 clinical episodes (urinary tract infection was significant at the 0.10 level) (Exhibit 182).

Episode payments could also be lower if fewer services within the bundle definition are provided. However, if this results in higher payments for services excluded from the bundle, overall Medicare payments may not be lower. The standardized allowed amount for services not included in the bundle for 90-day episodes increased in 8 of 11 clinical episodes relative to the change in the comparison group. For three of these episodes (CHF, chronic obstructive pulmonary disease (COPD), and stroke) the difference was statistically significant ($p<0.10$) (Exhibit 182). The increases in non-bundled payments were predominately due to readmissions that are excluded from the bundle. Across all three clinical episodes with significant increases in non-bundled costs, unplanned readmissions that are included in the episode also increased during the 90 days after the qualifying inpatient stay relative to the comparison group (stroke was statistically significant at the 0.10 level) (Exhibit 183). These results suggest that for these episodes, the EIs may have difficulty controlling readmissions in general, both those related and unrelated to the episodes.

Exhibit 182: Impact of BPCI on Allowed Payment Outcomes outside of the bundle, by Clinical Episode, Model 3 SNF, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of episodes initiated Q4 2013 – Q3 2015	Standardized allowed amount not included in bundle definition[1]	Standardized allowed amount Part A & B, Days 1-30 pre-bundle period	Standardized allowed amount Part A & B, Days 1-30 PBP
Congestive Heart Failure	1,382	$285	$10	$268
COPD, Bronchitis, Asthma	547	**$1,161**	$126	$1,238
Hip & Femur Procedures Except Major Joint	633	-$239	-$401	$128
Major Joint Replacement of the Lower Extremity	2,321	$24	-$282	-$406
Medical Non-Infectious Orthopedic	1,180	$147	-$221	$125
Other Respiratory	431	$464	$849	$805
Renal Failure	886	-$231	-$1,411	$381
Sepsis	1,858	$26	-$499	-$8
Simple Pneumonia & Respiratory Infections	1,192	$336	-$570	-$71
Stroke	659	**$662**	-$157	$99
Urinary Tract Infection	931	-$116	-$285	$902

Note: The estimates in this table are the result of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. 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. PBP = post-bundle period.

[1] The standardized allowed not included in bundle definition values include 90-day episodes only.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Exhibit 183: Impact of BPCI on the Rate of Readmissions Excluded from the Bundle, Model 3 SNFs, Q4 2013-Q3 2015

Clinical episode	Number of Intervention Episodes		BPCI		Comparison		DiD Estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P value
Congestive Heart Failure	1,301	1,300	2.3%	3.4%	3.8%	2.9%	2.0	0.0	4.0	0.05
COPD	527	526	2.8%	5.2%	4.8%	2.4%	4.8*	1.2	8.3	0.01*
Stroke	640	633	3.2%	4.2%	5.1%	2.7%	**3.4**	0.3	6.5	0.03

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. DiD = Difference-in-difference; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; pp = percentage points.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

How has quality of care changed under BPCI?

There is mixed and inconclusive evidence about the impact of BPCI on the quality of care under Model 3 SNF EIs (Exhibit 184). There was a statistically significant decline in the readmission rate for MJRLE episodes relative to the comparison group (-2.7 percentage points, p=0.03). The readmission rate increased relative to the comparison group for renal failure (6.3 percentage points, p=0.06) and stroke (6.1 percentage points, p=0.08) clinical episodes. There was a statistically significant increase in emergency department use for sepsis episodes (3.5 percentage points, p=0.08). For MJRLE episodes, there was a statistically significant decline in the mortality rate (-1.3 percentage points, p=0.04). There were statistically significant relative increases in the mortality rate for COPD (7.3 percentage points, p=0.02); renal failure (5.1 percentage points, p=0.04); and simple pneumonia and respiratory infection (SPRI) (3.8 percentage points, p=0.09) episodes. Across non-significant results, no patterns were observed. It is important to keep in mind that these claim-based quality outcomes are based on a small number of EIs and patients.

The assessment-based quality outcomes among the approximately three-quarters of patients who remained in a SNF long enough to have two patient assessments, are generally positive, indicating that for this subset of patients with Model 3 SNF episodes, BPCI is associated with an improvement in ADLs. CHF episodes exhibited statistically significant increases in the proportion of episodes that improved in all three ADL measures relative to the comparison group ($p<0.10$). Other respiratory, renal failure, SPRI, medical non-infectious orthopedic, and urinary tract infection episodes had statistically significant increases in the proportion of episodes that improved in two of the three ADL measures ($p<0.10$) (Exhibit 185).

There was only one ADL measure that had a statistically significant relative decline in the proportion of episodes that improved – there was a 9.2 percentage point (p=0.06) decline in the proportion of BPCI MJRLE episodes with improved self-care function. The survey findings from MJRLE patients with Model 3 SNF episodes, described below, indicate that beneficiary-reported changes in functional status from before to after the episode were similar between BPCI and comparison patients.

The next report, which will include an additional year under BPCI and the experience of additional SNF EIs will provide important insights into the impact of BPCI on quality of care.

Exhibit 184: Impact of BPCI on Claim-based and Assessment-based Quality Outcomes, by Clinical Episode, Model 3 SNF, Baseline to Intervention, Q4 2013 - Q3 2015

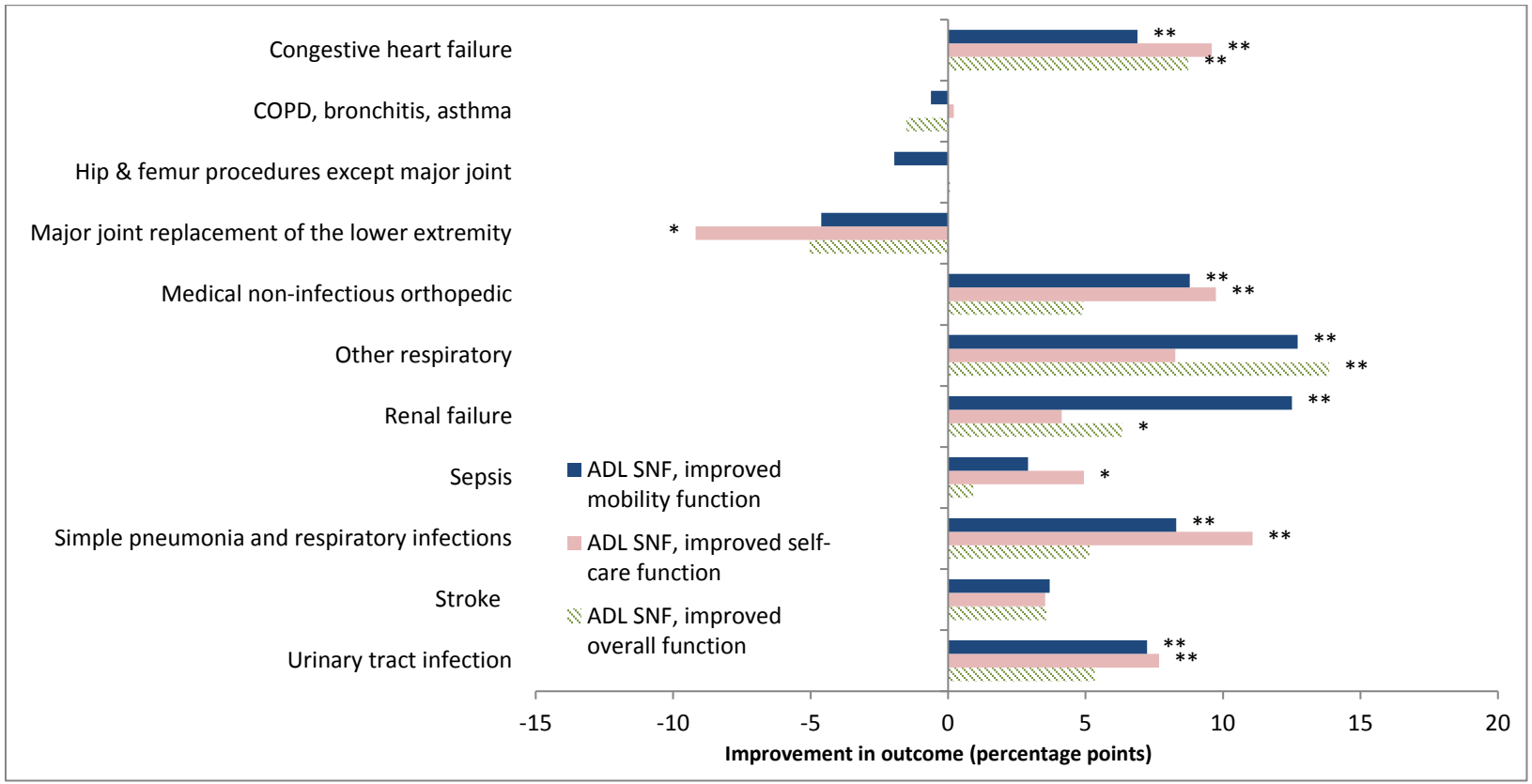
Clinical Episode	Number of episodes initiated Q4 2013 – Q3 2015	Unplanned readmission rate, 90 days from episode start date	Emergency Department Use, 90 days from episode start date	All-cause mortality rate, 90 days from episode start date	ADL, SNF, improved mobility function	ADL, SNF, improved overall function	ADL, SNF, improved self-care function
Congestive Heart Failure	1,382	1.8 pp	-0.4 pp	-2.6 pp	**6.9 pp**	**8.7 pp**	**9.6 pp**
COPD, Bronchitis, Asthma	547	5.1 pp	1.1 pp	7.3 pp	-0.6 pp	-1.5 pp	0.2 pp
Hip & Femur Procedures Except Major Joint	633	0.2 pp	-4.5 pp	0.5 pp	-2.0 pp	0.1 pp	0.6 pp
Major Joint Replacement of the Lower Extremity	2,321	**-2.7 pp**	-0.3 pp	**-1.3 pp**	-4.6 pp	-5.0 pp	-9.2 pp
Medical Non-Infectious Orthopedic	1,180	3.3 pp	3.2 pp*	-2.0 pp	**8.8 pp**	4.9 pp	**9.7 pp**
Other Respiratory	431	-5.2 pp	1.1 pp	3.6 pp	**12.7 pp**	13.9 pp	8.3 pp
Renal Failure	886	6.3 pp	-0.4 pp	5.1 pp	**12.5 pp**	6.3 pp	4.1 pp
Sepsis	1,858	-0.3 pp*	3.5 pp	-0.7 pp	2.9 pp	0.9 pp	4.9 pp
Simple Pneumonia & Respiratory Infections	1,192	3.6 pp*	0.1 pp	3.8 pp	**8.3 pp**	5.1 pp	**11.1 pp**
Stroke	659	6.1 pp	-0.3 pp	0.1 pp*	3.7 pp	3.6 pp	3.5 pp
Urinary Tract Infection	931	4.7 pp*	0.3 pp*	-1.1 pp	**7.2 pp**	5.3 pp	**7.7 pp**

Note: The estimates in this table are the result of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. PDP = post-discharge period.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate. Equal trends test was conducted for ED, readmission, and mortality outcomes.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Exhibit 185: Impact of BPCI on ADL Quality Outcomes, by Clinical Episode, Baseline to Intervention, Model 3 SNFs, Q4 2013 - Q3 2015



Note: The estimates in this table are the results of a difference-in-differences (DiD) model.

**Indicates DiD estimates are statistically significant at the 5% level.

*Indicates DiD estimates are statistically significant at the 10% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

b. BPCI-participating HHAs

Sample characteristics

Before discussing the impact of BPCI on payments, utilization, and quality, we present some basic statistics by clinical episode in Exhibit 186 to better understand the sample included in the analysis. By clinical episode, the number of matched HHA EIs ranged from 24 to 40; matched EIs initiated between 639 and 2,551 episodes over the first eight quarters of the initiative. Because providers were allowed to join BPCI over an extended period, these data represent an average of two quarters of participation. On average, 47% of a clinical episode's matched providers joined the initiative in Q3 2015 and 11 providers discontinued participation in the episode; nearly 90% of EIs with MJRLE episodes entered Phase 2 in Q3 2015. In the first two years of the initiative, a combined total of 34 EIs discontinued participation in CHF and simple pneumonia and respiratory infection.

Exhibit 186: Characteristics of the Matched BPCI Providers included in the DiD Estimates, Model 3 HHA, Q4 2013 – Q3 2015

Clinical Episodes	Matched EIs (#)	Matched Intervention Period Episodes (#)	Average length of participation (quarters)	EIs that joined in Q3 2015 (%)	EIs that joined in Q2 or Q3 2015 (%)	EIs that stopped participating in the episode (#)	Intervention Period Episodes from EIs that stopped participation (%)
Congestive heart failure	40	2,551	3	30%	48%	20	23%
Major joint replacement of lower extremity	24	639	1	88%	96%	0	0%
Simple pneumonia and respiratory infections	34	680	3	24%	50%	14	61%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2015 for BPCI providers. This table is limited to the matched BPCI providers used to calculate the difference-in-differences (DiD) results in the remainder of this section.

Key payment, utilization, and quality outcomes

How have the average standardized allowed amounts (Medicare payments and coinsurance/ copayments combined) changed under BPCI?

While payments decreased among BPCI-participating HHAs, few changes were statistically significant relative to the comparison group (Exhibit 187). The lack of many statistically significant payment outcomes could be in part due to the short tenure under BPCI, which averaged two quarters. For MJRLE, 88% of the HHA EIs had only one quarter in BPCI. The relative reductions in the total allowed amount and the HHA allowed amount across all three clinical episodes may suggest that with more time, HHA EIs may achieve statistically significant reductions in payments for these clinical episodes. For CHF episodes, the total amount included in the bundle definition for 90-day episodes decreased by $975 relative to the change in the comparison group (p=0.05). The decline in total payments for CHF episodes was driven by reductions in HHA, SNF, and readmission payments. The HHA standardized allowed amount for CHF episodes declined by $248 relative to the comparison group (p<0.01) (Exhibit 187).

Exhibit 187: Impact of BPCI on Allowed Payment Outcomes, by Clinical Episode, Model 3 HHA, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of episodes initiated Q4 2013 – Q3 2015	Total Amount Included in Bundle Definition[1]	HHA Standardized Allowed Amount, 90-day PDP[2]	SNF Standardized Allowed Amount, within the bundle[1,2]	Readmissions Standardized Allowed Amount, 90-day PDP[2]
Congestive Heart Failure	2,551	-$975	-$248	-$194	-$147
Major Joint Replacement of the Lower Extremity	639	-$957	-$184		
Simple Pneumonia & Respiratory Infections	680	-$403	-$112		$173

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. 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. Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. PDP = post-discharge period.

[1] The total amount included in bundle definition values are based on only the 90-day episodes.

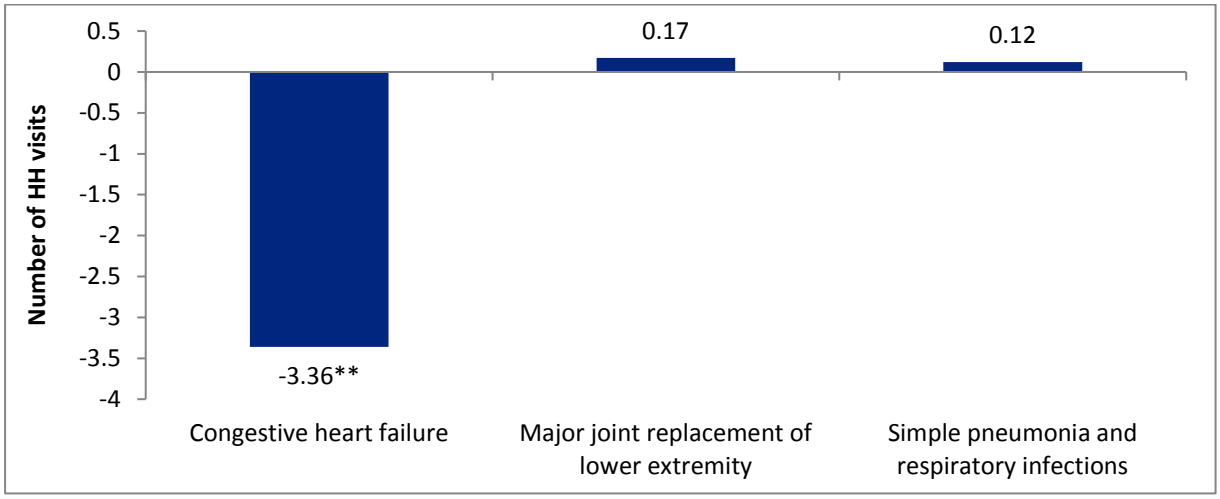
[2] These payment measures are not conditional upon use of the service.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How have the services changed under BPCI?

CHF episodes had a statistically significant decrease in the number of HH visits relative to the change in the comparison group (p=0.03) (Exhibit 188). Reducing the number of visits is the primary way for BPCI HHAs to reduce internal costs. Reducing home health visits will not necessarily reduce the Medicare FFS payment or contribute to NPRA because Medicare pays HHAs for a 60-day period. However, because the Medicare 60-day HHA payment depends in part on therapy categories that are dependent on the number of therapy visits, a reduction in therapy visits could account for the decline in CHF episode payments.

Exhibit 188: Impact of BPCI on Home Health Utilization, by Clinical Episode, Model 3 HHA, Baseline to Intervention, Q4 2013 - Q3 2015



Note: The estimates in this table are the results of a difference-in-differences (DiD) model.

**Indicates DiD estimates are statistically significant at the 5% level.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Are participants shifting services outside of the episode period or increasing services not covered in the bundle, which may reduce overall savings to Medicare?

Episode payments could be lower if fewer services within the bundle definition or within the bundle period are provided. This could reduce Medicare FFS payments for the episode, but may raise overall Medicare payments. Certain readmissions and services for specific conditions are excluded from the bundle, and the types of readmissions and conditions vary by clinical episode.[4] There was a statistically significant decline in payments for services provided during the 30 days after the CHF episode (p=0.03). There were no other changes in payments for services just before or after the episode period or for services excluded from the bundle (Exhibit 189).

[4] Centers for Medicare and Medicaid Services. (2016). Bundled Payments for Care Improvement Learning & Resources Area. Retrieved from https://innovation.cms.gov/initiatives/Bundled-Payments/learning-area.html

Exhibit 189: Impact of BPCI on Allowed Payment Outcomes Outside of the Bundle, by Clinical Episode, Model 3 HHA, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of episodes initiated Q4 2013 – Q3 2015	Total Amount not included in bundle definition[1]	Standardized allowed amount Part A & B, Days 1-30 pre-bundle period	Standardized allowed amount Part A & B, Days 1-30 PBP
Congestive Heart Failure	2,551	$17	-$216	-$756
Major Joint Replacement of the Lower Extremity	639	$24	-$161	$129
Simple Pneumonia & Respiratory Infections	680	-$363	$466	-$104

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark light orange shaded cells, respectively. 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. PBP = post-bundle period.

[1] The standardized allowed amount not included in bundle definition values include 90-day episodes only.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How has quality of care changed under BPCI?

Across clinical episodes, there were no patterns in readmissions rates, emergency department use, mortality rates, or the proportion of patients that improved in ADLs (Exhibit 190). There was a statistically significant increase in the proportion of BPCI patients with at least one emergency department visit during the first 90 days of the episode (CHF, p=0.04) and with improved bed transferring (MJRLE, p=0.10). We also found a statistically significant increase in the proportion of BPCI patients who had improved upper body dressing (CHF, p=0.06), although this may be a biased estimate. We rejected the null hypothesis that BPCI and the comparison group had parallel trends for this measure during the baseline period, which is required for an unbiased estimate.

Exhibit 190: Impact of BPCI on Claim-based and Assessment-based Quality Outcomes, by Clinical Episode, Model 3 HHA, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical episode	Number of episodes initiated Q4 2013 – Q3 2015	Unplanned readmission rate, 90 days from episode start date	Emergency Department Use, 90 days from episode start date	All-cause mortality rate, 90 days from episode start date	ADL, HHA, improved ambulation	ADL, HHA, improved bathing	ADL, HHA, improved bed transferring	ADL, HHA, improved lower body dressing	ADL, HHA, improved upper body dressing
Congestive Heart Failure	2,551	0.1 pp	**2.5 pp**	-1.7 pp	-0.9 pp	4.7 pp	2.2 pp	2.0 pp	5.4 pp*
Major Joint Replacement of the Lower Extremity	639	-2.1 pp	-3.2 pp		0.7 pp	0.7 pp	7.9 pp	-2.0 pp	1.2 pp
Simple Pneumonia & Respiratory Infections	680	0.7 pp	-4.8 pp	1.2 pp	-0.4 pp	-4.0 pp	0.4 pp	1.9 pp	1.8 pp

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. A blank cell indicates that the outcome cannot be presented, either due to insufficient sample size. PDP = post-discharge period.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate. Equal trends test was conducted for ED, readmission, mortality, and CHF ADL improved upper body dressing outcomes.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

2. Patient functional status, health status, and health care experience

Do patients treated under BPCI report differences in changes of functional status, health status, and health care experience compared with similar patients who were not in BPCI?

Although we have completed five waves of beneficiary surveys since the beginning of this evaluation, we only began sampling Model 3 at the clinical-episode level by EI type beginning in Wave 4.[5] In Waves 4 and 5, we collected enough survey responses to estimate results for respondents with a MJRLE episode initiated by Model 3 SNF EIs. Respondents were asked to report on their functional status, before and after the clinical episode, on overall mental and physical health, and on their health care experiences. We present risk adjusted results reflecting differences in outcomes between respondents receiving care from a BPCI SNF EI relative to a matched comparison group of respondents who received care in similar SNF providers that were not participating in BPCI. Estimates for the BPCI and the comparison group beneficiaries, along with confidence intervals for the estimated differences, are presented in **Appendix P.**

Exhibits 191-193 present point estimates of the differences in survey results between MJRLE respondents treated at Model 3 SNF EIs and respondents treated at matched comparison SNFs. Changes in functional status (Exhibit 191) are further classified as rates of improvement and rates of decline (as described above in Section II). Results across all measures are summarized below.

Overall, relative to comparison respondents, BPCI respondents with a MJRLE episode initiated by Model 3 SNF EIs reported similar changes in functional status from before to after the episode, and similar mental and physical health outcomes. However, BPCI respondents reported worse care experiences relative to comparison respondents for nine of ten measures, although only three were statistically significant ($p<0.05$).

a. Improvement and decline in functional status

BPCI respondents with a MJRLE episode initiated by Model 3 SNF EIs reported similar changes in functional status from before to after their MJRLE episode as the matched comparison group (Exhibit 191). In general, estimated differences in functional changes were small and statistically insignificant. The only exception was that BPCI respondents were 5.4 percentage points less likely to report decline in their ability to walk ($p=0.05$) relative to the comparison group.

b. Patient reported health status

BPCI respondents also reported similar mental and physical health outcomes relative to the matched comparison group (Exhibit 192). Estimated differences between BPCI and comparison respondents in depression, overall mental health, and physical health were small and statistically insignificant.

c. Health care experiences

BPCI respondents reported significantly worse care experiences for 3 of 10 health care experience measures (Exhibit 193). These included consideration of respondent preferences for PAC (-4.6 percentage points, $p=0.03$); medication instructions being clearly explained before

[5] We surveyed a sample of Model 3 BPCI and comparison beneficiaries across all BPCI-eligible clinical episodes and EI types in Wave 1 through Wave 3.

going home (-5.1 percentage points, p=0.01); and follow-up appointments or treatments being clearly explained before going home (-6.1 percentage points, p<0.01). Although BPCI respondents were less likely (-3.0 percentage points) to report satisfaction with their overall recovery, which was consistent with the generally negative estimates on care experience measures, the estimate on patient satisfaction was not statistically significant.[6]

[6] Subsequent surveys covering beneficiaries participating in the BPCI model in Wave 6 and Wave 7 of 2016 were analyzed together at the conclusion of Wave 7. Results indicated worse care experiences among BPCI respondents relative to comparison respondents for seven of ten measures, although only two differences were statistically significant. However, none of the three negative outcomes that were significant in Wave 5 remained significant in Wave 7. Moreover, unlike Wave 5, where BPCI respondents were 3.0 percentage points less likely to be satisfied with their overall recovery, BPCI respondents in Wave 7 were 3.0 percentage points more likely to be satisfied with their overall recovery. Thus, while there seems to be some persisting negative perception of care experience across waves, it is not specific to any particular outcome, nor does it have a continued impact on overall satisfaction.

Exhibit 191: Improvement and decline in functional status, Difference between BPCI and comparison survey respondents, Model 3 SNF EIs, 2015

Clinical episode	Wave	N BPCI responses*	Bathing, dressing, using the toilet, or eating		Planning regular tasks		Use of mobility device		Walking without rest		Using stairs		Physical/ emotional problems limit social activities		Pain limiting regular activities	
			Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.
MJRLE	5	407	1.37	-1.91	-1.22	0.79	0.46	3.22	2.31	**-5.36**	2.87	-2.59	2.31	-0.83	-1.12	-1.41

*The number of overall BPCI survey responses does not reflect item non-response (i.e., missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P.**

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* Dec., decline in functional outcomes or stay at the worst status; Imp., improvement in functional outcomes or stay at the best status; MJRLE, major joint replacement of the lower extremity.

Source: Lewin analysis of Model 3 BPCI and comparison beneficiary surveys collected in 2015.

Exhibit 192: Patient reported health outcomes, difference between BPCI and comparison survey respondents, Model 3 SNF EIs, 2015

Clinical episode	Wave	N BPCI responses*	Composite depression binary indicator†	Self-reported physical health binary indicator‡	Self-reported mental health binary indicator‡
MJRLE	5	407	-1.81	2.81	1.34

*The number of overall BPCI survey responses does not reflect item non-response (i.e., missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P.**

†The composite depression indicator is a binary measure equal to one when respondents reported a score of 3 or more on the Patient Health Questionnaire-2 (PHQ-2) and otherwise equals zero.

‡ The self-reported physical and health are binary measures equal one when respondents reported that their health was excellent, very good or good, and equal zero when respondents reported fair or poor health.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* MJRLE, major joint replacement of the lower extremity.

Source: Lewin analysis of Model 3 BPCI and comparison beneficiary surveys collected in 2015.

Exhibit 193: Health care experience, Difference between BPCI and comparison survey respondents, Model 3 SNF EIs, 2015

Clinical episode	Wave	N BPCI responses*	Never received Conflicting advice	Always received Appropriate care	Staff always used preferred language	Discharged at right time	Preferences considered for services after discharge†	Understand care of self†	Medication instructions clearly explained†	Follow-up explained†	Able to manage health†	Satis-faction with recovery ‡
MJRLE	5	407	-0.07	-2.85	-0.50	-2.01	**-4.58**	-0.70	**-5.09**	**-6.11**	1.64	-3.03

*The number of overall BPCI survey responses does not reflect item non-response (i.e., missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P.**

†Measure reflects that respondents either agreed or strongly agreed with this statement.

‡ Measure reflects that respondents were either "quite a bit satisfied" or "extremely satisfied" with their recovery.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* MJRLE, major joint replacement of the lower extremity.

Source: Lewin analysis of Model 3 BPCI and comparison beneficiary surveys collected in 2015.

3. Changes in Patient Mix

BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce episode payments below their target. Monitoring patient characteristics is particularly important under Model 3 because PAC providers can assess potential patients prior to admission and use this information in making admission decisions. While the impact analysis on payment, utilization, and quality presented in Section B.1 above controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix. In this section we present patient characteristics across Model 3 SNF and HHA episodes to determine whether there are any differences in patient case-mix between BPCI and comparison providers that might indicate patient selection. See Section II.D.2 for additional details on the DiD analysis.

There were indications from assessment-based measures that BPCI providers served healthier patients relative to the comparison group in two clinical episodes (SNF MJRLE and HHA CHF) and less healthy patients in SNF SPRI episodes. As presented earlier, there was a significant decline in payment and utilization and an improvement in quality outcomes for SNF MJRLE and HHA CHF and a significant decline in certain quality outcomes for SNF SPRI. To assess whether the results were due to changes in patient mix not accounted for by claim-based characteristics, we tested if the statistical significance and direction of the impact estimate changed after adding assessment-based measures to the risk adjustment model. We found that for most outcomes, even when the initial assessment-based measures were taken into account, BPCI had the same effect on the payment, utilization, and quality outcomes.

In the remaining clinical episodes, results were inconclusive in both claim- and assessment-based measures. However, the assessment-based functional status outcomes (e.g., measures on moving in bed, transferring, walking in room, and toileting) suggest that in 7 of the 12 clinical episodes, the BPCI-participating SNFs had a decrease in patients that required extensive assistance with these activities after joining BPCI relative to the change for the comparison group.

a. BPCI-participating SNFs

For the large majority of clinical episodes, there is no indication in claim- or assessment-based measures that providers preferentially selected healthier patients under BPCI. Claim-based patient characteristics suggest that 10 clinical episodes had either no significant results (3), indications of both healthier and less healthy patients (4), or at least one statistically significant result indicating less healthy patient mix (3) among BPCI participants relative to the comparison group. Two clinical episodes had changes toward healthier patients from baseline to intervention (non-fracture MJRLE and medical non-infectious orthopedic) but for no more than two claim-based measures (Exhibit 194).

The assessment-based results suggest that the patient mix changed toward healthier patients among BPCI SNF MJRLE participants (fractures and non-fractures) relative to comparison episodes, but the patient mix among BPCI SNF SPRI participants moved in the opposite direction (Exhibit 195). In MJRLE fracture episodes, BPCI patients had a relative decline in 18 of the 22 measures, of which six were statistically significant at the 5% level, indicating healthier patients. BPCI MJRLE fracture patients were less likely to need extensive assistance or be totally dependent moving in bed, transferring, walking in the room, and using the toilet, and were less

likely to frequently or always have urinary or bowel incontinence in the intervention period relative to the comparison group. For MJRLE non-fractures, all 22 outcomes were negative and 11 were statistically significant at the 10% level, suggesting a healthier patient mix (Exhibit 195). In SPRI, on the other hand, 20 of the 22 outcomes indicated a sicker patient mix in BPCI relative to the change for the comparison group, of which seven were statistically significant ($p<0.05$). BPCI patients in medical non-infectious and urinary tract infection episodes during the intervention period may have been sicker than patients in the baseline period relative to the comparison, though these patterns are not as clear. For the remaining strata, the large majority of measures had mixed or no statistically significant results (Exhibit 195).

Across all episodes, assessment-based functional status outcomes (e.g., measures on moving in bed, transferring, walking in room, and toileting) suggest that BPCI-participating SNFs were more likely to treat patients that less often required extensive assistance with these activities after joining BPCI relative to the change for the comparison group: 7 of the 12 clinical episodes had a decline in the proportion of patients who needed assistance in all of these functional status outcomes relative to the comparison group. Outside of MJRLE episodes, these declines were statistically significant for three measures ($p<0.10$). There were no observable patterns across assessment-based measures of patient intellectual impairment or active diagnoses of certain comorbidities. It should be noted, however, that functional data is more subjective than measures indicating the presence of comorbidities, so these results should be interpreted with caution.

Exhibit 194: Relative Changes of Claim-based Characteristics of BPCI and Comparison Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 3 SNF, Q4 2013 – Q3 2015

Clinical Episode	Number of Episodes, Q4 2013- Q3 2015	Demographics					Prior Health Status		Utilization Six Months Prior to Index Hospitalization									
		Age: 20-64 years	Age: 65- 79 years	Age: 80+ years	Female	Medicaid Eligibility	Disabled, no ESRD	Count of HCC indicators	Inpatient Acute Care Hospital	Emergency Room Admission	Home Health	Inpatient Rehabilitation Facility	Skilled Nursing Facility	Psychiatric Hospital	Long-term Care Hospital	Institutional Nursing Facility	No Institutional Care	No Post-acute Care
Congestive heart failure	1,382	2.0	2.4	-4.3	4.8	4.0	1.6	0.2	1.1	0.1	0.0	-0.1	1.0	-0.2	0.7	-0.3	-0.2	-0.5
COPD	547	1.8	4.2	-6.1	-0.7	0.0	3.3	0.3	3.3	-3.6	1.5	-0.4	6.7	0.3	-1.6	5.4	-1.8	0.9
Hip & femur procedures except major joint	633	-0.5	4.1	-3.6	3.5	3.0	-0.2	0.1	-2.3	-0.2	2.1	0.2	-2.4	0.1	-0.4	-4.4	1.7	1.1
Major joint replacement of lower extremity (fracture)	506	0.3	1.0	-1.3	2.8	3.1	1.2	0.0	-3.9	-3.3	-2.4	-0.1	-3.1	-0.4	-0.1	-3.3	5.1	0.4
Major joint replacement of lower extremity (non-fracture)	1,815	-1.7	4.2	-2.5	-1.6	-0.9	-1.2	0.0	-1.4	-1.7	-1.1	-0.2	-2.0	-0.1	0.1	-1.9	1.3	2.2
Medical non-infectious orthopedic	1,180	-0.9	3.0	-2.1	-0.6	1.0	-0.7	0.0	3.0	-4.3	-5.0	1.0	0.7	0.3	0.0	-1.6	-2.0	1.0
Other respiratory	431	-1.3	0.9	0.4	2.2	1.0	0.3	0.1	0.3	-7.1	3.9	1.5	0.1	-0.4	1.2	0.5	-0.6	3.2
Renal failure	886	2.5	-4.5	2.0	4.4	2.4	1.9	0.0	-1.6	-6.4	-0.7	-2.7	-4.7	-1.1	1.1	-5.2	3.3	3.1
Sepsis	1,858	1.9	0.7	-2.5	0.3	-0.9	2.4	0.3	5.7	-3.6	2.9	-0.2	2.6	0.1	-0.7	0.8	-4.4	-2.8
Simple pneumonia and respiratory infections	1,192	2.0	-1.7	-0.3	0.5	-0.8	2.2	0.1	4.2	-6.3	1.2	0.5	2.1	0.4	0.6	1.0	-2.7	0.4
Stroke	659	3.2	5.1	-8.3	-4.7	0.9	4.5	0.2	3.2	-1.4	4.6	0.5	-1.6	-0.9	-0.1	-3.1	-2.7	-1.7
Urinary tract infection	931	0.3	2.8	-3.1	-1.2	2.5	0.2	0.1	0.4	0.2	-3.0	0.8	5.4	0.0	1.0	1.2	-1.1	-2.2

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. DiD = difference-in-differences. SNF = skilled nursing facility. ESRD = end stage renal disease. HCC= Hierarchical Condition Category.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Exhibit 195: Relative Changes of Assessment-based Characteristics of BPCI and Comparison Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 3 SNF MDS, Q4 2013 – Q3 2015

Clinical Episode	Number of Episodes, Q4 2013- Q3 2015*	Not currently married	Intellectual or developmental disability (ID/DD)	Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	Sometimes, rarely, or never understands messages	Moderate to severe cognitive impairment	Rarely or never able to communicate	Moderate to severe depression	Rejected necessary evaluation or care at least once	Need extensive assistance or are totally dependent moving in bed	Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	Need extensive assistance or are totally dependent walking in room	Need extensive assistance or are totally dependent using the toilet	Do not believe they will improve enough to require less assistance with ADL in the future	Frequently or always had urinary incontinence	Frequently or always had bowel incontinence	Active diagnosis of heart failure	Active diagnosis of Alzheimer's	Active diagnosis of dementia	Active diagnosis of asthma, COPD, or chronic lung disease	Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	Experienced fracture due to a fall	Require special treatment[1]
Congestive heart failure	1,382	3.4	0.4	0.2	-0.5	-4.0	3.4	1.4	0.0	-4.0	-2.0	-1.6	-2.4	-4.8	1.3	-0.3	-1.7	0.1	2.9	-3.7	0.2	-1.1	0.4
COPD	547	2.5	-0.8	-0.8	-1.2	-7.2	0.1	-0.5	2.2	0.8	-0.3	1.5	2.1	-3.8	5.2	0.1	3.8	0.5	2.4	1.3	-4.2	2.3	-3.7
Hip & femur procedures except major joint	633	4.7	-0.9	0.2	-1.0	0.4	0.1	-4.9	0.4	-2.6	-0.4	1.6	-1.8	-2.7	0.6	1.1	0.6	-1.9	-2.8	-4.1	-1.9	0.6	1.4
Major joint replacement of lower extremity (fracture)	2,321	-4.1	-0.4	-2.7	-0.1	-5.5	0.1	-0.8	-0.9	-5.5	-8.4	-10.7	-8.4	0.8	-12.0	-12.3	-1.8	-1.7	-0.9	1.0	-2.3	2.0	-1.8
Major joint replacement of lower extremity (non-fracture)	1,180	-9.4	-0.1	0.0	-0.1	-2.1	-0.2	-0.7	-0.5	-11.4	-12.3	-6.5	-10.5	-4.3	-2.0	0.0	-1.4	-0.2	-1.9	-3.9	-0.4	-0.6	-3.5

Clinical Episode	Number of Episodes, Q4 2013- Q3 2015*	Not currently married	Intellectual or developmental disability (ID/DD)	Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	Sometimes, rarely, or never understands messages	Moderate to severe cognitive impairment	Rarely or never able to communicate	Moderate to severe depression	Rejected necessary evaluation or care at least once	Need extensive assistance or are totally dependent moving in bed	Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	Need extensive assistance or are totally dependent walking in room	Need extensive assistance or are totally dependent using the toilet	Do not believe they will improve enough to require less assistance with ADL in the future	Frequently or always had urinary incontinence	Frequently or always had bowel incontinence	Active diagnosis of heart failure	Active diagnosis of Alzheimer's	Active diagnosis of dementia	Active diagnosis of asthma, COPD, or chronic lung disease	Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	Experienced fracture due to a fall	Require special treatment[1]
Medical non-infectious orthopedic	431	1.1	0.0	-0.2	2.2	2.2	**3.5**	-0.3	2.2	-2.8	-3.6	-0.6	-3.6	0.7	2.8	2.3	0.1	-0.2	**5.7**	-2.9	-1.0	2.4	-0.9
Other respiratory	886	**9.5**	1.5	-1.0	1.4	1.0	2.9	**-4.6**	0.8	1.1	1.2	4.8	0.3	1.1	5.3	5.7	-3.9	2.0	5.4	-5.9	-3.0	-3.0	-2.2
Renal failure	1,858	2.6	0.6	0.0	-0.7	-1.1	2.3	0.2	0.6	-2.8	-2.2	-1.8	-2.9	1.4	0.4	4.9	-0.2	-2.2	-3.4	-1.5	-1.3	-1.8	-2.5
Sepsis	1,192	3.7	0.4	-0.3	1.5	-3.3	**4.5**	-0.9	1.1	-2.8	-0.5	-0.1	-3.0	**-4.6**	0.8	3.1	-2.0	0.7	1.1	0.6	-0.2	0.1	2.4
Simple pneumonia and respiratory infections	659	0.4	0.3	-0.9	**5.2**	**8.4**	**3.9**	0.0	1.7	2.0	**5.1**	2.7	3.1	0.5	**7.2**	**11.0**	0.9	2.1	**7.8**	0.9	1.7	0.3	0.7
Stroke	931	-3.3	-0.2	0.1	3.4	2.1	-2.5	0.2	-0.4	0.8	0.8	-3.1	0.3	-1.1	0.8	0.1	-0.7	-0.5	-1.1	-2.6	-2.4	-0.5	-0.2
Urinary tract infection	1,382	-2.6	0.5	0.4	0.7	1.0	2.4	1.3	0.6	-0.5	-1.0	-0.4	-1.4	1.2	4.2	**7.7**	3.8	-2.2	1.1	-1.1	1.0	-0.4	-1.9

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the **5%** or **10%** significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the **5%** or **10%** significance level are indicated by dark and light orange shaded cells, respectively). MDS = minimum data set.

*Assessment data was not available for all episodes.

[1]Examples include chemotherapy, oxygen therapy, and transfusions.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

b. BPCI-participating HHAs

There is some evidence that the change in CHF patients in BPCI-participating HHAs was to a healthier mix during the intervention period relative to the change in the comparison group. Although claim-based results show a statistically significant relative increase in the proportion of BPCI-participating HHA CHF patients who were eligible for Medicaid and disabled, the assessment-based results suggest a relative increase in the share of healthier patients (Exhibit 197). There were statistically significant declines in six measures – the proportion of BPCI CHF patients who had poor overall health status, were often or constantly confused, were short of breath from moderate to no exertion, required assistance with toileting, transferring, or ambulating ($p<0.10$) – all of which indicate healthier patients.

There is mixed evidence that BPCI-participating HHAs served healthier MJRLE patients during the intervention relative to the change in the comparison group. Claim-based results indicate a relative increase in the share of resource-intensive patients from the baseline to the intervention period as indicated by greater utilization of acute care in the six months prior to the index hospitalization and a higher count of Hierarchical Condition Categories (HCC) indicators ($p<0.05$) (Exhibit 196). Assessment-based results for MJRLE align with claim-based results for eight statistically significant outcomes, which suggest patients were less healthy ($p<0.10$). However, there were also three statistically significant results suggesting a relative increase in the share of healthier patients ($p<0.05$).

Finally, the results for SPRI patients were also mixed across both claim- and assessment-based measures. Claim-based outcomes indicate that BPCI patients were younger, more disabled, and had more IRF but less institutional nursing facility utilization in the six months prior to the index hospitalization. Assessment-based results seem to indicate a healthier patient mix, with statistically significant reductions in the proportion of patients with pre-existing conditions, poor overall status, poor ability to understand verbal content, and depressive symptoms ($p<0.10$), though there was also a statistically significant increase in the proportion of patients requiring assistance transferring ($p=0.04$).

Exhibit 196: Relative Changes of Claim-based Characteristics of Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 3 HHA, Q4 2013 – Q3 2015

Clinical Episode	Number of Episodes, Q4 2013- Q3 2015	Demographics					Prior Health Status		Utilization Six Months Prior to Index Hospitalization									
		Age: 20-64 years	Age: 65- 79 years	Age: 80+ years	Female	Medicaid Eligibility	Disabled, no ESRD	Count of HCC indicators	Inpatient Acute Care Hospital	Emergency Room Admission	Home Health	Inpatient Rehabilitation Facility	Skilled Nursing Facility	Psychiatric Hospital	Long-term Care Hospital	Institutional Nursing Facility	No Institutional Care	No Post-acute Care
Congestive heart failure	2,551	**2.3**	1.3	-3.6	1.2	**4.6**	**2.4**	0.0	0.7	-0.8	-2.9	0.4	-1.7	0.2	-0.1	-1.8	-0.8	2.8
Major joint replacement of lower extremity	639	0.0	-2.0	2.0	2.4	-1.2	0.5	**0.2**	**4.2**	0.9	0.4	-0.9	-0.3	0.6	0.2	-4.4	**-4.4**	-0.6
Simple pneumonia and respiratory infections	680	**4.4**	-3.1	-1.3	6.0	3.2	**4.7**	0.0	-0.9	0.3	0.5	1.4	0.0	0.1	**1.5**	-5.3	-0.5	0.8

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. DiD = difference-in-differences. HHA = home health agency. ESRD = end stage renal disease.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Exhibit 197: Relative Changes of Assessment-based Characteristics of Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 3 HHA Oasis, Q4 2013 – Q3 2015

Clinical Episode	Number of Episodes Q4 2013- Q3 2015*	Pre-existing condition(s)[1]	Poor overall status[2]	Severely impaired vision or hearing	Sometimes, rarely, or never understands verbal content	One or more stage 2 or higher unhealed pressure ulcers	Short of breath from moderate to no exertion[3]	Require use of bedside commode or are totally dependent in toileting	Require assistance transferring or are unable to transfer (e.g., from bed to wheelchair)	Require walker or more assistance ambulating	Dependent in maintaining self-care[4]	Dependent in ambulating[4]	Dependent in transferring[4]	Not likely to receive assistance in ADL[5]	Caregiver needs training to provide supervision and safety, is unlikely to provide help, or is not present[5]	Urinary incontinence	Bowel incontinence	Poor cognition[6]	Often or constant confusion	Depressive symptoms[7]
Congestive heart failure	2,551	3.1	-17.4	-0.1	0.1	-1.3	-5.3	-4.9	-24.3	-6.4	-1.4	-1.0	-1.5	1.6	-0.9	-0.3	0.0	-0.1	-2.1	0.2
Major joint replacement of lower extremity	639	9.7	1.1	1.1	0.5	0.3	8.0	-19.4	-17.1	0.3	1.4	1.2	1.0	-3.6	-3.7	2.5	1.4	0.7	2.0	2.1
Simple pneumonia and respiratory infections	680	-7.0	-6.6	-0.7	-5.7	-0.3	-0.1	4.4	7.2	-1.5	1.6	2.0	1.6	1.9	0.1	-2.7	1.3	-2.5	-4.5	-3.4

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the **5%** or **10%** significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the **5%** or **10%** significance level are indicated by dark and light orange shaded cells, respectively.

*Assessment data was not available for all episodes

[1] Urinary incontinence, indwelling/suprapubic catheter, intractable pain, impaired decision-making, disruptive or socially inappropriate behavior, or memory loss to the extent that supervision is required.

[2] Likely to remain in fragile health and have ongoing high risk(s) of serious complications and death, or has serious progressive conditions that could lead to death within a year.

[3] Moderate exertion includes dressing, using a commode or bedpan, and walking distances less than 20 feet.

[4] Prior to patient's most recent illness, exacerbation, or injury.

[5] Current and future willingness and ability of non-agency caregivers to provide assistance

[6] Requires at least some assistance and direction in specific situations (e.g., on all tasks involving shifting of attention), consistently requires low stimulus environment due to distractibility, or is totally dependent due to disturbances such as constant disorientation, coma, persistent vegetative state, or delirium.

[7] Had little interest in doing things or were feeling down, depressed, or hopeless more than half of the days in the last two weeks, or meets criteria for further evaluation for depression using a different standardized assessment.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

4. Market dynamics

Has the market share of BPCI-eligible episodes changed under BPCI?

For the Model 3 market dynamics analysis, we calculated the market share of Model 3 SNF and HHA EIs to determine whether EIs captured a greater share of BPCI-eligible episodes under the initiative. The numerator of the market share for SNF EIs is the number of BPCI-eligible episodes admitted to SNF EIs 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 to all SNFs in the EI's market. The market share for HHA EIs was calculated in the same manner. As discussed in Section II above, we defined the BPCI market for SNF and HHA EIs as the CBSA in which it was located. We removed EIs that had a 100% market share for a clinical episode in the baseline period (i.e., there were no other SNFs (HHAs) in the CBSA with any Medicare patients in that clinical episode), and those that did not have at least one episode in each of the relevant time periods. As we did for Model 2 EIs, we also identified five CBSAs that were too large to accurately define a local health care market (Chicago, New York, Detroit, Cincinnati, and Los Angeles) and we removed all Model 3 EIs located in these markets from the analysis of market share. As discussed above, the removal of the largest markets, and all their EIs, limits the generalizability of our findings, but still allows us to draw meaningful inferences. Since the EIs located in the five largest markets each had small baseline market share (due to the many competitors in these large CBSAs) it is difficult to detect meaningful change in their market shares over time. Including these EIs in the market share analyses would obfuscate the changes occurring in more typical markets that are well-defined by CBSA boundaries.

We calculated market shares separately for MJRLE, CHF, and sepsis episodes, since market shares may be considerably different for different clinical conditions. We tracked changes in the average market share from the baseline period (Q4 2011-Q3 2012) to the intervention period across SNF and HHA EIs. Because participants joined the initiative at different times, EIs were stratified by the calendar quarter in which they began participating in the clinical episode. We segmented the baseline period and each EI's BPCI participation period into six-month intervals using data through Q3 2015, and then calculated the change in each measure from the baseline period to each six-month BPCI participation period, to observe changes over time. The analyses focus on the largest SNF and HHA EI cohorts for each clinical episode for which we had at least six months of complete claims data in the BPCI participation period at the time of this analysis. For SNF EIs, we focused on EIs that started participating in MJRLE, CHF, and sepsis episodes in Q1 2014 and Q2 2015. For HHA EIs, we focused on EIs that started participating in CHF episodes in Q1 2014, Q2 2014, and Q2 2015. Some cohorts include a relatively small number of EIs, and our exclusion restrictions further reduce the number of EIs that are analyzed in each cohort, so conclusions should be interpreted with caution.

The changes in market share that we observed do not suggest that Model 3 EIs increased their market share of BPCI-eligible episodes over time (Exhibit 198). Market shares for Model 3 SNF EIs that started participating in BPCI in Q1 2014 tended to decline over the BPCI participation periods compared with the baseline, for all three episodes. Although the percentage point change in average market share was small for any BPCI participation period, the magnitudes were sometimes fairly large relative to the average market share at baseline. For the Q2 2015 cohort of EIs, the market share changed little from baseline to the BPCI participation periods for any of the three clinical episodes.

Exhibit 198 also includes results for the average change in market share for the Q1 2014, Q2 2014, and Q2 2015 cohorts of Model 3 HHA EIs. The CHF market share for the Q1 2014 cohort of HHA EIs declined in the BPCI participation period, while the market share of the Q2 2014 cohort of HHA EIs increased. For HHA EIs that started participating in CHF in Q2 2015, market shares declined in the first six months after starting BPCI.

Exhibit 198: Average Change in Market Share of MJRLE, CHF, and Sepsis Episodes for Model 3 SNF EIs, and CHF Episodes for Model 3 HHA EIs

Clinical Episode	Statistic	Q1 2014 Cohort				Q2 2014 Cohort				Q2 2015 Cohort	
		Average Market Share in Baseline	Average Percentage Point Change from Baseline ‡			Average Market Share in Baseline	Average Change from Baseline‡			Average Market Share in Baseline	Average Change from Baseline‡
			Q1 2014 / Q2 2014	Q3 2014 / Q4 2014	Q1 2015 / Q2 2015		Q2 2014/ Q3 2014	Q4 2014/ Q1 2015	Q2 2015/ Q3 2015		Q2 2015 / Q3 2015
MJRLE*	N (SNF)	12	12	12	12	-	-	-	-	75	75
	Mean (SNF)	4.17	-0.65	-0.22	-0.18	-	-	-	-	9.73	0.09
CHF**	N (SNF)	23	23	23	23	-	-	-	-	46	46
	Mean (SNF)	4.36	-0.7	-0.97	-1.14	-	-	-	-	12.76	-0.15
	N (HHA)	13	13	13	13	8	8	8	8	7	7
	Mean (HHA)	15.35	-1.07	-3.53	-0.69	1.86	0.18	0.74	0.37	23.79	-4.82
Sepsis†	N (SNF)	23	23	23	23	-	-	-	-	44	44
	Mean (SNF)	4.99	-0.84	-0.61	-1.04	-	-	-	-	10.69	0.25

* *MJRLE*: 21 SNF EIs in the Q2 2015 cohort were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. Six more SNF EIs in the Q1 2014 cohort were excluded because they did not have at least one episode in all baseline and BPCI participation periods.

** *CHF*: 22 SNF EIs and one HHA EI in the Q1 2014 cohort, and seven SNF EIs in the Q2 2015 cohort were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. Two HHA EIs in the Q1 2014 cohort were excluded because they were not located in a CBSA. One SNF EI in the Q2 2015 cohort had 100% market share at baseline and was excluded. Nine more SNF EIs and one more HHA EI in the Q1 2014, and 15 more SNF EIs in the Q2 2015 cohort, were excluded because they did not have at least one episode in all baseline and BPCI participation periods.

† *Sepsis*: 22 SNF EIs in the Q1 2014 cohort and eight SNF EIs in the Q2 2015 were located in Chicago, New York, Detroit, Cincinnati, or Los Angeles, and were excluded from the analysis. Two SNF EIs in the Q2 2015 cohort was excluded because they did not have at least one episode in all baseline and BPCI participation periods.

Source: Lewin analysis of Medicare claims data for discharges from Q4 2011 through Q2 2015 for BPCI and non-BPCI providers within CBSAs that have at least two BPCI providers and meet our market inclusion criteria.

C. Model 3 discussion

1. Participants

There were 873 SNF, 144 PGP, 116 HHA, 9 IRF, and 1 LTCH episode initiators (EIs) under 136 Awardees that were active in Model 3 during the first two years of the initiative. These participants initiated over 35,000 episodes of care during that period, or approximately 12% of all Model 2, 3, and 4 episodes.

The SNFs and HHAs that participated in BPCI had higher Medicare Part A payments for the 90-day episode prior to the announcement of BPCI than providers that did not participate in BPCI. In addition, participating SNFs and HHAs tended to be larger, for-profit, and part of a chain. Combined, these characteristics indicate that the SNFs and HHAs that chose to participate may have had more resources to devote to the initiative and their higher episode payments may have made it easier for them to reduce episode payments and receive positive net payment reconciliation amount (NPRA) than those that did not participate.

Similar to Model 2 participants, Model 3 interviewees reported that they joined BPCI for the opportunity to learn how to use financial incentives to drive change, manage care beyond discharge, and improve care transitions. Interviewees described bundled payments as the future of payment reform. According to Awardee-reported data, only 3% of SNFs and 1% of HHAs had prior experience with bundled payments. Their lack of prior experience may have been a factor in the large proportion that joined BPCI under a Convener. Over 85% of the Model 3 SNF and HHA EIs joined BPCI under a Convener, compared to 62% of Model 2 hospitals. Participants we interviewed indicated that the Convener shared financial risk and provided administrative and analytic support.

Interviews with participating providers revealed the importance of their relationships with hospitals, particularly in competitive PAC markets. With the expectation of reducing the SNF length of stay and HHA visits under BPCI, they noted a greater reliance on hospitals for more patient referrals.

Key implementation challenges included difficulties coordinating care with other providers and changing the mindsets of patients. Model 3 participants found that referring hospitals, especially those outside of their market, were often unwilling to collaborate. In particular, they often found it difficult to obtain data from referring hospitals to determine if the patient was in one of their BPCI clinical episodes. Interviewees also noted difficulties in changing the mindset of patients who felt entitled to 21 days in the SNF, even if it was not medically necessary. Coordinating with other providers and resetting patient expectations were also key implementation challenges for Model 2 participants. Also similar to Model 2, a small share of participants used the various waivers.

A number of participants stopped participating in some or all of their BPCI clinical episodes. About 15% of HHA and 5% of SNF EIs withdrew completely from Model 3 during the first two years of the initiative. Awardees cited a lack of internal capacity to support the initiative as a reason for withdrawing. Financial reasons were also noted, such as receiving an insufficient volume of patients or not realizing positive NPRA.

Although there were a large number of SNF and HHA EIs in the first two years of the initiative, we had sufficient sample size to conduct impact analyses on only 11 SNF and 3 HHA clinical episodes.

2. Impact of BPCI on Costs and Utilization

Only two of the 14 clinical episodes analyzed, SNF major joint replacement of the lower extremity (MJRLE) and HHA congestive heart failure (CHF), achieved statistically significant reductions in episode payments relative to the comparison group. For BPCI-participating SNFs in particular, one reason may have been that the reduction in SNF payments was offset by an increase in readmission payments for some clinical episodes.

Also, as with Model 2 participants, the lack of payment effects for the majority of Model 3 clinical episodes could be due to the short average time in BPCI. Model 3 results reflect an average of three quarters under BPCI. Finally, few of these providers had any prior experience with bundled payment models.

Although lacking statistical significance in most clinical episodes, SNFs appear to have reduced the services that were within their control - the number of days in a SNF. The average number of days in a SNF declined in 9 of 11 (81%) clinical episodes relative to the comparison group (although only 2 were statistically significant). Fewer SNF days is consistent with the observed reduction in total SNF payments. However, reducing SNF days may be a risky strategy for BPCI-participating SNFs because Medicare pays them a daily rate. As a result, fewer SNF days reduces their Medicare revenue. In addition, the foregone SNF days would be those at the end of the stay, which probably are the lowest cost days for the SNF to provide. Therefore, to determine the most advantageous strategy in response to BPCI, a SNF would need to assess the relationship among the reduced Medicare per diem payments, potential NPRA, and changes in operating costs. The SNF representatives we have interviewed have not provided us with information on how they make this tradeoff.

Among Model 3 participating HHAs we have discerned no pattern across the three clinical episodes with regard to changes in the number of HH visits, although there was a statistically significant decline in HH visits for CHF episodes ($p=0.03$). Because HHAs generally are paid for a 60-day episode of care, reducing the number of visits reduces their internal costs, but it does not always lower Medicare payments and, therefore, may not contribute to achieving NPRA.

3. Impact of BPCI on Quality

Various data sources were used to gather evidence on the impact of BPCI on the quality of care delivered to patients. Among HHAs, BPCI did not appear to have a systematic effect, either positive or negative, on the quality of care delivered. However, across the results for BPCI-participating SNFs, there were some concerning declines in quality of care, though these were accompanied by improvements in several ADL measures among the subgroup of patients who remained in the SNF or received care from a HHA long enough to complete two patient assessments.

With the exception of SNF MJRLE episodes, all significant changes in mortality, emergency department use and readmissions pointed toward declines in quality for BPCI SNF participants relative to the comparison group. There were increases in 90-day readmission rates for renal failure and stroke episodes; 90-day emergency department use for sepsis; and 90-day mortality for COPD, renal failure, and simple pneumonia and respiratory infections (SPRI) episodes relative to the comparison group. Mortality increased from baseline to intervention for BPCI SNF COPD patients, but declined for BPCI SNF SPRI patients from baseline to intervention. On the other hand, even with the reduction in Medicare payments and SNF days for SNF MJRLE episodes, BPCI SNFs were able to significantly reduce mortality and unplanned readmission rates relative to comparison providers.

Among SNF-initiated episodes, the assessment-based quality outcomes, which are based on the approximately three-quarters of patients who remain in the SNF long enough to have two patient assessments, provide a picture of improved quality in BPCI SNFs. All three ADL measures improved for CHF, and two of the three measures improved for another five clinical episodes. The only statistically significant decline was in the proportion who had improved self-care among SNF MJRLE episodes. For HHA-initiated episodes, there were two statistically significant ADL outcomes. There were increases in the proportion of episodes with improved bed transferring for MJRLE episodes and improved upper body dressing for CHF episodes. However, the increase in the proportion of BPCI HHA CHF patients with improved upper body dressing may be biased because we rejected the null hypothesis that BPCI and the comparison group had parallel trends in the baseline period. As a whole, the ADL results suggest that BPCI is associated with an improvement in ADLs among patients that receive SNF or HHA care long enough to have two assessments. However, these patients may not be representative of all patients discharged to PAC due to the reasons they do not have two patient assessments. Patients were not included in the assessment-based quality outcomes if the patient was readmitted to the hospital, died, had little or no cognitive function, or if the episode occurred later during our measurement period and therefore a second assessment was not yet available in the data.

Surveys of SNF MJRLE patients gathered information on reported changes in functional status, health status, and care experience. BPCI respondents with a MJRLE episode initiated by Model 3 SNF EIs reported similar changes in functional status from before to after the episode, similar mental and physical health outcomes, but worse care experiences on three of ten measures of health care experiences.

It is unclear what may be driving these statistically significant changes in quality outcomes. The current results are based on a small number of EIs and patient episodes. In addition, there were large differences in the baseline period rates for many of the clinical episodes, despite our attempts to find a good comparison group. These underlying baseline differences contributed to these findings and indicate the need for further analysis. We will monitor if these outcomes are consistent over time and will include them in the next annual report with an additional year of experience.

4. Unintended effects

There was a decline in the proportion of SNF and HHA patients that needed extensive assistance with ADLs, based on the initial patient assessment. This may be an indication that SNFs and HHAs had healthier patients under BPCI. SNFs and HHAs can decide whether or not to admit a

patient and they may have access to clinical information that can help them in selecting a healthier, and therefore less costly, mix of patients. This would be an important issue under BPCI because the target amount is based on historical episode spending and it is not adjusted for changes in patient mix. As a result, a SNF or HHA with less costly patients in the intervention period than in the baseline could have episode costs below the target and earn NPRA without making any care redesign changes.

Patients in MJRLE episodes (fractures and non-fractures) initiated at BPCI-participating SNFs had better functional status upon admission to the SNF during the intervention than during the baseline period relative to the change for the comparison group. It is notable that this is a surgical episode because it would be more likely that patient selection occur for surgical episodes. It would be easier for a SNF to identify a surgical patient as a BPCI episode because the patient's MS-DRG assignment (and thus BPCI participation) would be less ambiguous than for patients with medical conditions. Claim- and assessment-based characteristics also suggest that the BPCI-participating HHAs' CHF patient population became significantly healthier from baseline to intervention relative to the comparison group.

The evidence about potential patient selection was not consistent across episodes, however. There was a significant increase in the severity of the patients for SNF SPRI episodes, based on evidence from demographic, prior health care use, and functional status upon admission to the SNF. If this decline was associated with higher costs, then the SNFs would have more difficulty reducing Medicare payments below the target amount. This could lead to BPCI-participating SNFs declining to admit these patients, leading to possible access problems, if SNFs determine that they cannot achieve savings for these episodes. In the remaining episodes, we observed no consistent pattern of changes in patient characteristics.

Our analyses of the impact of BPCI on payment, utilization, and quality use patient characteristics from the claims data to control for patient differences between BPCI and the comparison group. The patient assessment data provide additional information about the functional status of patients upon admission to the post-acute care setting. Although we cannot use patient assessment data in the risk adjustment across all Model 3 episodes due to missing data and in some instances, challenges with degrees of freedom and perfect prediction, we did use it to further explore the impact of changes in patient mix for some clinical episodes. We conducted sensitivity analyses to determine if the changes observed in the patient assessment data for SNF MJRLE, SNF SPRI, and HHA CHF explain the significant BPCI impacts on payment and quality for these clinical episodes. We found that for most outcomes, even after adjusting for functional status and other characteristics present at admission, the impact estimates remained statistically significant. This adds confidence to the findings that BPCI had an impact on key outcomes as presented above.

5. Summary

Only SNF MJRLE and HHA CHF episodes achieved statistically significant reductions in Medicare payments relative to the comparison group. For both of these strata, the BPCI providers decreased the services that were within their control – SNF days and HHA visits, respectively. At the same time, the functional status and other characteristics upon admission to the SNF or HHA indicated that SNFs were admitting healthier MJRLE patients and HHAs were admitting healthier CHF patients in the intervention period than the baseline, relative to the

comparison group. Additional analysis, however, suggests that even when these functional status and other characteristics present at admission were taken into account, the outcomes remain statistically significant.

D. Clinical episode issue brief chapters

1. Congestive Heart Failure, HHA

Medicare beneficiaries received 2,703 congestive heart failure (CHF) episodes of care from 50 BPCI-participating home health agencies (HHAs) in the first eight quarters of the initiative (from October 2013 through September 2015).[7] Total standardized Part A and B payments during 90-day episodes were $975 lower for episodes initiated at BPCI-participating HHAs relative to matched comparison HHAs from the baseline to the intervention period (p=0.05). The decline was due to a relative decrease in skilled nursing facility (SNF) and HHA payments. BPCI did not have an adverse effect on quality as measured by claim-based outcomes and there was a significant relative improvement in one measure of functional status outcomes among the patients who received two patient assessments.[8] Although there were no relative changes in service use in the six months before the episode, the patient population did shift such that patients in BPCI-participating HHAs were relatively younger and differed on a range of characteristics that may indicate that they were less costly to treat at the start of their episodes compared with patients in those HHAs during the baseline. Additional analysis, however, suggests that even when these functional status and other characteristics present at admission were taken into account, payments and quality outcomes still decline significantly.

a. Participants

During the first eight quarters of the BPCI initiative, 50 HHAs (43% of all Model 3 BPCI HHAs) participated in the CHF episode and 45 of them chose 90-day episodes (90%). Two HHAs chose 30-day episodes and three chose 60-day CHF episodes. There were 2,703 CHF episodes initiated by these providers during the BPCI initiative, comprising approximately 38% of all HHA episodes across the 48 clinical episodes. Because participants were allowed to join BPCI over an extended period, the CHF results are based on an average of three quarters of experience. Approximately 32% of the HHAs began participating in CHF in Q3 2015 and 46% joined in either Q2 or Q3 2015. Furthermore, of the 50 participating HHAs, 26 (52%) stopped participating in the CHF episode within the first eight quarters of the initiative. This is the highest episode drop rate of any BPCI episode. The HHAs that dropped this episode did so after participating for at least three quarters, and contributed 23% of all CHF episodes. Results from all HHAs that stopped participating in the episode are included in our analysis (other than six for which baseline volume was too small).

There is a possibility that less successful HHAs were more likely to terminate their participation, which could increase the impact of the successful HHAs on the overall results (i.e., a provider selection effect). One BPCI interviewee[9] that terminated this episode attributed the decision to the financial risk associated with readmissions, which can be frequent for CHF patients. Even though

[7] Because participants were allowed to join BPCI over an extended period, these participants had an average of 3 quarters of experience in Model 3.

[8] This estimate may be biased because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

[9] See Methods section for more information on the BPCI participant site visits and interviews.

the target amount accounts for the historical rate of readmissions, participants would face financial risk if readmissions increased. They may also have been concerned about volatility in readmission rates, which for low-volume facilities in particular, may pose financial risk. Interviewees also reported difficulties identifying patients with CHF episodes due to delays in getting the hospital-assigned Medicare Severity-Diagnosis Related Group (MS-DRG) because patients with CHF frequently had other diagnoses that could affect MS-DRG assignment.

Compared to HHAs *not* participating in BPCI, HHAs that participated in the CHF episode were more likely to be for-profit, not government-owned, and part of a chain (see Exhibits 199a & 199b). Additionally, BPCI-participating HHAs were more likely to have more nurses and more CHF discharges prior to joining BPCI. Indeed, several interviewees noted that they chose this episode because they had a relatively large volume of CHF patients.

Matched comparison HHAs were identified for 40 of the 50 BPCI-participating HHAs in the sample.[10] While this is a relatively low match rate, it is unlikely to distort the results, because the unmatched HHAs were all excluded due to low CHF volume (less than 3% of the total CHF episodes).[11] Of the 10 that were excluded, six also dropped the episode; 20 of the 40 HHAs included in subsequent analyses dropped the episode during the first eight quarters of BPCI. As one BPCI HHA is much larger and different than other HHAs on several important characteristics, we identified its comparison providers separately from the other participants and excluded it from the Matched CHF BPCI HHAs column in Exhibits 199a and 199b. This outlier provider contributed 49% of all Model 3 HHA CHF episodes in the baseline and intervention period.[12] The matched BPCI-participating HHAs were not significantly different from the comparison HHAs in any measures except that they were more likely to be part of a chain.

[10] Mahalanobis distance matching was used to match each BPCI-participating HHA to comparison HHAs. There were 10 HHAs that were not included in the analysis because they had fewer than five relevant discharges in 2011 and 2012.

[11] The outlier provider is excluded from the matched group (although included in subsequent analyses) but is shown in the "All CHF BPCI HHAs" column in Exhibits 199a & 199b.

[12] This provider later withdrew, but not during the eight quarters covered in this report.

Exhibits 199a & 199b: Characteristics of BPCI-participating HHAs with Congestive Heart Failure Episodes, Compared with Non-participating HHAs, All BPCI-participating HHAs, and Comparison HHAs, Model 3, Q4 2013 – Q3 2015

Characteristic		All CHF BPCI HHAs (N=50)		Non-participating HHAs (N=9,769)		Matched CHF BPCI HHAs (N=39‡)		Comparison HHAs (N=585+)	
		N	%	N	%	N	%	N	%
Ownership	Non-Profit	7	14%	1,699	17%	6	15%	108	18%
	Government	0	0%	612	6%*	0	0%	0	0%
	For Profit	43	86%	7,458	76%*	33	85%	477	82%
Urban/Rural	Urban	41	82%	7,883	81%	32	82%	492	84%
Part of Chain	Yes	36	72%	3,110	32%*	28	72%	254	43%*

Characteristic	All CHF BPCI HHAs (N=50)	Non-participating HHAs (N=9,769)	Matched CHF BPCI HHAs (N=39‡)	Comparison HHAs (N=585+)
	Mean	Mean	Mean	Mean
Number of Employed Nurses in HHA	51	9*	23	20
Number of CHF HHA Episodes of Care, 2011	52	9*	33	31
Standardized Part A Allowed Payment 90 days from start of receiving home health, CHF, 2011	**$11,109**	**$11,645**	**$11,856**	**$11,764**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating HHAs are defined as Model 3 EIs, Q4 2013 – Q3 2015. Non-participating HHAs are all other HHAs.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" HHAs are compared to "All CHF BPCI" HHAs. "Comparison" HHAs are compared to "Matched CHF BPCI" HHAs.

‡ The outlier provider and its matched comparison providers are not included in this column. However, it is included in the "All CHF BPCI HHAs" column, where it has the effect of skewing the number of employed nurses and episodes of care.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI HHAs. The unique number of matched comparison providers is 454.

b. Medicare Payments

Total Part A and B standardized allowed payments (2015$) included in the bundle for 90-day episodes decreased by $975[13] (p = 0.05) for BPCI relative to the comparison group (see Exhibit 200).[14] Total payments during the episode declined by 6% between the baseline and intervention period for BPCI HHAs, while they increased by 1% for comparison HHAs.

[13] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

[14] The outlier provider, which contributed 70% of episodes in the intervention period after exclusions, had comparable unadjusted results to other BPCI providers.

There were significant reductions in total Part A and B payments in the 1-30 day post-bundle period for episodes initiated through Q3 2015 ($p = 0.03$),[15] and in the 1-90 ($p<0.01$), 1-120 ($p<0.01$), and 1-180 ($p = 0.01$) day post-bundle periods for episodes initiated through Q2 2015.[16] Although an influx of 16 Awardees in Q3 2015 resulted in 13% more episodes for the 1-30 day measure compared to the 1-90 day measure, there is a consistent effect of lower payments during the post-episode period. One explanation is that BPCI-participating HHAs may have improved some aspect of care that has enabled their CHF patients to reduce spending on post-episode care. Another explanation may be that BPCI HHAs may have had more potential to reduce costs, given that post-episode spending during the baseline period was higher for BPCI than the comparison group.

[15] The outlier provider had less of a reduction on this measure, unadjusted.
[16] There was not sufficient claims run out time to include Q3 2015 episodes in all the post bundle outcomes.

Exhibit 200: Impact of BPCI on Total Part A and B Standardized Allowed Payments (2015$) for Congestive Heart Failure Episodes, Model 3 Home Health Agencies, Q4 2011 - Q3 2015

Total Allowed Amount	Number of intervention episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Total payment included in bundle definition, 90-day episodes	2,414	2,476	$15,999	$15,155	$14,989	$15,119	-$975	-$1,958	$8	0.05
Total payment, post bundle, days 1-30	2,057	2,122	$5,063	$4,272	$4,698	$4,663	-$756	-$1,418	-$95	0.03
Total payment, post bundle, days 1-90	1,825	1,908	$14,432	$12,761	$13,713	$13,949	-$1,906	-$3,305	-$508	<0.01
Total payment, post bundle, days 1-120	1,511	1,598	$18,600	$16,092	$17,755	$17,714	-$2,467	-$4,319	-$615	<0.01
Total payment, post bundle, days 1-180	1,490	1,590	$25,980	$22,922	$24,881	$24,789	-$2,966	-$5,262	-$669	0.01

DiD = difference-in-differences estimate; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; PDP = post-discharge period.

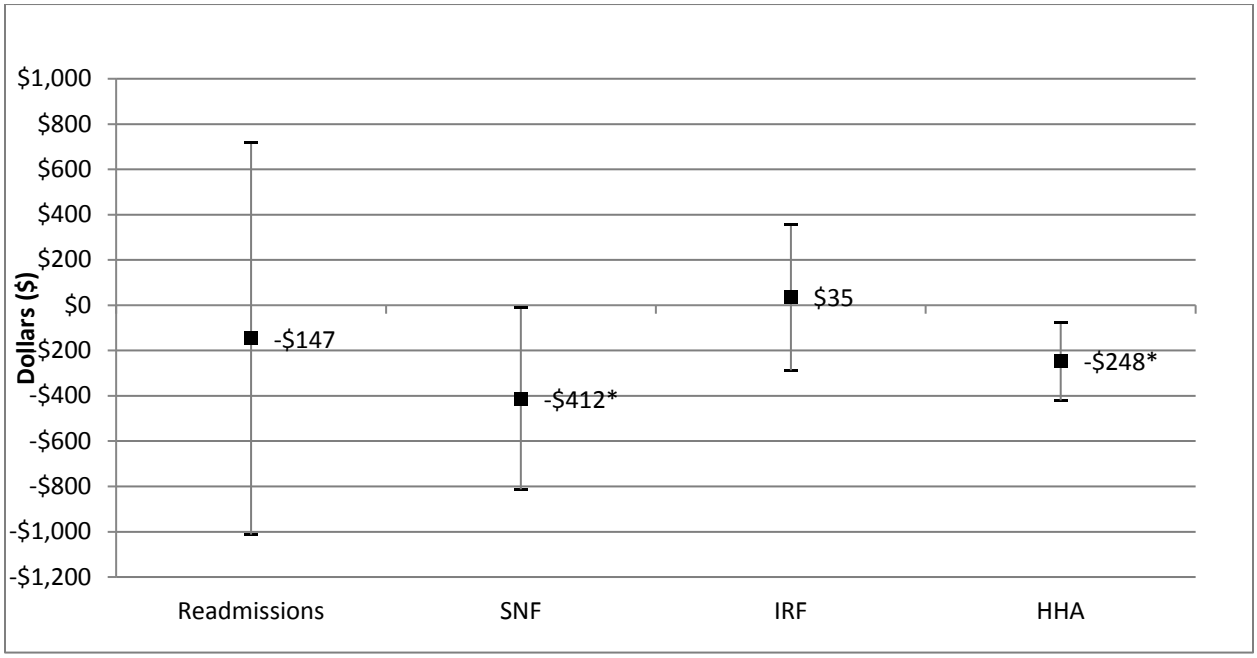
These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015 for the total payment included in the bundle definition, 90–day episodes and total payment, post bundle, l day 1-30 outcomes. The intervention period for total payment, post bundle, days 1-90, 1-120, and 1-180 is defined as Q4 2013 through Q2 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

The reductions in total payments were due to changes in post-acute care (PAC) utilization. There were relative decreases in SNF (-$412, p = 0.04) and HHA payments (-$248, p<0.01), while changes in payments for readmissions and inpatient rehabilitation facilities (IRF) were not significant (see Exhibit 201).

Exhibit 201: Difference-in-Differences Estimates and 95% Confidence Intervals for Medicare Payments (2015$) during the 90-day Post-discharge Period by Care Setting for Model 3 Home Health Agency episodes



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.
* Indicates statistical significance at the 5% level

BPCI-participating HHAs also reduced the number of visits by 3.4 visits per patient, relative to the comparison providers (p = 0.03). While this reduction would not lower the Medicare payments (unless there was a reduction in the number of therapy visits for patients receiving 4 or more visits) because HHAs receive an episode payment, it could generate internal cost savings for the HHA.

c. *Quality of Care*

The impact of BPCI on quality of care was mixed. The proportion of BPCI patients with an emergency department (ED) visit within the 90-day post-discharge period (PDP) rose relative to comparison patients (2.5 percentage points, p = 0.04), which was due to a reduction in the share of patients with an ED visit between the baseline and intervention period for the comparison group and an increase in the BPCI group.[17] However, there was no relative change in ED visits during the first 30 days of the episode, in readmissions within 30 or 90 days, or in mortality rates within 30 or 90 days. There was a significant relative improvement in the proportion of patients with improved upper body dressing (5.4 percentage points, p=0.06). However, this estimate may

[17] The outlier provider had a smaller decline in the unadjusted 90-day ED rate than other BPCI providers.

be biased as we rejected the null hypothesis that BPCI and the comparison group had parallel trends in the baseline period.

d. Change in Patient Mix

We monitor changes in patient characteristics because BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce their episode payments below their target. Monitoring patient characteristics is particularly important under Model 3 because PAC providers can assess potential patients prior to admission and use this information in making admission decisions. While the impact analysis on payment, utilization, and quality controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix. A comparison of the CHF patient population of BPCI participants from baseline to intervention relative to the same time periods for the comparison CHF patient population suggests that the CHF patients in BPCI-participating HHAs became less intensive or severe, relative to comparison providers.

There was a significant relative change in the mix of patients by age group (see Exhibits 202a & 202b). BPCI HHAs had no change in the proportion of patients in each age group, while comparison HHAs had a decrease in patients aged 20-64 years (2.3 percentage points, $p = 0.02$) and an increase in patients aged 80 years and older (-3.6 percentage points, $p = 0.06$). In contrast, BPCI-participating HHAs had a significant increase in the proportion of patients that were eligible for Medicare due to disability (and not end stage renal disease) (2.4 percentage points, $p = 0.02$).

Exhibits 202a & 202b: Characteristics of BPCI and Comparison Beneficiaries with a HHA Admission for Congestive Heart Failure, Model 3, Q4 2013 - Q3 2015

Characteristic		BPCI HHAs				Comparison HHAs				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	160	7%	170	7%	157	6%	109	4%	2.3	0.4	4.1	0.02
	65-79	743	30%	770	30%	722	29%	714	28%	1.3	-2.2	4.9	0.46
	80+	1,558	63%	1,611	63%	1,582	64%	1,728	68%	-3.6	-7.4	0.1	0.06
Gender	Female	1,494	61%	1,548	61%	1,442	59%	1,464	57%	1.2	-2.7	5.0	0.55
Medicaid and Disability	% Eligible for Medicaid	687	28%	788	31%	544	22%	523	21%	4.6	1.2	8.0	0.01
	% Disability, no ESRD	187	8%	191	7%	190	8%	133	5%	2.4	0.4	4.4	0.02
Utilization Six Months Prior to Index Hospitalization*	Inpatient Acute Care Hospital	1,301	53%	1,314	52%	1,299	53%	1,293	51%	0.7	-3.2	4.7	0.71
	Emergency Room Admission	802	33%	825	32%	822	33%	865	34%	-0.8	-4.4	2.9	0.69
	Home Health	1,163	47%	1,147	45%	1,068	43%	1,123	44%	-2.9	-6.8	1.0	0.14
	Inpatient Rehabilitation Facility	96	4%	97	4%	85	3%	75	3%	0.4	-1.0	1.9	0.57
	Skilled Nursing Facility	345	14%	339	13%	390	16%	429	17%	-1.7	-4.5	1.1	0.23
	Psychiatric Hospital	11	0%	11	0%	13	1%	8	0%	0.2	-0.3	0.7	0.45
	Long-term Care Hospital	17	1%	12	0%	15	1%	12	0%	-0.1	-0.7	0.5	0.78
	Institutional Nursing Facility	744	30%	648	25%	756	31%	707	28%	-1.8	-5.4	1.7	0.31
	No Institutional Care	1,126	46%	1,188	47%	1,114	45%	1,197	47%	-0.8	-4.7	3.1	0.67
	No Post-acute Care	666	27%	756	30%	728	30%	748	29%	2.8	-0.7	6.4	0.12

Characteristic	BPCI HHAs				Comparison HHAs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	4.03	2.42	4.13	2.56	4.09	2.47	4.22	2.48	-0.03	-0.23	0.16	0.75

Note: Count represents the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period)

*Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

+ The values for the matched comparison providers are weighted to take into account that some are matched to multiple participants.

The Outcome and Assessment Information Set (OASIS) assessment, taken at the initiation of home health care, indicates that the health and functional status of BPCI CHF patients at the beginning of their episode were better in the intervention period relative to the baseline relative to the change for the comparison group (see Exhibit 203). The share of BPCI patients with poor overall status declined by 17.4 percentage points ($p<0.01$) relative to a matched comparison group. The proportion of patients with poor overall status decreased from 40% to 27% for BPCI EIs between the baseline and intervention periods, and it increased from 54% to 59% for comparison HHAs. The proportion of patients requiring assistance or unable to transfer from the bed increased from 25% to 28% for BPCI EIs and from 19% to 46% for comparison HHAs between the baseline and intervention periods (-24.3 percentage points, $p<0.01$). BPCI EIs also experienced a relative decrease in the percent of patients that required the use of a bedside commode or were totally dependent in toileting (-4.9 percentage points, $p = 0.02$), the percent that required a walker or more assistance ambulating (-6.4 percentage points, $p<0.01$), and the percent short of breath from moderate to no exertion (-5.3 percentage points, $p=0.01$).

While almost all of these findings indicate that CHF patients treated by both BPCI and comparison HHAs experienced declines in health status between the intervention and baseline period, BPCI patients were generally healthier during the intervention than the comparison group. The relative difference is particularly pertinent for the measure of shortness of breath because of the importance of this symptom for patients with CHF. Further, while the HHA case-mix index decreased for both BPCI and comparison patients between the baseline and intervention periods, the BPCI-participating HHAs saw a larger decline in case-mix index (DiD -0.04, $p = 0.04$). However, even when these functional status and other characteristics present at admission were taken into account, payments and quality outcomes still decline significantly.

Exhibit 203: OASIS-based Characteristics of BPCI Beneficiaries and Matched Comparison Beneficiaries with an HHA Admission for Congestive Heart Failure, Model 3, Q4 2013 - Q3 2015

Characteristic	BPCI HHAs				Comparison HHAs				DiD	95% Confidence Interval		.05
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Pre-existing condition(s)*	897	43%	1006	46%	1067	50%	1076	50%	3.1	-1.1	7.3	0.15
Poor overall status**	836	40%	603	27%	1164	54%	1280	59%	-17.4	-21.5	-13.3	<0.01
Severely impaired vision or hearing	77	4%	63	3%	65	3%	51	2%	-0.1	-1.6	1.3	0.86
Sometimes, rarely, or never understands verbal content	97	5%	123	6%	113	5%	133	6%	0.1	-1.8	2.0	0.94
Unhealed pressure ulcer stage 2 or greater	102	5%	89	4%	84	4%	95	4%	-1.3	-3.0	0.4	0.14
Short of breath from moderate to no exertion***	1190	56%	1159	52%	1350	63%	1386	64%	-5.3	-9.4	-1.1	0.01
Require use of bedside commode or are totally dependent in toileting	374	18%	445	20%	278	13%	437	20%	-4.9	-8.1	-1.7	<0.01
Require assistance transferring or are unable to transfer (e.g., from bed to wheelchair)	535	25%	631	28%	398	19%	993	46%	-24.3	-28.1	-20.6	<0.01
Require walker or more assistance ambulating	1746	82%	1943	87%	1746	82%	2003	93%	-6.4	-9.3	-3.5	<0.01
Dependent in maintaining self-care (e.g., dressing, bathing)[4]	240	11%	240	11%	194	9%	213	10%	-1.4	-3.9	1.2	0.30
Dependent in ambulating+	182	9%	191	9%	160	7%	183	8%	-1.0	-3.3	1.3	0.39
Dependent in transferring+	160	8%	163	7%	141	7%	170	8%	-1.5	-3.7	0.7	0.18
Not likely to receive assistance in ADL++	157	7%	113	8%	159	7%	90	7%	1.6	-1.0	4.1	0.23
Caregiver needs training to provide supervision and safety, is unlikely to provide help, or is not present++	64	3%	39	3%	70	3%	54	4%	-0.9	-2.6	0.8	0.29
Urinary incontinence	1067	50%	1238	56%	1081	50%	1210	56%	-0.3	-4.5	3.9	0.88
Bowel incontinence	302	14%	353	16%	243	11%	281	13%	0.0	-2.9	2.8	0.97
Poor cognition+++	197	9%	246	11%	212	10%	254	12%	-0.1	-2.7	2.5	0.93

Characteristic	BPCI HHAs				Comparison HHAs				DiD	95% Confidence Interval		.05
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Often or constant confusion	179	8%	202	9%	187	9%	248	11%	-2.1	-4.6	0.3	0.09
Depressive symptoms‡	83	4%	86	4%	74	3%	70	3%	0.2	-1.4	1.7	0.84

Source: Lewin analysis of HHA OASIS assessment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

*Urinary incontinence, indwelling/suprapubic catheter, intractable pain, impaired decision-making, disruptive or socially inappropriate behavior, or memory loss to the extent that supervision is required.

**Likely to remain in fragile health and have ongoing high risk(s) of serious complications and death, or has serious progressive conditions that could lead to death within a year.

***Moderate exertion includes dressing, using a commode or bedpan, and walking distances less than 20 feet.

+Prior to patient's most recent illness, exacerbation, or injury.

++Current and future willingness and ability of non-agency caregivers to provide assistance.

+++Requires at least some assistance and direction in specific situations (e.g., on all tasks involving shifting of attention), consistently requires low stimulus environment due to distractibility, or is totally dependent due to disturbances such as constant disorientation, coma, persistent vegetative state, or delirium.

‡Had little interest in doing things or was feeling down, depressed, or hopeless more than half of the days in the last two weeks, or meets criteria for further evaluation for depression using a different standardized assessment.

e. Conclusions

Payments declined more for CHF patients in BPCI-participating HHAs than CHF patients in comparison HHAs. This was due to relative declines in SNF and HHA payments. BPCI had a statistically significant increase on ED visits during the first 90 days of the episodes but had no significant impact on readmissions or mortality. There was a relative improvement in one measure of functional status among BPCI beneficiaries relative to the comparison group (although this estimate may be biased as we rejected the null hypothesis that there were parallel trends between the two groups during the baseline period). The change in patient characteristics indicated that BPCI participating HHAs admitted healthier patients under the initiative. However, additional analysis suggests that even when these functional status and other characteristics present at admission were taken into account, payments and utilization outcomes still decline significantly.

2. Congestive Heart Failure, SNF

Medicare beneficiaries received 1,625 congestive heart failure (CHF) episodes of care in 205 BPCI-participating skilled nursing facilities (SNFs) in the first eight quarters of the initiative (from October 2013 through September 2015). An analysis of the 118 SNFs with a sufficient number of cases indicated that BPCI was not associated with statistically significant changes in payment; however, there was an increase in the proportion of BPCI patients with improvement in Activities of Daily Living (ADL) measures relative to a matched comparison group.

a. Participants

During the first eight quarters of the BPCI initiative, 205 SNFs (23% of all Model 3 BPCI SNFs) participated in the CHF episode. Thirty-five participants (17% of SNFs participating in this episode) chose a 30-day episode length, 15 (7%) chose a 60-day episode length, and 155 (76%) chose a 90-day episode length. There were 1,625 CHF episodes initiated by these providers during the BPCI initiative (approximately 7% of Model 3 SNF episodes across the 48 BPCI clinical episodes), which is an average of 4.8 episodes per quarter of participation. Because participants were allowed to join BPCI over an extended period, the CHF results are based on an average of three quarters of experience. Approximately 37% of SNFs participating in this episode began participating in Q3 2015 and 70% joined in either Q2 or Q3 2015. Out of the 205 SNFs, 20 stopped participating in the CHF episode within the first eight quarters of the initiative. BPCI interviewees indicated several reasons for dropping this episode including that they had difficulties identifying their BPCI CHF patients because of delays in the hospital assigning the final Medicare Severity-Diagnosis Related Group (MS-DRG), managing complex patients, reducing readmissions, and managing the associated financial risk, all of which have been raised as concerns for other medical episodes.[18]

Compared to SNFs *not* participating in BPCI, SNFs that participated in the CHF episode were larger and more likely to be located in an urban area (see Exhibits 204a & 204b). The participating SNFs were less likely to be hospital-based and had higher average overall nursing home quality scores. Prior to joining BPCI, participating SNFs had a higher number of CHF

[18] See Methods section for more information on the BPCI participant site visits and interviews.

discharges (11 vs. 6 discharges in 2011). One interviewee mentioned that relatively high CHF volume was a reason for selecting the episode.

Matched comparison SNFs were identified for 118 of the 205 BPCI-participating SNFs in the sample.[19] The SNFs that were excluded from the matching process had very low volumes of CHF episodes (an average of 3.3 episodes for excluded SNFs, compared with 12.9 episodes for included SNFs across all quarters of participation). The matched BPCI-participating SNFs were not significantly different from the comparison SNFs.

[19] Each BPCI-participating SNF was matched with up to 15 comparison SNFs based on a propensity score model that considered market and SNF characteristics and baseline outcomes. There were 87 participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the CHF episode or had fewer than five relevant discharges in 2011 or 2012.

Exhibits 204a & 204b: Characteristics of BPCI-participating SNFs with Congestive Heart Failure Episodes, compared with Non-participating SNFs, All BPCI-participating SNFs, and Comparison SNFs Model 3, Q4 2013 - Q3 2015

Characteristic		All CHF BPCI SNFs (N=205)**		Non-participating SNFs (N=13,302)		Matched CHF BPCI SNFs (N=118)		Comparison SNFs (N=1,725+)	
		N	%	N	%	N	%	N	%
Ownership	Non-Profit	43	21%	3,311	25%	25	21%	374	22%
	Government	1	0%	617	5%*	0	0%	0	0%
	For Profit	161	79%	9,374	70%	93	79%	1,351	78%
Urban/Rural	Urban	183	89%	9,364	70%*	110	93%	1,610	93%
IRF in CBSA	Yes	111	54%	7,238	54%	74	63%	1097	64%
Hospital-Based	Yes	2	1%	586	4%*	2	2%	40	2%
Part of Chain	Yes	49	24%	2,946	22%	30	25%	442	26%

Characteristic	All CHF BPCI SNFs (N=205)**	Non-participating SNFs (N=13,302)	Matched CHF BPCI SNFs (N=118)	Comparison SNFs (N=1,725+)
	Mean	Mean	Mean	Mean
Bed Count	127	112*	138	140
SNF Market Share	6%	6%	6%	6%
Nursing Home Overall Score‡	3.6	3.3*	3.7	3.7
Number of Congestive Heart Failure Discharges, 2011	11	6*	15	16
Standardized Part A Allowed Payment 90 days following SNF admission, Congestive Heart Failure, 2011	**$24,036**	**$24,665**	**$25, 970**	**$26,064**

PDP = post-discharge period; CHF = congestive heart failure.

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating SNFs are defined as Model 3 EIs, Q4 2013 – Q3 2015. Non-participating SNFs are all other SNFs.

* Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" SNFs are compared to "All CHF BPCI" SNFs. "Comparison" SNFs are compared to "Matched CHF BPCI" SNFs.

** BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

‡ Number of points out of 5 in overall rating in three areas: Quality, Survey/Health Inspections, and Staffing

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple participants. The unique number of matched comparison providers is 1,139.

b. Change in Patient Mix

We monitor changes in patient characteristics because BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce their episode payments below their target. Monitoring patient characteristics is particularly important under Model 3 because post-acute care (PAC) providers can assess potential patients prior to admission and use this information in making admission decisions. While the impact analysis on payment,

utilization, and quality controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix.

There were few differences in the change in CHF patient characteristics between the BPCI and comparison SNFs from the baseline to the intervention period, possibly due to the limited number of observations (see Exhibits 205a & 205b). There was a relative increase in the proportion of patients aged 20 to 64 (2.0 percentage points, $p = 0.07$) and a relative decline in the proportion over 80 (-4.3 percentage points, $p = 0.09$). There was also a relative increase in the proportion of patients that were Medicaid-eligible (4.0 percentage points, $p = 0.06$).[20] There was no statistically significant difference in the change in prior health care services use or in the average number of medical conditions.

[20] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

Exhibits 205a & 205b: Characteristics of BPCI and Comparison Beneficiaries with a SNF Admission for Congestive Heart Failure, Model 3, Q4 2013 - Q3 2015

Characteristic		BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	50	4%	75	5%	44	4%	41	3%	2.0	-0.2	4.1	0.07
	65-79	291	24%	390	28%	287	23%	353	26%	2.4	-2.4	7.1	0.33
	80+	889	72%	917	66%	899	73%	988	71%	-4.3	-9.3	0.6	0.09
Gender	Female	724	59%	828	60%	779	63%	823	60%	4.8	-0.5	10.1	0.07
Medicaid and Disability	% Eligible for Medicaid	203	17%	239	17%	273	22%	262	19%	4.0	-0.2	8.3	0.06
	% Disability, no ESRD	58	5%	84	6%	46	4%	48	3%	1.6	-0.6	3.9	0.16
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	731	59%	826	60%	709	58%	786	57%	1.1	-4.3	6.5	0.69
	Emergency Room Admission	442	36%	471	34%	443	36%	471	34%	0.1	-5.1	5.3	0.98
	Home Health	591	48%	724	52%	533	43%	659	48%	0.0	-5.4	5.4	1.00
	Inpatient Rehabilitation Facility	48	4%	65	5%	50	4%	68	5%	-0.1	-2.3	2.2	0.96
	Skilled Nursing Facility	429	35%	434	31%	462	38%	457	33%	1.0	-4.1	6.2	0.70
	Psychiatric Hospital	11	1%	6	0%	9	1%	6	0%	-0.2	-1.0	0.7	0.71
	Long-term Care Hospital	20	2%	23	2%	20	2%	13	1%	0.7	-0.6	2.0	0.28
	Institutional Nursing Facility*	661	54%	600	43%	686	56%	632	46%	-0.3	-5.7	5.1	0.92
	No Institutional Care	460	37%	523	38%	489	40%	558	40%	-0.2	-5.5	5.1	0.95
	No Post-acute Care	268	22%	285	21%	298	24%	326	24%	-0.5	-5.1	4.0	0.82

Characteristic	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	4.48	2.68	4.65	2.80	4.40	2.72	4.36	2.62	0.21	-0.09	0.50	0.17

Note: Count represents the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Nursing Home Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

The Minimum Data Set (MDS) assessment, taken at the beginning of SNF care, provides information about the health and functional status of patients. As with the claims and enrollment-based measures, MDS assessment data did not reveal many differences in the change in patient mix from baseline to intervention between the BPCI and comparison SNFs. There was a relative decline of 1.1 percentage points in the proportion of BPCI patients that do not believe they would improve enough to require less assistance with ADLs in the future ($p = 0.01$), and a relative decline of 4.0 percentage points in the proportion of patients who needed extensive assistance or were totally dependent moving in bed ($p = 0.09$) (see Exhibit 206). Conversely, matched comparison SNFs saw a greater reduction, by 3.4 percentage points ($p<0.01$), in the proportion of patients that were rarely or never able to communicate.

Exhibit 206: Additional characteristics of BPCI and Comparison Beneficiaries with a SNF Admission for Congestive Heart Failure, from Minimum Data Set Assessments, Model 3, Q4 2013 - Q3 2015

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Not currently married	802	70%	883	69%	827	72%	898	68%	3.4	-1.8	8.5	0.20
Intellectual or developmental disability (ID/DD)	1	0%	3	0%	2	0%	1	0%	0.4	-0.3	1.0	0.25
Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	9	1%	10	1%	9	1%	7	1%	0.2	-0.7	1.1	0.63
Sometimes, rarely, or never understands messages	56	5%	42	3%	61	5%	55	4%	-0.5	-2.7	1.8	0.67
Moderate to severe cognitive impairment	362	35%	351	31%	395	37%	456	37%	-4.0	-9.6	1.6	0.17
Rarely or never able to communicate	34	3%	42	3%	75	6%	43	3%	3.4	1.3	5.6	<0.01
Moderate to severe depression	58	5%	38	3%	95	8%	61	5%	1.4	-1.1	3.9	0.29
Rejected necessary evaluation or care at least once	62	5%	49	4%	75	6%	64	5%	0.0	-2.4	2.4	0.99
Need extensive assistance or are totally dependent moving in bed	915	78%	1066	79%	869	74%	1071	80%	-4.0	-8.6	0.6	0.09
Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	890	76%	1053	78%	868	74%	1059	79%	-2.0	-6.7	2.7	0.40
Need extensive assistance or are totally dependent walking in room	780	67%	899	68%	768	66%	915	68%	-1.6	-6.8	3.6	0.55
Need extensive assistance or are totally dependent using the toilet	932	79%	1101	82%	905	77%	1106	82%	-2.4	-6.8	2.0	0.28
Do not believe they will improve enough to require less assistance with ADL in the future	87	11%	50	5%	62	9%	72	8%	-4.8	-8.5	-1.1	0.01
Frequently or always had urinary incontinence	340	29%	405	30%	311	27%	356	27%	1.3	-3.7	6.3	0.60
Frequently or always had bowel incontinence	335	29%	423	32%	278	24%	362	27%	-0.3	-5.2	4.7	0.92
Active diagnosis of heart failure**	1024	87%	1188	88%	1014	87%	1203	90%	-1.7	-5.3	1.9	0.35
Active diagnosis of Alzheimer's	31	3%	25	2%	37	3%	30	2%	0.1	-1.6	1.9	0.86
Active diagnosis of dementia	191	16%	238	18%	203	17%	213	16%	2.9	-1.2	7.1	0.17
Active diagnosis of asthma, COPD, or chronic lung disease	452	38%	491	36%	402	34%	485	36%	-3.7	-9.0	1.6	0.17

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	537	46%	533	40%	536	46%	536	40%	0.2	-5.3	5.7	0.94
Experienced fracture due to a fall+	57	6%	57	5%	52	5%	66	5%	-1.1	-3.7	1.6	0.42
Require special treatment‡	680	58%	790	59%	667	57%	772	57%	0.4	-5.1	5.9	0.89

DiD=difference-in-differences.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* During 7-day lookback period (1st week of stay) unless otherwise noted.

** In order to have an active diagnosis of heart failure during the admission assessment (which occurs within 14 days of SNF admission), there must be a physician-documented diagnosis of heart failure within the last 60 days and the diagnosis must have a direct relationship to the resident's current functional, cognitive, mood or behavior status, medical treatments, nursing monitoring, or risk of death during the 7-day look-back period. The finding that less than 100% of the CHF patients in this episode had an active diagnosis of heart failure is likely due to differing definitions.

+ Within 6 months prior to admission.

‡ Examples include chemotherapy, oxygen therapy, and transfusions, measured after SNF admission using the five-day PPS assessment (regardless of resident status).

Source: Lewin analysis of SNF MDS assessment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

c. Medicare Payments and Quality of Care

BPCI was not associated with statistically significant changes in total payments (see Exhibit 207), payments covered during the bundle, or utilization. There was no statistically significant change in readmissions, emergency department (ED) use, or mortality from baseline to intervention for CHF BPCI patients relative to the comparison group.

Among patients that remained in the SNF long enough to receive two assessments, the proportion of BPCI patients with an improvement in functional status outcomes increased relative to the comparison group for all three ADL measures (see Exhibit 208). These measures are based on a subset of the patients as they require two MDS patient assessments conducted during the SNF stay. The proportion of BPCI patients with an improvement in overall function increased by 8.7 percentage points relative to the comparison group ($p=0.01$). The proportion of BPCI patients with an improvement in self-care function increased by 9.6 percentage points relative to the comparison group ($p<0.01$). The proportion of BPCI patients with an improvement in mobility function increased by 6.9 percentage points relative to the comparison group ($p=0.05$).

Exhibit 207: Impact of BPCI on Total Payments ($2015) for Qualifying Hospitalization plus 90-day Post-discharge Period, Congestive Heart Failure Episodes, Model 3 SNFs, Q4 2011 - Q3 2015

Total Amount	Number of intervention episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Total allowed amount, qualifying hospitalization plus 90-day PDP	1,372	1,373	$42,697	$40,623	$42,204	$40,535	-$405	-$2,484	$1,674	0.70

DiD = Difference-in-Differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; PDP =post-discharge period.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Exhibit 208: Functional Status of BPCI and Comparison CHF Patients, Model 3 SNFs, Q4 2011 - Q3 2015

Outcome	Number of intervention episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	95% LCI	95% UCI	p-value
Improved overall function	1,027	1,080	61.2%	65.8%	57.6%	53.5%	8.7	2.0	15.5	0.01
Improved self-care function	1,043	1,080	39.9%	48.3%	36.9%	35.7%	9.6	3.1	16.1	<0.01
Improved mobility function	1,018	1,080	53.8%	60.1%	49.4%	48.9%	6.9	0.1	13.7	0.05

DiD = Difference-in-Differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare Minimum Data Set (MDS) and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Note: "Improvement" means improvement or staying in the best possible status.

d. Conclusions

There were no statistically significant changes in payment or utilization in CHF episodes initiated in Model 3-participating SNFs relative to a matched comparison group. There were, however, relative improvements in all three functional status measures among the patients who remained in the SNF long enough to receive two patient assessments. This suggests that for those patients that remain in the SNF, BPCI was associated with an improvement in ADLs.

3. Major Joint Replacement of the Lower Extremity, SNF

There were 3,322 major joint replacement of the lower extremity (MJRLE) episodes of care in 308 BPCI-participating skilled nursing facilities (SNF) in the first eight quarters of the initiative (from October 2013 through September 2015).[21] We distinguish between planned joint replacements not due to fracture (MJRLE-NF) and emergent replacements due to fracture (MJRLE-F) because care needs and types of patients are quite different according to the Technical Expert Panel on MJRLE episodes. For non-fractures, hospitals and physicians reported that because these are typically scheduled surgeries, patients often receive pre-surgical education and decide upon the post-hospital discharge site of care in consultation with their clinicians before the surgery. Additionally, beneficiaries receiving a scheduled lower joint replacement typically are healthier than the average Medicare beneficiary. In contrast, hip replacements due to a fracture are emergent procedures and typically do not allow for patient education or planning for the episode of care. Further, fractures more often involve frailer individuals who have more comorbidities. Because of these differences, we analyzed MJRLE episodes with and without fracture separately.

Based on the subset of episodes for which we found comparison providers (N=153) and patients (N=2,321), the average Medicare payments for the services included in the 90-day episodes decreased more for BPCI episodes than the matched comparison episodes.[22] Payments for BPCI fracture episodes declined by $5,419 ($p<0.01$) relative to episodes from matched comparison group providers. Payments for non-fracture episodes declined by $1,978 ($p<0.01$) compared with a matched sample of episodes. The reduction in Medicare payments was due primarily to reduced SNF payments as a result of reductions in SNF days during the 90-day post-hospital discharge period. SNF payments declined $4,199 ($p<0.01$) more for BPCI fracture episodes and $1,658 ($p<0.01$) more for BPCI non-fracture episodes than for comparison episodes.

There was no difference in readmissions, emergency department (ED) visits, or mortality among BPCI episodes relative to the comparison group. Non-fracture patients, however, experienced a statistically significant decline in self-care functional status relative to the comparison group during the SNF stay among the patients who received two SNF patient assessments. For non-fracture episodes, the proportion of BPCI beneficiaries with improvement in self-care function during the SNF stay declined 11.9 percentage points ($p=0.03$) relative to the comparison group.

[21] Because participants were allowed to join BPCI over an extended period, these early participants had an average of 2 quarters of experience in Model 3.

[22] Results presented in this brief are based on total Medicare payments, standardized to remove the effect of geographic and other adjustments and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. These results do *not* take into account the BPCI payment reconciliation process. For a complete discussion of the reconciliation process see https://innovation.cms.gov/Files/x/BPCI_Model3Background.pdf.

The shorter time in the SNF may have contributed to the lower improvement in self-care function for these patients. There were no statistically significant differences in functional improvement for patients with a fracture during the SNF stay. At the same time, BPCI respondents across all Model 3 SNF MJRLE episodes experienced a statistically significant improvement in one functional status measure from the beneficiary survey relative to the comparison group. They were 5.4 percentage points less likely to report a decline in their ability to walk (p=0.05) relative to the comparison group. However, BPCI respondents reported worse health care experiences in three out of 10 measures in the beneficiary survey relative to comparison respondents: consideration of respondent preferences for post-acute care (PAC), clear explanations of medication instructions before going home, and clear explanations of follow-up appointments or treatments before going home.

SNFs have considerable discretion to determine who they admit to their facilities. According to site visit interviewees,[23] one strategy to improve coordination between hospitals and SNFs under BPCI is for the SNFs to send a representative to visit patients in the hospital prior to the SNF admission. SNFs may use information gathered during these visits to decide who to admit and who not to admit. Particularly in markets where there are not excess SNF beds, SNFs may be more selective in their admissions. Under BPCI, they have strong incentives not to admit patients who are more likely to have higher care needs than the average patient. A comparison of patient characteristics available in claims data and the initial SNF patient assessment suggests that BPCI participants may have selected healthier patients after joining BPCI relative to the comparison group. We conducted sensitivity analyses to determine if the changes observed in the patient assessment data explain the significant BPCI impacts on payments. We found that even after adjusting for functional status and other characteristics present at admission, the impact estimates remained statistically significant.

a. Participants

During the first eight quarters of the BPCI initiative, 308 SNFs (35% of all BPCI-participating SNFs) participated in the MJRLE episode. Most of these SNFs chose an episode length of 90 days (71%), while 15% chose the 60-day length and 14% chose the 30-day episode length. The most commonly cited reason, gathered through participant interviews and site visits, for why BPCI SNFs chose MJRLE was due to the high volume of patients in this episode. BPCI providers initiated 3,322 MJRLE episodes during the initiative (approximately 15% of all Model 3 SNF episodes across the 48 BPCI clinical episodes). Because participants were allowed to join BPCI over an extended period, the MJRLE results are based on an average of two quarters of experience. Approximately 61% of the SNFs began participating in MJRLE in Q3 2015; 94% joined in either Q2 or Q3 2015. Furthermore, of the 308 SNFs, 32 (10%) terminated their participation in MJRLE within the first eight quarters of the initiative.

Compared with SNFs not participating in BPCI, SNFs that participated in the MJRLE episode were less likely to be government owned and hospital based and more likely to be in urban locations (Exhibits 209a & 209b). In addition, prior to joining BPCI, participating SNFs had more MJRLE admissions.

[23] See Methods section for more information on the BPCI participant site visits and interviews.

Compared with all Model 3 BPCI-participating SNFs, the 308 SNFs participating in MJRLE were more likely to be not for profit and more likely to be part of a chain (Exhibits 209a & 209b). The SNFs participating in MJRLE also had higher standardized Part A Medicare payments ($2015) for MJRLE episodes in 2011 than all Model 3 BPCI SNFs ($22,469 vs. $19,459). Matched comparison SNFs were identified for 153 of the 308 BPCI-participating SNFs in the sample.[24] Matched BPCI-participating SNFs were not statistically different from the comparison SNFs with respect to any measures of note.

[24] Each BPCI-participating SNF was matched with up to 15 comparison SNFs based on a propensity score model that considered market and SNF characteristics and baseline outcomes. There were 154 participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one quarter before they stopped participating in the MJRLE episode or had fewer than five relevant discharges in 2011 or 2012. There was 1 other participant for which there was not a match within the caliper.

Exhibits 209a & 209b: Characteristics of BPCI-participating SNFs with Major Joint Replacement of the Lower Extremity Episodes, compared with Non-participating SNFs, All BPCI-participating SNFs, and Comparison SNFs, Model 3, Q4 2013 - Q3 2015

Characteristic		All MJRLE BPCI SNFs (N=308)		Non-participating SNFs (N=13,302)		All BPCI SNFs (N=864)**		Matched MJRLE BPCI SNFs (N=153)		Comparison SNFs (N=2,430+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	73	24%	3,311	25%	121	14%*	51	33%	767	34%
	Government	1	0%	617	5%*	3	0%	1	1%	22	1%
	For Profit	234	76%	9,374	70%	740	86%*	101	66%	1,641	65%
Urban/Rural	Urban	263	85%	9,364	70%*	724	84%	140	92%	2,039	91%
IRF in CBSA	Yes	165	54%	7,238	54%	488	56%	77	50%	1,221	48%
Hospital-Based	Yes	2	1%	586	4%*	7	1%	2	1%	40	2%
Part of Chain	Yes	90	29%	2,946	22%	168	19%*	48	31%	873	35%

Characteristic	All MJRLE BPCI SNFs (N=308)	Non-participating SNFs (N=13,302)	All BPCI SNFs (N=864)**	Matched MJRLE BPCI SNFs (N=153)	Comparison SNFs (N=2,430+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	117	112	122	125	124
SNF Market Share	6%	6%	6%	8%	8%
Nursing Home Overall Score++	3.45	3.32	3.45	3.83	3.81
Number of MJRLE Admissions, 2011	18	10*	24	29	30
Standardized Part A Allowed Payment 90 days following SNF admission, MJRLE, 2011	**$22,469**	**$21,621**	**$19,459***	**$18,812**	**$18,587**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating SNFs are defined as Model 3 SNFs, Q4 2013 – Q3 2015. Non-participating SNFs are all other SNFs.

*Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" SNFs are compared to "All MJRLE BPCI" SNFs. "Comparison" SNFs are compared to "Matched MJRLE BPCI" SNFs.

** BPCI-participating SNFs that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI SNFs. The unique number of matched comparison providers is 1,489.

++ Number of points out of 5 in overall rating and in three areas: Quality, Survey/Health Inspections, and Staffing.

b. Change in Patient Mix

We monitor changes in patient characteristics because BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce their episode payments below their target. While the impact analysis on payment, utilization, and quality controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix. A comparison of patient characteristics available in the claims and the initial SNF patient assessment (the five-day Minimum Data Set assessment) suggest that BPCI participants may have selected healthier patients after joining BPCI relative to the comparison group,

particularly for non-fracture episodes (Exhibits 210a & 210b and 212a & 212b). Among the fracture episodes, there was a statistically significant decrease in the proportion of BPCI patients that upon admission needed extensive assistance or were totally dependent moving in bed (-5.5 percentage points, p=0.03); transferring, such as between a bed and a wheelchair (-8.4 percentage points, p<0.01); walking in room (-10.7 percentage points, p<0.01); or using the toilet (-8.4 percentage points, p<0.01); as well as a significant decrease in the proportion who frequently or always had urinary (-12.0 percentage points, p<0.01) or bowel incontinence (-12.3 percentage points, p<0.01) (Exhibit 211).[25]

There were similar changes in patient characteristics for the non-fracture episodes (Exhibits 212a & 212b). There was a statistically significant increase of 4.2 percentage points (p=0.05) in the proportion of BPCI patients between the ages of 65 and 80 relative to the matched comparison group. There was also a statistically significant decline of 2.0 percentage points (p=0.09) in the proportion of patients who had a SNF stay during the six months prior to the qualifying hospitalization. The initial SNF patient assessment indicates that BPCI SNFs had a greater decline in the proportion of MJRLE-NF patients who needed extensive assistance or were totally dependent moving in bed (-11.4 percentage points, p<0.01), transferring (-12.3 percentage points, p<0.01), walking in room (-6.5 percentage points, p<0.01), or using the toilet (-10.5 percentage points, p<0.01), and frequently or always had urinary incontinence (-2.0 percentage points, p=0.04) (Exhibit 213). In addition, BPCI SNFs had a greater decline in the proportion of patients who had an active diagnosis of dementia (-1.9 percentage points, p<0.01) or an active diagnosis of asthma, chronic obstructive pulmonary disease (COPD), or chronic lung disease (-3.9 percentage points, p=0.01) relative to the change for the comparison group. There was also a decrease in the proportion of patients that do not believe they will improve enough to require less assistance with activities of daily living in the future (-4.3 percentage points, p<0.01), the proportion of patients with moderate to severe cognitive impairment (-2.1 percentage points, p=0.07), and the proportion of patients who require special treatment (-3.5 percentage points, p=0.02).

[25] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a comparison group. The changes are observed before and after the intervention began.

Exhibits 210a & 210b: Characteristics of BPCI and Comparison Patients with a SNF Admission for Major Joint Replacement of the Lower Extremity due to a Fracture, Model 3, Q4 2013 - Q3 2015

Characteristic		BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	8	1%	11	2%	13	2%	13	3%	0.3	-1.9	2.5	0.78
	65-79	149	21%	126	25%	156	22%	126	25%	1.0	-5.8	7.8	0.78
	80+	555	78%	369	73%	543	76%	367	73%	-1.3	-8.2	5.7	0.72
Gender	Female	522	73%	372	74%	535	75%	367	73%	2.8	-4.3	9.9	0.44
Medicaid and Disability	% Eligible for Medicaid	76	11%	54	11%	102	14%	57	11%	3.1	-2.2	8.3	0.25
	% Disability, no ESRD	9	1%	13	3%	19	3%	14	3%	1.2	-1.2	3.6	0.32
Utilization Six Months Prior to Qualifying Hospitalization	Inpatient Acute Care Hospital	176	25%	115	23%	147	21%	114	23%	-3.9	-10.6	2.9	0.26
	Emergency Room Admission	213	30%	142	28%	180	25%	135	27%	-3.3	-10.5	3.9	0.38
	Home Health	177	25%	110	22%	181	25%	125	25%	-2.4	-9.3	4.5	0.50
	Inpatient Rehabilitation Facility	10	1%	10	2%	8	1%	9	2%	-0.1	-2.1	1.9	0.93
	Skilled Nursing Facility	105	15%	66	13%	93	13%	73	14%	-3.1	-8.6	2.5	0.28
	Psychiatric Hospital	10	1%	2	0%	7	1%	2	0%	-0.4	-1.9	1.1	0.58
	Long-term Care Hospital	0	0%	1	0%	1	0%	2	0%	-0.1	-0.7	0.6	0.86
	Institutional Nursing Facility*	215	30%	96	19%	227	32%	121	24%	-3.3	-10.4	3.9	0.37
	No Institutional Care	519	73%	378	75%	550	77%	374	74%	5.1	-1.9	12.1	0.15
	No Post-acute Care	357	50%	262	52%	368	52%	268	53%	0.4	-7.7	8.4	0.93

Characteristic	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	2.46	1.78	2.49	1.87	2.43	1.81	2.44	1.82	0.02	-0.28	0.31	0.91

DiD = difference-in-differences.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and matched comparison providers.

Exhibit 211: Additional Characteristics of BPCI and Comparison Patients with a SNF Admission for Major Joint Replacement of the Lower Extremity due to a Fracture, Model 3, Q4 2013 - Q3 2015

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Not currently married	473	71%	309	66%	455	68%	317	67%	-4.1	-11.9	3.7	0.30
Intellectual or developmental disability (ID/DD)	6	1%	2	1%	3	1%	2	0%	-0.4	-2.0	1.2	0.61
Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	13	2%	3	1%	3	0%	9	2%	-2.7	-4.5	-0.9	0.00
Sometimes, rarely, or never understands messages	62	9%	35	7%	75	11%	45	9%	-0.1	-4.9	4.7	0.97
Moderate to severe cognitive impairment	250	41%	154	34%	250	41%	180	40%	-5.5	-13.9	2.9	0.20
Rarely or never able to communicate	42	6%	16	3%	56	8%	26	5%	0.1	-3.8	4.0	0.98
Moderate to severe depression	50	8%	17	4%	61	9%	29	6%	-0.8	-5.0	3.4	0.71
Rejected necessary evaluation or care at least once	50	7%	22	5%	46	7%	24	5%	-0.9	-4.9	3.0	0.64
Need extensive assistance or are totally dependent moving in bed	621	91%	420	87%	608	90%	443	90%	-5.5	-10.5	-0.4	0.03
Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	618	91%	402	83%	609	90%	443	90%	-8.4	-13.6	-3.2	0.00
Need extensive assistance or are totally dependent walking in room	553	81%	325	67%	554	82%	383	78%	-10.7	-17.4	-3.9	0.00
Need extensive assistance or are totally dependent using the toilet	627	92%	413	85%	609	90%	447	91%	-8.4	-13.4	-3.5	0.00
Do not believe they will improve enough to require less assistance with ADL in future	87	15%	48	13%	63	13%	40	10%	0.8	-5.5	7.1	0.80
Frequently or always had urinary incontinence	265	39%	146	30%	237	35%	186	38%	-12.0	-19.9	-4.1	0.00
Frequently or always had bowel incontinence	234	35%	120	25%	219	32%	170	35%	-12.3	-19.9	-4.6	0.00
Active diagnosis of heart failure	91	13%	49	10%	92	14%	59	12%	-1.8	-7.2	3.7	0.52
Active diagnosis of Alzheimer's	57	8%	28	6%	45	7%	28	6%	-1.7	-5.8	2.4	0.42
Active diagnosis of dementia	168	25%	99	20%	173	25%	108	22%	-0.9	-7.8	6.1	0.81

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Active diagnosis of asthma, COPD, or chronic lung disease	111	16%	81	17%	112	16%	78	16%	1.0	-5.1	7.0	0.76
Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	90	13%	58	12%	90	13%	70	14%	-2.3	-7.9	3.3	0.42
Experienced fracture due to a fall**	578	90%	404	87%	599	90%	407	86%	2.0	-3.4	7.3	0.47
Require special treatment +	158	23%	100	21%	168	25%	117	24%	-1.8	-8.7	5.2	0.62

DiD = difference-in-differences, LCI = lower 95% confidence interval, UCI = upper 95% confidence interval.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* During 7-day lookback period (1st week of stay) unless otherwise noted.

** Within 6 months prior to admission

+ Examples include chemotherapy, oxygen therapy, and transfusions; measured after SNF admission using the five-day PPS assessment (regardless of resident status).

Source: Lewin analysis of SNF MDS assessment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and matched comparison providers.

Exhibit 212a & 212b: Characteristics of BPCI and Comparison Patients with a SNF Admission for Major Joint Replacement of the Lower Extremity, Non-fracture, Model 3, Q4 2013 - Q3 2015

Characteristic		BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	173	8%	117	6%	150	7%	130	7%	-1.7	-3.9	0.5	0.13
	65-79	1,345	59%	1,147	63%	1,373	60%	1,092	60%	4.2	0.0	8.5	0.05
	80+	781	34%	551	30%	776	34%	593	33%	-2.5	-6.6	1.6	0.22
Gender	Female	1,713	75%	1,328	73%	1,679	73%	1,331	73%	-1.6	-5.5	2.2	0.40
Medicaid and Disability	% Eligible for Medicaid	270	12%	166	9%	309	13%	213	12%	-0.9	-3.7	1.9	0.53
	% Disability, no ESRD	217	9%	144	8%	193	8%	147	8%	-1.2	-3.6	1.2	0.33
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	318	14%	237	13%	279	12%	231	13%	-1.4	-4.3	1.6	0.36
	Emergency Room Admission	429	19%	317	17%	421	18%	341	19%	-1.7	-5.0	1.7	0.33
	Home Health	259	11%	208	11%	202	9%	183	10%	-1.1	-3.8	1.5	0.42
	Inpatient Rehabilitation Facility	15	1%	9	0%	12	1%	10	1%	-0.2	-0.8	0.5	0.58
	Skilled Nursing Facility	185	8%	130	7%	164	7%	150	8%	-2.0	-4.3	0.3	0.09
	Psychiatric Hospital	10	0%	6	0%	5	0%	3	0%	-0.1	-0.5	0.4	0.83
	Long-term Care Hospital	0	0%	1	0%	3	0%	2	0%	0.1	-0.2	0.3	0.53
	Institutional Nursing Facility*	417	18%	217	12%	395	17%	234	13%	-1.9	-5.0	1.2	0.24
	No Institutional Care	1,955	85%	1,557	86%	1,988	86%	1,559	86%	1.3	-1.7	4.4	0.39
	No Post-acute Care	1,583	69%	1,280	71%	1,611	70%	1,263	70%	2.2	-1.8	6.2	0.29

Characteristic	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	1.13	1.34	1.20	1.36	1.04	1.29	1.13	1.32	-0.03	-0.14	0.09	0.66

DiD=difference-in-differences.

Note: "Count" represent the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and matched comparison providers.

Exhibit 213: Additional Characteristics of BPCI and Comparison Patients with a SNF Admission for Major Joint Replacement of the Lower Extremity, Non-fracture, Model 3, Q4 2013 - Q3 2015

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Not currently married	1225	56%	888	51%	1143	52%	984	57%	-9.4	-13.9	-5.0	0.00
Intellectual or developmental disability (ID/DD)	6	0%	2	0%	2	0%	0	0%	-0.1	-0.6	0.5	0.84
Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	12	1%	5	0%	10	0%	4	0%	0.0	-0.6	0.5	0.91
Sometimes, rarely, or never understands messages	11	0%	3	0%	11	0%	5	0%	-0.1	-0.7	0.4	0.68
Moderate to severe cognitive impairment	159	7%	87	5%	164	8%	126	7%	-2.1	-4.4	0.2	0.07
Rarely or never able to communicate	38	2%	14	1%	40	2%	19	1%	-0.2	-1.2	0.8	0.72
Moderate to severe depression	63	3%	18	1%	76	3%	41	2%	-0.7	-2.1	0.7	0.31
Rejected necessary evaluation or care at least once	46	2%	27	2%	46	2%	35	2%	-0.5	-1.7	0.8	0.47
Need extensive assistance or are totally dependent moving in bed	1467	66%	917	52%	1475	67%	1123	64%	-11.4	-15.6	-7.1	0.00
Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	1450	65%	897	51%	1437	65%	1103	63%	-12.3	-16.5	-8.0	0.00
Need extensive assistance or are totally dependent walking in room	989	44%	601	34%	1067	48%	778	44%	-6.5	-10.9	-2.2	0.00
Need extensive assistance or are totally dependent using the toilet	1405	63%	876	50%	1416	64%	1068	61%	-10.5	-14.8	-6.1	0.00
Do not believe they will improve enough to require less assistance with ADL in future	131	9%	44	5%	82	6%	71	6%	-4.3	-7.1	-1.5	0.00
Frequently or always had urinary incontinence	107	5%	77	4%	107	5%	112	6%	-2.0	-3.9	-0.1	0.04
Frequently or always had bowel incontinence	64	3%	70	4%	72	3%	77	4%	0.0	-1.7	1.6	0.97
Active diagnosis of heart failure	108	5%	67	4%	103	5%	87	5%	-1.4	-3.2	0.5	0.15
Active diagnosis of Alzheimer's	13	1%	6	0%	10	0%	8	0%	-0.2	-0.9	0.4	0.42
Active diagnosis of dementia	65	3%	26	1%	45	2%	44	3%	-1.9	-3.2	-0.6	0.00
Active diagnosis of asthma, COPD, or chronic lung disease	322	14%	209	12%	303	14%	263	15%	-3.9	-7.0	-0.9	0.01
Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	153	7%	107	6%	162	7%	122	7%	-0.4	-2.7	1.8	0.69

Characteristic*	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Experienced fracture due to a fall**	53	3%	38	2%	35	2%	33	2%	-0.6	-2.0	0.7	0.35
Require special treatment [+]	248	11%	152	9%	272	12%	232	13%	-3.5	-6.3	-0.7	0.02

DiD=difference-in-differences.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

* During 7-day lookback period (1st week of stay) unless otherwise noted.

** Within 6 months prior to admission

[+] Examples include chemotherapy, oxygen therapy, and transfusions; measured after SNF admission using the five-day PPS assessment (regardless of resident status).

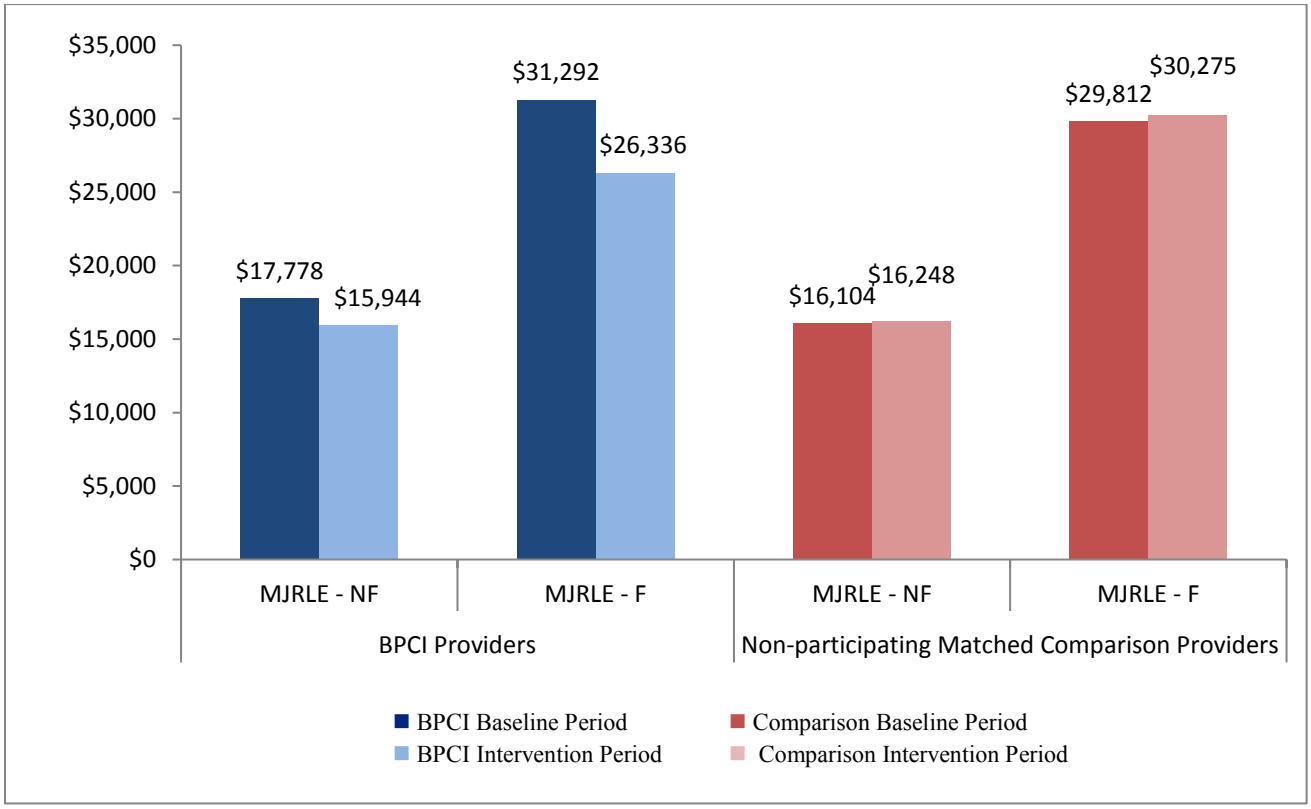
Source: Lewin analysis of SNF Minimum Data Set (MDS) assessment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and matched comparison providers.

c. Medicare Payments and Utilization

Total payments for the BPCI fracture episodes were $26,336 compared with $15,944 for the BPCI non-fracture episodes during the intervention, indicating the differences in the severity and care needs for patients with a joint replacement due to a fracture and those with an elective procedure.

Total Medicare payments during the 90-day episodes were statistically significantly lower during the intervention than the baseline for both fracture and non-fracture BPCI episodes compared to episodes from a matched comparison group of providers not participating in BPCI (Exhibit 214). For the 316 fracture episodes during the intervention period, total Medicare payments ($2015) for the 90-day episode declined by $5,419 more for BPCI episodes than for episodes from matched comparison providers ($p<0.01$). Total payments were $31,292 during the baseline period for BPCI fracture episodes, and declined by 16% to $26,336 during the intervention period. Total payments for comparison episodes were $29,812 during the baseline, and increased by 2% to $30,274 during the intervention period. Based on the 982 non-fracture 90-day episodes during the intervention period, total Medicare payments ($2015) for the 90-day episode period also had a statistically significant decline relative to the comparison group. Total Medicare payments ($2015) for 90-day episodes declined by $1,978 more for BPCI non-fracture episodes than episodes from the matched comparison group ($p<0.01$). Total payments were $17,778 during the baseline period for BPCI non-fracture episodes and declined by 10% to $15,944 during the intervention period. For comparison SNF non-fracture episodes, total payments were $16,104 during the baseline, and increased to $16,248 during the intervention period.

Exhibit 214: Medicare Payments ($2015) for 90-day episodes, MJRLE-F and MJRLE NF, Model 3, Q4 2011-Q3 2015



MJRLE-NF= Major Joint Replacement of the Lower Extremity (Non-Fracture); MJRLE-F= Major Joint Replacement of the Lower Extremity (Fracture).

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Notes: Medicare payments for the services included in the 90-day episode are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. The intervention period includes episodes that began from the fourth quarter of 2013 through the third quarter of 2015.

The greater reductions in Medicare payments for BPCI MJRLE episodes were driven by reduced SNF payments. SNF payments for BPCI fracture episodes declined $4,199 ($p<0.01$) more from the baseline to the intervention period than for comparison fracture episodes. Prior to BPCI, SNF payments for fracture episodes were $22,087 for BPCI providers, declining to $18,065 during the intervention period. For comparison providers, payments increased from $21,244 to $21,421. The relative change in SNF payments for non-fracture episodes was -$1,658 ($p<0.01$).

The changes in payments align with a reduction in the number of SNF days (Exhibit 215). The relative decline in SNF payments was greater for fracture patients than non-fracture payments due to a greater reduction in the number of SNF days among fracture patients. For fracture episodes, the number of SNF days declined from 40.1 to 32.4 days for BPCI episodes and were virtually unchanged for comparison episodes (39.2 days to 39.1 days), for a relative difference of 7.6 days ($p<0.01$). For non-fracture patients, there was a 2.9 day relative decline in SNF days

(p<0.01), from 19.2 to 16.5 days for the BPCI episodes, compared with 18.8 to 19.1 days for the comparison episodes.

While the proportion of patients using home health agency (HHA) services did not change between BPCI-participating SNFs and the comparison group from baseline to the intervention period, there was a relative increase in the number of HHA visits by 1.1 visits (p=0.04) for non-fracture episodes, suggesting SNFs were substituting HHA visits for SNF days in non-fracture episodes (Exhibit 215). The number of home health visits increased from 15.0 visits during the baseline to 16.7 visits during the intervention for non-fracture BPCI episodes and from 15.1 to 15.8 visits for comparison episodes. There was no statistically significant change in HHA visits for BPCI fractures patients relative to the comparison group.

The reduction in SNF days and the increase in HHA visits align with information we received through quarterly calls and site visits with BPCI participants. Model 3 interviewees mentioned that they try to reduce SNF length of stay and refer patients to HHAs as soon as possible to reduce their episode payments. However, interviewees indicated that convincing patients and family members that the patient should go home before the maximum number of SNF days covered by Medicare is challenging for the SNFs. One way BPCI-participating SNFs are addressing this challenge is to give one clear message and to educate patients and staff to expect shorter stays and that patients will be discharged home with home health care as soon as possible.

Exhibit 215: Impact of BPCI on Number of SNF Days and Home Health Visits, MJRLE Episodes, Fracture and Non-Fracture, Model 3 SNF, Q4 2011-Q3 2015

	Outcome	Number of intervention episodes		BPCI		Comparison		DID estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DID	LCI	UCI	P-value
Non-fracture	Number of SNF days	1,813	1,814	19.2	16.5	18.8	19.1	-2.9	-4.3	-1.6	<0.01
	Number of home health visits	1,145	1,115	15.0	16.7	15.1	15.8	1.1	0.1	2.1	0.04
Fracture	Number of SNF days	504	506	40.1	32.4	39.2	39.1	-7.6	-11.7	-3.4	<0.01
	Number of home health visits	319	298	18.6	20.4	18.4	19.2	1.0	-1.6	3.5	0.46

DiD = difference-in-differences, LCI = lower 95% confidence interval, UCI = upper 95% confidence interval at the 5% level.

Notes: Number of SNF days and HHA visits are based on patients who have at least one day/visit. The intervention period includes episodes that began from the fourth quarter of 2013 through the second quarter of 2015.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and matched comparison providers.

d. Quality of Care

There were no statistically significant changes in the claim-based or patient assessment-based outcomes for the BPCI-participating SNF fracture episodes relative to the comparison group. BPCI-participating SNFs achieved mixed results for the non-fracture episodes with respect to quality outcomes relative to the comparison SNFs (Exhibit 216). For non-fracture episodes, there was a relative reduction of 2.4 percentage points in unplanned readmissions (p=0.05) and a relative reduction of 1.0 percentage points in mortality during the 90 days from the start of the episode (p<0.01). The unplanned readmission rate decreased for BPCI-participating SNFs from 10.8% to 9.6% compared to an increase for the comparison group from 9.5% to 10.8%.

Exhibit 216: Impact of BPCI on Claim-based Quality Outcomes, MJRLE Episodes, Non-Fracture, Model 3 SNF, Q4 2011-Q3 2015

	Outcome	Number of intervention episodes		BPCI		Comparison		DID estimate			
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DID	LCI	UCI	P-value
Non-fracture	Unplanned readmission rate, 90 days from episode start date	1,807	1,804	10.8%	9.6%	9.5%	10.8%	-2.4	-4.9	0.0	0.05
	Mortality rate, 90 days from episode start date	1,813	1,813	0.6%	0.4%	0.3%	1.2%	-1.0	-1.7	-0.3	<0.01

DiD = difference-in-differences, LCI = lower 95% confidence interval, UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

Change in beneficiary function was calculated based on the change between two SNF patient assessments and indicated that BPCI fracture patients experienced no statistically significant change in functional status relative to the comparison patients, while non-fracture patients experienced a relative decline in functional status (Exhibit 217). For non-fracture episodes, the proportion of BPCI patients with improvement in self-care function during the SNF stay declined by 11.9 percentage points relative to the comparison group (p=0.03). The average proportion of BPCI non-fracture patients with improvement in self-care function during the SNF stay declined from 63.8% during the baseline period to 54.0% during the intervention period. For comparison group patients, the proportion with improvement in self-care function increased from 61.9% to 64.0%. It should be noted that non-fracture episodes had fewer SNF days under BPCI, that is, less time to recover function in the SNF. This shorter time may have contributed to these results. It is not possible to know if patients with shorter stays achieved the same level of function after leaving the SNF as they would have had they remained in the SNF longer.

Exhibit 217: Impact of BPCI on Functional Status, MJRLE Episodes, Fracture and Non-Fracture, Model 3 SNF, Q4 2011-Q3 2015

	Outcome	Number of intervention episodes		BPCI		Comparison		DID estimate	UCI	LCI	P-value
		BPCI	Comparison	Baseline	Intervention	Baseline	Intervention				
MJRLE-NF	Improved overall function	1,136	1,497	76.5%	69.7%	76.4%	76.8%	-7.2	-18.5	4.0	0.21
	Improved self-care function	1,136	1,498	63.8%	54.0%	61.9%	64.0%	-11.9	-22.4	-1.4	0.03
	Improved mobility function	1,136	1,496	71.7%	65.1%	72.1%	71.9%	-6.3	-16.8	4.3	0.24
MJRLE-F	Improved overall function	375	448	65.1%	69.8%	65.6%	69.5%	0.8	-8.6	10.3	0.86
	Improved self-care function	377	448	45.2%	48.3%	44.8%	47.8%	0.2	-9.4	9.7	0.97
	Improved mobility function	375	448	62.8%	64.1%	64.7%	67.1%	-1.0	-11.2	9.2	0.85

DiD = difference-in-differences, LCI = lower 95% confidence interval, UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare Minimum Data Set (MDS) and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Note: Improved means improvement or staying in the best possible status.

At the same time, according to the beneficiary survey, BPCI SNF respondents with a MJRLE episode reported similar changes for most functional status, mental and physical health, and health care experience outcomes from before to after their episode of care as the matched comparison group.[26] The only exception was that BPCI respondents were 5.4 percentage points less likely to report decline in their ability to walk (p=0.05) relative to the comparison group (see Exhibit 218). In addition, BPCI respondents reported significantly worse care experiences relative to the matched comparison group, for three outcomes: consideration of respondent preferences for post-acute care (-4.6 percentage points, p=0.03); medical staff clearly explained medication before patient went home (-5.1 percentage points, p=0.01); and follow-up appointments or treatments being clearly explained before going home (-6.1 percentage points, p<0.01). However, there was no statistically significant difference in overall satisfaction with the recovery between BPCI and comparison respondents.

[26] Subsequent surveys covering beneficiaries participating in the BPCI model in the spring (Wave 6) and summer (Wave 7) of 2016 were analyzed together at the conclusion of Wave 7. Results indicated worse care experiences among BPCI respondents relative to comparison respondents for seven of ten measures, although only two differences were statistically significant. However, none of the three negative outcomes that were significant in Wave 5 remained significant in Wave 7. Moreover, unlike Wave 5, where BPCI respondents were 3.0 percentage points less likely to be satisfied with their overall recovery, BPCI respondents in Wave 7 were 3.0 percentage points more likely to be satisfied with their overall recovery. Thus, while there seems to be some persisting negative perception of care experience across waves, it is not specific to any particular outcome, nor does it have a continued impact on overall satisfaction.

Exhibit 218: Impact of BPCI on Select Survey Outcomes, MJRLE Episodes, Model 3 SNF, May-June 2015 and October-November 2015

Measure		Wave	Number of Episodes		Risk adjusted Rates		Estimated Difference			
			BPCI	Comparison	BPCI	Comparison	Point Estimate	95% LCI	95% UCI	p-value
Functional Status	Decline in walking without rest	5	387	393	23.4	28.7	-5.4	-10.7	-0.0	0.05
Patient Care Experience	Patient's preferences considered for services after discharge	5	367	371	89.2	93.8	-4.6	-8.6	-0.5	0.03
	Medical staff clearly explained medication instructions	5	345	333	89.7	94.8	-5.1	-9.0	-1.1	0.01
	Medical staff clearly explained follow-up care	5	344	343	88.9	95.0	-6.1	-10.1	-2.1	<0.01

LCI = lower 95% confidence interval, UCI = upper 95% confidence interval.

Note: This table reports all functional status and care experience measures that have a statistically significant difference between the intervention and comparison groups at the 10% or 5% level.

The improvement indicator takes a value of 1 if a patient moved to a better functional status level after the episode (e.g., from "complete help needed" before the episode to "no help needed" after the episode) or if the patient recalled having the highest functional status prior to hospitalization and remained in that status at the time of survey response (e.g., "no help needed" both before hospitalization and after the episode). 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 status group after the episode or if the patient recalled having the lowest functional status prior to hospitalization and remained in that status at the time of the survey.

Source: Lewin analysis of beneficiary survey data for episodes that began May, June, October, and November 2015 for BPCI EIs and the matched comparison providers.

e. Conclusion

Under BPCI, SNFs treating patients discharged from an acute care hospital due to an MJRLE achieved statistically significant payment reductions relative to a comparison group, which is the result of lower SNF payments. BPCI-participating SNFs had lower SNF payments due to reduced days in the SNF, which were off-set in part by an increase in HHA visits. This is consistent with data gathered through interviews and site visits; BPCI interviewees stated they were focused on reducing SNF length of stay and referring patients to HHAs as soon as possible to reduce their episode payments. The payment reductions were greater for patients whose MJRLE was due to a fracture than for patients with a planned (non-fracture) MJRLE, because BPCI participants managed to achieve larger reductions in SNF days for fracture patients compared to planned MJRLE procedures. Further, while BPCI non-fracture patients showed less improvement in function than comparison patients upon discharge from the SNF, overall MJRLE BPCI respondents were less likely to report a decline in their ability to walk in the beneficiary survey.

Given that SNFs have considerable discretion in who they admit to their facilities, patient selection is a concern in Model 3. There are indications that the MJRLE patients treated by BPCI-participating SNFs may have required less intensive care than prior to the initiative relative to the change in patient mix at the comparison providers, as measured by prior health care use, functional status upon admission, and other indicators. We conducted sensitivity analyses to determine if the changes observed in the patient assessment data explain the significant BPCI impacts on payments. We found that even after adjusting for functional status and other characteristics present at admission, payments still decline significantly for BPCI-participating SNFs relative to the comparison group.

4. Simple Pneumonia and Respiratory Infections, SNF

Medicare patients received 1,437 simple pneumonia and respiratory infections (SPRI) episodes of care in 230 BPCI-participating skilled nursing facilities (SNF) in the first eight quarters of the initiative (from October 2013 through September 2015). We observed no statistically significant changes in Medicare payments or utilization for BPCI episodes relative to comparison episodes. However, we observed mixed effects on quality of care. BPCI participants experienced an increase in hospital readmission and mortality rates between the baseline and intervention period relative to the comparison group. The subset of BPCI patients who had at least two SNF assessments (79% of all SPRI patients), however, experienced relative improvements in functional status. These results suggest that for those patients who remained in the SNF long enough to have two assessments, BPCI had a positive impact on their functional status.

a. Participants

During the first eight quarters of the BPCI initiative, 230 SNFs (26% of all BPCI-participating SNFs) participated in the SPRI episode, 169 of which chose 90-day episodes. Thirty-eight SNFs opted to participate in 30-day episodes and the remaining 23 SNFs participated in 60-day episodes. There were 1,437 SPRI episodes initiated by the BPCI-participating SNFs during the BPCI initiative (approximately 7% of episodes across the 48 BPCI clinical episodes). The SPRI results are based on an average of three quarters of experience because 96 SNFs began participating in SPRI in Q3 2015; 181 (79%) joined in either Q2 or Q3 2015. Furthermore, 30 SNFs (13% of SPRI EIs) terminated their participation in SPRI episode within the first eight quarters of the initiative.

Compared to SNFs that did not participate in BPCI, SNFs that participated in the SPRI episode were larger and more likely to be for-profit facilities, located in urban locations, and had a higher nursing home overall score (Exhibits 219a & 219b). The participating SNFs were less likely to be government owned or hospital-based. Further, prior to joining BPCI, participating SNFs had more admissions for SPRI.

Matched comparison SNFs were identified for 159 of the 230 BPCI-participating SNFs.[27] The matched BPCI-participating SNFs were not statistically different from the comparison SNFs in any measure (Exhibits 219a & 219b).

[27] Each BPCI-participating SNF was matched with up to 15 comparison SNFs based on a propensity score model that considered market and SNF characteristics, as well as baseline outcomes. There were 71 participants that were not included in the analysis because they either were only enrolled in Phase 2 of BPCI for one-quarter before they stopped participating in the SPRI episode or had fewer than five relevant discharges in 2011.

Exhibits 219a & 219b: Characteristics of BPCI-participating SNFs with Simple Pneumonia and Respiratory Infection (SPRI) Episodes, compared with Non-participating SNFs, All BPCI-participating SNFs, and Comparison SNFs, Model 3, Q4 2013 - Q3 2015

Characteristic		All SPRI BPCI SNFs (N=230)**		Non-participating SNFs (N=13,302)		All BPCI SNFs (N=864)		Matched SPRI BPCI SNFs (N=159)		Comparison SNFs (N=2,378++)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	42	18%	3,311	25%	121	14%	30	19%	422	18%
	Government	2	1%	617	5%*	3	0%	1	1%	7	0%
	For Profit	186	81%	9,374	70%*	740	86%	128	81%	1,949	82%
Urban	Urban	201	87%	9,364	70%*	724	84%	144	91%	2,171	91%
IRF in CBSA	Yes	128	56%	7,238	54%	488	56%	94	59%	1,404	59%
Hospital-Based	Yes	2	1%	586	4%*	7	1%	1	1%	10	0%
Part of Chain	Yes	56	24%	2,946	22%	168	19%	37	23%	544	23%

Characteristic	All SPRI BPCI SNFs (N=230)**	Non-participating SNFs (N=13,302)	All BPCI SNFs (N=864)	Matched SPRI BPCI SNFs (N=159)	Comparison SNFs (N=2,378++)
	Mean	Mean	Mean	Mean	Mean
Bed Count	128	112*	122	136	137
SNF Market Share	6%	6%	6%	6%	6%
Nursing Home Overall Score+	3.59	3.32*	3.45	3.65	3.66
Number of SPRI Admissions, 2011	12	8*	13	14	14
Standardized Part A Allowed Payment 90 days following SNF admission, SPRI episodes, 2011	**$24,792**	**$23,952**	**$24,529**	**$25,074**	**$25,110**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating SNFs are defined as Model 3 SNFs, Q4 2013 – Q3 2015. Non-participating SNFs are all other SNFs.

*Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" SNFs are compared to "All SPRI BPCI" SNFs. "Comparison" SNFs are compared to "Matched SPRI BPCI" SNFs.

** Please note that BPCI-participating SNFs that received Medicare certification after 2011 are not included in this table.

+ Number of points out of 5 in overall rating and in three areas: Quality, Survey/Health Inspections, and Staffing.

++ This column represents the weighted number of matched comparison providers to take into account the fact that some are matched to multiple participants. The distinct number of matched comparison providers is 1,770.

b. Change in Patient Mix

We monitor changes in patient characteristics because BPCI participants have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce their episode payments below their target. Additionally, SNFs have considerable discretion to determine who they admit to their facilities. While the impact analysis on payment, utilization, and quality controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix. There is some evidence that BPCI SPRI patients were *less* healthy during the

intervention relative to the comparison group. The BPCI patient population was younger (there was a relative increase in patients aged 20-64, p=0.06) and more likely to be eligible for Medicare due to disability (and not due to end stage renal disease) (p=0.07), but less likely to use the emergency room in the six months prior to the qualifying hospitalization (p=0.02) after joining BPCI relative to the comparison group (Exhibits 220a & 220b).

A comparison of patient characteristics available in the initial SNF patient assessment (the five-day Minimum Data Set assessment) also suggest that BPCI participants treated relatively less healthy patients after joining BPCI (Exhibit 221). BPCI-participating SNFs had relative increases from the baseline to intervention periods in the proportions of SPRI patients who sometimes, rarely, or never understood messages (5.2 percentage points, p<0.01); had moderate to severe cognitive impairment (8.4 percentage points, p<0.01); were rarely or unable to communicate (3.9 percentage points, p=0.01); needed extensive assistance or were totally dependent transferring (5.1 percentage points, p=0.02); frequently or always had urinary incontinence (7.2 percentage points, p=0.01); frequently or always had bowel incontinence (11.0 percentage points, p<0.01); and had an active diagnosis of dementia (7.8 percentage points, p<0.01).

Exhibits 220a & 220b: Characteristics of BPCI and Comparison Patients with a SNF Admission for Simple Pneumonia and Respiratory Infection (SPRI) Episodes, Model 3, Q4 2011 - Q3 2015

Characteristic		BPCI SPRI SNFs				Comparison SPRI SNFs				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Age	20-64	60	4%	68	6%	63	4%	46	4%	2.0	-0.1	4.2	0.06
	65-79	420	26%	308	26%	390	24%	306	26%	-1.7	-6.3	2.9	0.47
	80+	1,125	70%	816	68%	1,152	72%	840	70%	-0.3	-5.2	4.5	0.89
Gender	Female	894	56%	656	55%	930	58%	677	57%	0.5	-4.8	5.7	0.86
Medicaid and Disability	% Eligible for Medicaid	372	23%	213	18%	439	27%	272	23%	-0.8	-5.2	3.7	0.73
	% Disability, no ESRD	76	5%	79	7%	73	5%	51	4%	2.2	-0.1	4.5	0.07
Utilization Six Months Prior to Qualifying Hospitalization	Inpatient Acute Care Hospital	728	45%	565	47%	757	47%	537	45%	4.2	-1.1	9.4	0.12
	Emergency Room Admission	615	38%	411	34%	574	36%	456	38%	-6.3	-11.4	-1.2	0.02
	Home Health	586	37%	482	40%	540	34%	434	36%	1.2	-3.9	6.3	0.66
	Inpatient Rehabilitation Facility	65	4%	48	4%	54	3%	34	3%	0.5	-1.5	2.5	0.63
	Skilled Nursing Facility	502	31%	359	30%	539	34%	361	30%	2.1	-2.8	7.1	0.39
	Psychiatric Hospital	22	1%	20	2%	25	2%	17	1%	0.4	-0.9	1.7	0.50
	Long-term Care Hospital	21	1%	18	2%	25	2%	14	1%	0.6	-0.7	1.8	0.35
	Institutional Nursing Facility**	919	57%	561	47%	966	60%	584	49%	1.0	-4.3	6.3	0.71
	No Institutional Care	801	50%	591	50%	778	48%	606	51%	-2.7	-8.0	2.6	0.32
	No Post-acute Care	470	29%	360	30%	487	30%	368	31%	0.4	-4.5	5.3	0.88

Characteristic	BPCI SNFs				Comparison SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		LCI	UCI	
Average Count of HCCs	3.40	2.74	3.60	2.96	3.39	2.72	3.50	2.72	0.10	-0.20	0.39	0.52

DiD = difference-in-differences; ESRD = end stage renal disease; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

** Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Notes: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

Exhibit 221: Additional Characteristics of BPCI and Comparison Patients with a SNF Admission for Simple Pneumonia or Respiratory Infection (SPRI) Episodes, Model 3, Q4 2013 - Q3 2015

Characteristic*	BPCI SPRI SNFs				Comparison SPRI SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Not currently married	1,004	67%	757	67%	1,040	69%	771	68%	0.4	-4.7	5.5	0.86
Intellectual or developmental disability (ID/DD)	6	1%	2	0%	8	1%	2	0%	0.3	-0.7	1.3	0.57
Admitted from nursing home, swing bed, psychiatric hospital, ID/DD facility, or hospice	20	1%	5	0%	17	1%	13	1%	-0.9	-2.0	0.2	0.10
Sometimes, rarely, or never understands messages	151	10%	131	12%	222	15%	125	11%	5.2	1.7	8.7	0.00
Moderate to severe cognitive impairment	584	45%	426	44%	679	53%	446	44%	8.4	2.6	14.2	0.00
Rarely or never able to communicate	109	7%	69	6%	192	13%	85	7%	3.9	0.9	6.9	0.01
Moderate to severe depression	101	7%	40	4%	129	9%	60	5%	0.0	-2.7	2.6	0.99
Rejected necessary evaluation or care at least once	99	6%	58	5%	128	8%	59	5%	1.7	-0.9	4.4	0.20
Need extensive assistance or are totally dependent moving in bed	1,211	79%	969	83%	1,221	79%	938	81%	2.0	-2.2	6.3	0.35
Need extensive assistance or are totally dependent transferring, e.g., between bed and wheelchair	1,190	78%	973	84%	1,221	79%	923	80%	5.1	0.8	9.4	0.02
Need extensive assistance or are totally dependent walking in room	1,067	71%	850	74%	1,109	72%	837	73%	2.7	-2.2	7.5	0.28
Need extensive assistance or are totally dependent using the toilet	1,245	82%	992	85%	1,272	83%	963	84%	3.1	-1.0	7.1	0.14
Do not believe they will improve enough to require less assistance with ADL in the future	134	14%	74	10%	150	18%	91	13%	0.5	-4.2	5.2	0.83
Frequently or always had urinary incontinence	614	40%	510	44%	671	44%	462	40%	7.2	1.9	12.6	0.01
Frequently or always had bowel incontinence	577	38%	511	44%	650	42%	433	38%	11.0	5.7	16.3	0.00
Active diagnosis of heart failure	440	29%	381	33%	436	28%	362	31%	0.9	-4.0	5.9	0.71
Active diagnosis of Alzheimer's	93	6%	59	5%	124	8%	57	5%	2.1	-0.5	4.7	0.11
Active diagnosis of dementia	402	26%	338	29%	466	30%	292	25%	7.8	2.9	12.6	0.00

Characteristic*	BPCI SPRI SNFs				Comparison SPRI SNFs				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Count	%	Count	%	Count	%	Count	%		LCI	UCI	
Active diagnosis of asthma, COPD, or chronic lung disease	632	41%	479	41%	631	41%	461	40%	0.9	-4.4	6.2	0.74
Short of breath due to exertion (e.g., walking, bathing, or transferring), sitting at rest, or lying flat	541	36%	376	33%	627	41%	417	36%	1.7	-3.5	6.9	0.52
Experienced fracture due to a fall**	83	6%	61	6%	95	7%	71	7%	0.3	-2.6	3.2	0.84
Require special treatment+	876	57%	639	55%	899	59%	639	55%	0.7	-4.7	6.0	0.80

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Note: "Count" represents the numerator for the given characteristic. The "%" is the numerator divided by the total number of episodes with non-missing values.

*During 7-day lookback period (3rd week of stay) unless otherwise noted.

**Within 6 months prior to admission

+Examples include chemotherapy, oxygen therapy, and transfusions; measured after SNF admission using the five-day PPS assessment (regardless of resident status).

Source: Lewin analysis of SNF Minimum Data Set (MDS) assessment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

c. Medicare Payments and Utilization

Total Medicare allowed payments ($2015) for services included in the 90-day bundle did not significantly change for BPCI SNFs relative to the comparison group (Exhibit 222).[28] Similarly, the BPCI SNFs did not have a statistically significant change in any of the payment or utilization measures estimated in this study relative to the comparison group.

Exhibit 222: Impact of BPCI on Total Allowed Payments ($2015) Included in Bundle for Simple Pneumonia and Respiratory Infection, 90-day Episodes, Model 3, Q4 2011-Q3 2015

Measure	BPCI (N=1,085)		Comparison (N=1,077)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total allowed payments included in the bundle, 90-day episodes	$29,091	$27,485	$28,255	$28,225	-$935	-$3,139	$1,269	0.41

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome. The change in total payments included in the bundle for the 59 30-day BPCI intervention episodes and the 46 60-day BPCI intervention episodes was also not statistically significant relative to the comparison group.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

d. Quality of Care

BPCI-participating SNFs achieved mixed results with respect to quality outcomes relative to the comparison SNFs. BPCI SPRI patients experienced relative increases in unplanned hospital readmissions and mortality rates. At the same time, the subset of SPRI BPCI patients who had at least two patient assessments experienced relative improvements in measures of functional status.

BPCI SPRI patients experienced an increase of 4.6 percentage points in unplanned readmissions during the first 30 days of the episode relative to the comparison group (p=0.03) (Exhibit 223). Among BPCI patients, the unplanned readmission rate during the first 30 days of the episode increased from 20.4% to 21.8% between the baseline and intervention periods, while comparison group SPRI patients experienced a decline from 23.7% to 20.4%. The unplanned readmission rate within 90 days from the start of the episode declined from 37.1% to 36.2% for BPCI patients and from 40.1% to 35.6% for comparison group patients, however because we rejected the null hypothesis that BPCI and comparison SNFs had parallel trends for this outcome during the baseline period (with 90% confidence), this estimate may be biased. Therefore, it is not clear whether the relative increase present in the 30-day measure persists for the 90-day measure.

[28] These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Exhibit 223: Impact of BPCI on Unplanned Hospital Readmission Rates for Simple Pneumonia and Respiratory Infection Episodes, Model 3, Q4 2011-Q3 2015

Outcome	BPCI (N= 1,181)		Comparison (N= 1,187)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
30-Day Unplanned Hospital Readmission Rate	20.4%	21.8%	23.7%	20.4%	4.6	0.4	8.8	0.03
90-Day Unplanned Hospital Readmission Rate	37.1%	36.2%	40.1%	35.6%	3.6*	-1.2	8.3	-0.14

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

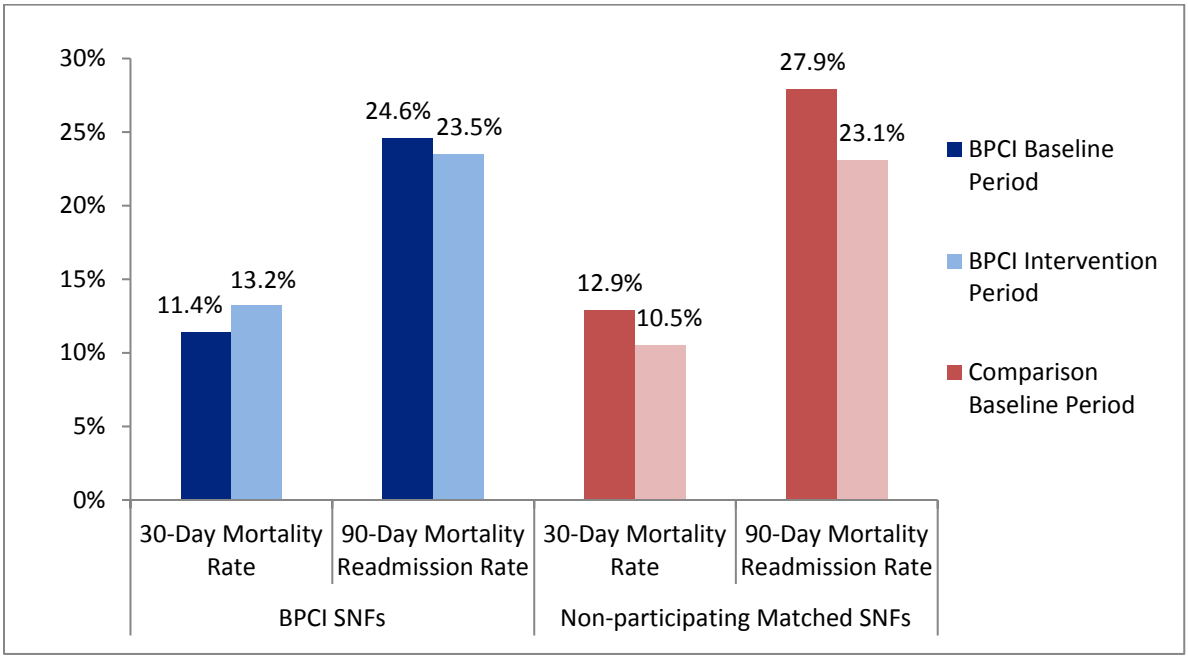
This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q2 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

In addition, there was an increase in 30-day and 90-day mortality rates for BPCI episodes relative to comparison episodes. The 30-day mortality rate increased 4.2 percentage points (p=0.02) and the 90-day rate increased 3.8 percentage points (p=0.09). Among patients in BPCI-participating SNFs, the mortality rate within the first 30 days of the episode increased from 11.4% to 13.2%, while it declined from 12.9% to 10.5% for the comparison group. For BPCI episodes, the 90-day mortality rate declined from 24.6% to 23.5% and it declined from 27.9% to 23.1% for comparison episodes (Exhibit 224).

Exhibit 224: Impact of BPCI on Mortality Rates for Simple Pneumonia and Respiratory Infection Episodes, Model 3, Q4 2011 - Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI SNFs and the matched comparison providers.

Change in beneficiary function was calculated based on the change between two SNF patient assessments and indicated that BPCI SPRI patients experienced improvements in functional status relative to the comparison group (Exhibit 225). The proportion of BPCI patients with an improvement in self-care function during the SNF stay increased by 11.1 percentage points ($p<0.01$) relative to the comparison group, and improvement in mobility function increased by 8.3 percentage points ($p=0.01$). The proportion of BPCI patients with an improvement in self-care increased from 34.5% in the baseline period to 43.0% in the intervention period, whereas the comparison group declined from 31.8% to 29.3%. The proportion of BPCI patients with an improvement in mobility increased from 51.1% to 60.0% between the baseline and intervention periods, whereas the comparison remained the same around 46.1% (Exhibit 225). Although not significant, BPCI patients with improved overall function increased relative to the comparison group by 5.1 percentage points ($p=0.10$).

The evidence on the impact of BPCI on quality of care is inconsistent. While the relative increases in mortality and readmission rates for BPCI patients may suggest a decline in quality, the relative increase in the proportion of BPCI patients with improvements in functional status suggests otherwise. The functional status measures are based on patients who had at least two assessments during their SNF stay (79% of SPRI patients), so the measures do not include patients who died during their SNF stay or were discharged before they received the second assessment. As a result, the observed mixed effects in quality may be due to a different patient sample.

Exhibit 225: Impact of BPCI on Activities of Daily Living for Simple Pneumonia and Respiratory Infection Episodes, Model 3 SNF, Q3 2013-Q3 2015

Measure	Number of intervention episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Patients with improved overall function	925	923	57.4%	62.0%	51.1%	50.6%	5.1	-1.0	11.3	0.10
Patients with improved self-care function	935	923	34.5%	43.0%	31.8%	29.3%	11.1	5.2	16.9	<0.01
Patients with improved mobility function	916	922	51.1%	59.6%	46.1%	46.3%	8.3	1.9	14.7	0.01

DiD = difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

Source: Lewin analysis of Medicare Minimum Data Set (MDS) and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Note: "Improved" means improvement or staying in the best possible status.

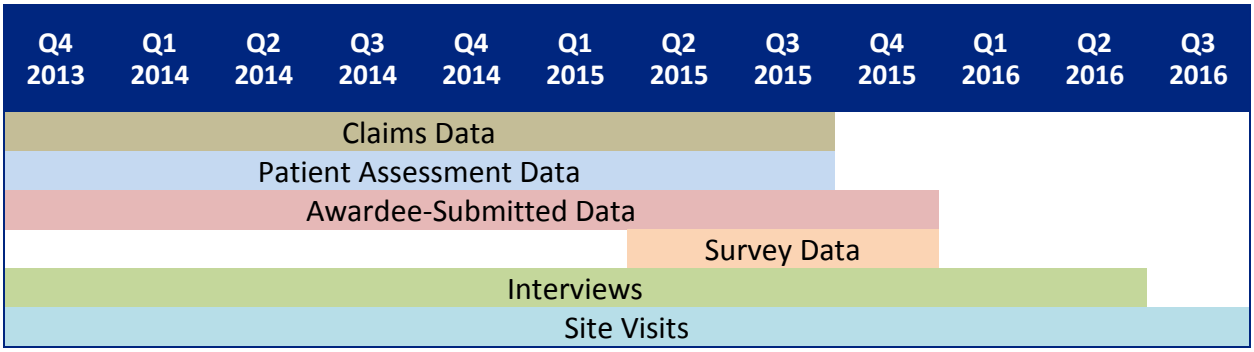
e. Conclusion

Results from the first eight quarters indicate that the BPCI initiative had no statistically significant effect on payments or utilization for SPRI patients with episodes initiated in a SNF. However, we observed mixed effects on quality of care. BPCI participants experienced a relative increase in readmission and mortality rates. We also observed that Model 3 SNF SPRI patients were *less* healthy during the intervention period compared with the baseline, relative to the change for the comparison group. However, even when we accounted for functional status and other characteristics present at admission, the relative decline in quality outcomes remained. In contrast, for the three-quarters of patients who had at least two assessments during their SNF stay we observed relative improvements in the functional status measures. It is unclear what may be driving these statistically significant changes in quality outcomes. We will monitor if these outcomes are consistent over time and will include them in the next annual report with an additional three quarters of experience, which may provide more insights into the impact of BPCI.

V. Model 4 Results and Discussion

This section presents information about the experience of Model 4 BPCI participants and their episodes of care. Exhibit 226 presents the Phase 2 BPCI time period reflected in each of the data sources used throughout the Model 4 results section. The quantitative outcomes are risk adjusted, as described in Section II.D.2.

Exhibit 226: BPCI Quantitative and Qualitative Data Sources Used for Model 4 Results



Note: The risk-bearing phase (Phase 2) of BPCI began Q4 2013.

A. Characteristics of the initiative and participants

This section summarizes the characteristics of the BPCI participants during the first two years of the initiative. Where relevant, we summarized the qualitative results from the last BPCI evaluation annual report in a call-out box in the beginning of a section. The narrative that follows the call-out box complements those results with additional insights gathered through the last year's qualitative data collection and analysis.

1. Key takeaways on Model 4 characteristics of the initiative and participants

- 16 Awardees that represented 23 acute care hospital (ACH) episode initiators (EIs) were in the risk-bearing phase of Model 4 of BPCI. These participants initiated over 7,000 episodes of care during the first two years of the initiative.
- BPCI-participating hospital EIs were less likely to be part of a chain, but more likely to be larger, to be non-profit, to be located in urban areas, and to have a smaller proportion of total inpatient days attributed to Medicare patients compared to non-participating hospitals. They also had a lower disproportionate share percentage on average, indicating fewer Medicare Supplemental Security Income, Medicaid, or other low-income patients were admitted to BPCI Model 4 hospital EIs compared to non-participating hospitals.
- BPCI-participating hospitals were located in larger, more urban and more competitive markets, with fewer skilled nursing facility (SNF) beds per 10,000 residents than markets with no Model 2 or 4 BPCI hospitals.
- Seventy-five percent of Model 4 Awardees entered into gainsharing agreements between Q4 2013 and Q2 2015 and three Awardees distributed approximately $904,000 in internal cost savings (ICS) to their gainsharing partners through Q1 2016. No EIs provided beneficiary incentives.

- Because Model 4 participants are financially responsible for readmissions, they expressed the importance of developing relationships with post-acute care (PAC) providers to enhance communication and coordination. Working with primary care providers (PCPs) was also noted as important for ensuring patient follow-up.
- Interviewees indicated that it was helpful to work with external contractors to provide data analysis and to pay claims.
- Interviewees reported providing patient education, including formal pre-operative classes on specific procedures, informal discussions about PAC options, and written educational materials to set patient expectations and build relationships with patients.
- Interviewees reported standardizing care protocols and surgical devices to achieve cost savings.
- Interviewees noted several challenges facing Model 4 participants, such as accurately identifying BPCI patients and sharing data. It was also difficult to get SNFs to reduce lengths of stay because they are paid by Medicare on a per diem basis.
- Interviewees noted that patients who are discharged earlier than Medicare's geometric mean length of stay are not included in their BPCI episodes. Interviewees stated that this results in the inclusion of higher acuity, higher cost patients in a BPCI episode, making it more difficult to achieve cost savings.
- During the first two years of the BPCI initiative, the majority of Model 4 hospitals (61%) have withdrawn from BPCI.

2. Participants

a. Entry decisions

Interview Findings Reported Previously

- Interviewees joined BPCI:
  - **For the financial opportunities available, including lowering the cost of care and gainsharing.**
  - **To align incentives among practitioners.**
  - **For the opportunity to learn and gain experience with bundled payment models.**
- Interviewees commonly reported that they did not select Model 2 because of the need to establish strong relationships with PAC providers in Model 2. Interviewees noted the decision to leverage their strengths, rather than to develop new relationships.

One Model 4 interviewee mentioned joining BPCI for the opportunity to learn about improving patient hand-offs and payment reform. This interviewee also joined BPCI because their physicians were supportive of the initiative. Respondents indicated that Model 4 was selected because they were uncertain of their ability to effectively manage PAC utilization. Another hospital respondent stated that it selected Model 4 because it had been participating in a gainsharing demonstration project in its state, which was the precursor to BPCI's Model 1. Interviewees stated that Model 4 was selected over Model 2 because they operate in an integrated system, which meant that reducing costs in the PAC portion of the clinical episode would have resulted in reduced revenues to its SNF and HHA.

b. Participant characteristics

This section discusses the characteristics of the Model 4 BPCI-participating hospitals that joined the initiative through Q3 2015. There were 16 Model 4 Awardees with 23 EI hospitals that participated in BPCI at some point during the first two years of the initiative, but 14 of the 23 EIs withdrew from BPCI prior to September 2015.

Exhibits 227a & 227b compare Model 4 BPCI-participating hospital EIs that joined between Q4 2013 and Q3 2015 to non-participating hospitals. In regard to ownership, 70% of BPCI-participating hospitals in Model 4 were non-profit compared to 57% of non-participants. A lower proportion of Model 4 participating hospitals were part of a chain (43%) than non-participating hospitals (53%).

Nearly all Model 4 BPCI-participating hospitals were located in urban areas (91%), compared with 69% of non-participating hospitals. The average bed count for participating hospitals was more than double that of non-participants (405 vs. 175). Participating hospitals had a higher average resident-to-bed ratio than non-participating hospitals (0.14 vs. 0.05) and over twice as many admissions during 2011 for MS-DRGs included in BPCI episodes (3,460 vs. 1,598).

Participating hospitals had a smaller proportion of total inpatient days attributed to Medicare patients compared to non-participating hospitals (31% vs. 42%). BPCI-participating hospitals also had a lower disproportionate share percentage on average than non-participating hospitals (21% vs. 29%), indicating proportionally fewer Medicare Supplemental Security Income, Medicaid, or other low-income patients were admitted in BPCI Model 4 hospital EIs.

Exhibits 227a & 227b: Baseline Characteristics of BPCI-participating Hospital EIs and Non-participating Hospitals, Model 4, Q4 2013 - Q3 2015

Characteristic		BPCI Hospital EIs (N=23)		Non-participating Hospitals (N=2,971)	
		N	%	N	%
Ownership	For Profit	6	26%	638	23%
	Government	1	4%	542	20%
	Non-Profit	16	70%	1,594	57%
Urban/Rural	Rural	2	9%	872	31%
	Urban	21	91%	1,902	69%
Part of Chain	Yes	10	43%	1,469	53%

Characteristic	BPCI Hospital EIs (N=23)	Non-participating Hospitals (N=2,971)
	Mean	Mean
Bed Count	405	175
Number of Admissions for BPCI Episode MS-DRGs, 2011	3,460	1,598
Medicare Days Percent	31%	42%
Resident-bed ratio	0.14	0.05
Disproportionate Share Percent	21%	29%

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 2015. Non-participating hospitals are all other hospitals not participating in any BPCI initiative that reported values for all measures listed above and are not from Maryland. Please note that BPCI-participating Model 4 hospital EIs that received Medicare certification after 2011 are not included in this table.

Prior to joining BPCI, 2011 standardized Part A payments during the inpatient stay and the 90-day post-discharge period (PDP) for patients discharged after a coronary artery bypass graft were 5% higher among Model 4 BPCI-participating hospital EIs than non-participating hospitals (Exhibit 227c). The 2011 standardized Part A payments for patients discharged after a major joint replacement of the lower extremity (MJRLE), however, were 3% lower in Model 4 BPCI-participating hospitals than non-participating hospitals.

Exhibit 227c: 2011 Standardized Part A Payments for the Inpatient Stay and the 90 day PDP, BPCI-participating Hospital EIs and Non-participating Hospitals, by Clinical Episode,[1] Model 4, Q4 2013 – Q3 2015

Clinical Episode	BPCI Hospital EIs			Non-participating Hospitals		
	Number of Hospital EIs	Number of Discharges	Mean Payment Across EIs	Number of Hospitals	Number of Discharges	Mean Payment Across hospitals
Coronary artery bypass graft	9	620	$38,773	832	39,030	$36,975
Major joint replacement of the lower extremity	15	2,719	$22,717	2,322	289,882	$23,467

Source: Lewin analysis of 2013 Provider of Service (POS) files and 2011 Medicare claims. Standardized payments are based on 2011 claims and trended to 2015 dollars. BPCI participating hospitals are defined as Model 4 Episode Initiators, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals not participating in any BPCI initiative that reported values for all measures listed in Exhibit 227a, had at least five discharges in 2011 in the episode of relevance, and are not from Maryland. Please note that BPCI-participating Model 4 hospital EIs that received Medicare certification after 2011 are not included in this table.

[1]The clinical episodes included in this table are those that had enough sample size to include in the risk adjusted Difference-in-differences analysis presented in section V.B.

c. Participant Readiness

Prior Experience

As illustrated in Exhibit 228, Model 4 EIs reported whether they participated in care redesign or payment incentives initiatives prior to their participation in BPCI. Enhancement in care delivery was the most commonly reported care redesign experience for Model 4 EIs, reported by 87.5% of participants. The majority of Model 4 EIs also reported prior experience with system changes to support care (81.3%), the redesign of care pathways (75.0%), patient activation, engagement, and risk management (68.8%), and care coordination (68.8%).

The most common type of prior payment incentives experience among Model 4 EIs was pay-for-performance, with 81.3% of participants reporting prior participation. Conversely, only two participants reported prior experience with bundled payments (12.5%).

Exhibit 228: Care Redesign and Payment Incentives Experience, Model 4, Prior to BPCI Participation*

		Model 4 ACH EIs (N=16)	
		N	%
Prior experience in care redesign initiatives	Redesign of Care Pathways	12	75.0%
	Enhancements in Care Delivery	14	87.5%
	Patient Activation, Engagement, & Risk Management	11	68.8%
	Care Coordination	11	68.8%
	System Changes to Support Care	13	81.3%
	Other Redesign Activities	1	6.3%
Prior experience in payment incentives	Bundled Payments	2	12.5%
	Pay-for-Performance	13	81.3%
	Shared Savings	4	25.0%
	Other Payment Incentives	3	18.8%

Source: Lewin analysis of Awardee-submitted data collected February 2015 through February 2016 for Model 4 episode initiators participating in BPCI between Q4 2013-Q4 2015.

*Prior experience in care redesign and payment incentive initiatives was reported when episode initiators first became active in Phase 2 of the initiative. Therefore, episode initiators that changed models are only included in the sample for the first model they participated in. This table only includes information about episode initiators that submitted data. Episode initiators that did not submit complete data are excluded from the counts in the table.

Electronic Health Record (EHR) Use

Sixteen Model 4 participants reported use of EHRs with the following functionalities: e-Prescribing, computerized physician order entry, and discharge instructions and care summary documents. Fourteen of these EIs (87.5%) also reported having medication management and clinical decision support functionalities (Exhibit 229).

Exhibit 229: Electronic Health Record (EHR) Use, Model 4, Baseline*

Meaningful-use functionalities	Model 4 EIs (N=16)	
	N	%
EIs with an EHR	16	100.0%
Automated Quality Reporting	12	75.0%
Discharge Instructions and Care Summary Documents	16	100.0%
Medication Management	14	87.5%
e-Prescribing	16	100.0%
Computerized Physician Order Entry	16	100.0%
Clinical Decision Support	14	87.5%

Source: Lewin analysis of Awardee-submitted data collected February 2015 and February 2016 for Model 4 episode initiators participating in BPCI between Q4 2013-Q4 2015.

*EHR measures were analyzed using data from the first reporting period in which episode initiators were required to report annual measures. This table only includes information about episode initiators that submitted data. Episode initiators that did not submit complete data are excluded from the counts in the table.

d. Episode selection

The count of EIs participating in BPCI clinical episodes during the first eight quarters of the initiative is shown in Exhibit 230. At least one Model 4 EI participated in 19 of the 48 clinical episodes from Q4 2013 to Q3 2015, and the average Model 4 EI participated in 4 clinical episodes. The most popular clinical episode among Model 4 participants was MJRLE, in which 65% of EIs participated. Coronary artery bypass graft (CABG) and double joint replacement of the lower extremity were the next most common clinical episodes and were each chosen by 39% of EIs. During the first eight quarters of the initiative, only 12 clinical episodes had three or more Model 4 EIs.

Exhibit 230: Participation of Episode Initiators by Clinical Episode, Model 4, Q4 2013 - Q3 2015

Clinical Episode	Episode Initiators (N=23)	
	N	% of EIs
Major joint replacement of the lower extremity	15	65%
Coronary artery bypass graft	9	39%
Double joint replacement of the lower extremity	9	39%
Cardiac defibrillator	7	30%
Pacemaker	7	30%
Percutaneous coronary intervention	7	30%
Cardiac valve	6	26%
Pacemaker device replacement or revision	6	26%
Back & neck except spinal fusion	4	17%
Cervical spinal fusion	4	17%
Spinal fusion (non-cervical)	4	17%
Revision of the hip or knee	3	13%
Congestive heart failure	2	9%
Other knee procedures	2	9%
Combined anterior posterior spinal fusion	2	9%
Complex non-cervical spinal fusion	2	9%
Acute myocardial infarction	1	4%
Sepsis	1	4%
AICD generator or lead	1	4%

Source: Lewin Analysis of CMS' BPCI database, June 2016.

Note: The sum of the total EIs participating in each of the clinical episodes exceeds the total number of EIs because EIs can participate in more than one clinical episode.

Interview Findings Reported Previously

- Interviewees selected episodes based on the opportunity to improve quality and reduce costs.
- Interviewees contracted with consultants to identify episodes with an opportunity for savings.

In interviews, Model 4 respondents shared their rationales for selecting episodes. One respondent stated that the ability to easily identify patients who would be in the clinical episode was an important factor, which led them to choose CABG. Another Model 4 facility selected episodes on the basis of predictability of clinical outcomes, patient volume, alignment/strong relationship with physician groups, and areas of historical high performance. One interviewee specifically rejected the idea of including medical episodes because they believed such conditions would be too difficult to manage effectively.

3. Market characteristics

The Model 2 results 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.3 for a summary of the results.

4. Model incentive structure characteristics

a. Partners

Interview Findings Reported Previously

- Interviewees described several benefits to partnership, including improved communication with partners, coordination of efforts and accountability under BPCI, and alignment of physician incentives.
- Interviewees continued to partner informally with PAC providers to improve patient care under the BPCI initiative.
- Interviewees also described partnerships with contractors to conduct data analytics, administer the BPCI initiative, and distribute any gains through gainsharing.

Similar to findings reported previously, Model 4 interviewees discussed partnering with PAC providers to improve care management and collaboration. However, engagement levels among partners varied. Interviewees noted that some partners were willing to discuss care redesign activities and follow protocols to improve care, while others showed no interest in aligning incentives. Model 4 interviewees also described partnerships with third party vendors that provided data analysis on costs and quality metrics. One participant described working with a third party vendor that had prior experience with Medicare's Acute Care Episode (ACE) demonstration to help pay claims. In addition, some interviewees reported partnerships with physician practices in the community. Interviewees described relationships with physician practices as being positive, but physician engagement was not always strong.

b. Waiver use

In order to utilize waivers of certain fraud and abuse laws which allow for the distribution of beneficiary incentives and the engagement in gainsharing arrangements, Model 4 Awardees must include their plans for their waiver use in their Implementation Protocols (IP). Even if an Awardee has intended to use the waivers, its EIs may not elect to provide incentives or engage in gainsharing. Most Model 4 Awardees (75%) included gainsharing plans in their IPs, though none provided beneficiary incentives.

Gainsharing

Interview Findings Reported Previously

- Interviewees reported that gainsharing improved communication and coordination among physicians. Interviewees also noted that gainsharing could result in more comprehensive and coordinated care.

Among the 12 Awardees with gainsharing plans from Q4 2013 through Q2 2015, three distributed approximately $904,000 in internal cost savings (ICS) from Q4 2013 through Q1 2016. Physicians were the most likely gainsharing partner to receive a distribution (Exhibit 231). No Awardee shared gains with institutional PAC providers, HHAs, or other types of providers.

Exhibit 231: Gainsharing Distributions Received by Partner Type, Model 4, Q4 2013 - Q1 2016

Gainsharing Partners	Awardees that reported eligible partners of this type	Number of Partners Receiving a Distribution	Number of Partners receiving ICS
Physicians	5	91	91
PGPs	2	16	16
Hospitals	2	2	2
Institutional PAC providers	0	-	-
HHAs	0	-	-
Other	0	-	-

Source: Lewin analysis of Awardee-submitted data collected February 2015 and May 2016 for Model 4 EIs participating in BPCI between Q4 2013 – Q1 2016.

All Model 4 participants who were interviewed this past year were gainsharing with physicians or considering gainsharing with physicians. Some interviewees believed that gainsharing could influence physician behavior, while others doubted that gainsharing would motivate change. One interviewee mentioned that when applying for BPCI, they expected gainsharing to be a "carrot" to encourage physician engagement. However, this interviewee found that the extensive amount of time it took them to calculate gainsharing amounts and distribute payments to their partners weakened the influence of gainsharing on physician behavior.

5. Care redesign and cost savings

Site visits and interviews provided insights on care redesign and care coordination approaches implemented by a small sample of Model 4 interviewees while participating in BPCI. Tactics discussed included patient education and communication, coordination with PAC and PCP providers, patient risk stratification, care and device standardization, and staff changes.

Interview Findings Reported Previously

- **Care Redesign Activities:**
  - Interviewees described standardizing care protocols as a part of their care redesign, with streamlining protocols, device standardization, and creating guidelines for downstream partners (e.g. SNFs) being mentioned most often.
- **Care Coordination:**
  - Interviewees cited using case managers or care navigators to reduce readmissions. Specific roles of these staff varied by participant; however, it was noted that these staff are generally responsible for establishing relationships with beneficiaries, leading education classes prior to surgery, discussing post-discharge options, and conducting follow-up calls throughout the episode.
- **Cost Savings Strategies:**
  - Reducing readmissions was the most commonly reported cost-saving strategy among interviewees. Identifying high-risk patients and scheduling follow-up appointments were also cited as common approaches to reduce readmissions.
  - Standardizing care protocols was an additional strategy used to achieve cost savings. However, participants reported that they often encountered resistance from physicians regarding the adoption of new protocols.

All Model 4 interviewees discussed patient education during their interviews, and the majority mentioned the importance of educational materials and classes for their patients. Interviewees described several approaches to patient education: educational books for patients before admission, pre-operative classes for patients undergoing spinal and joint surgery, and inpatient educational materials. Interviewees described patient education covering all aspects of care, from pre-operative classes to discussions about PAC options. Some interviewees described sharing a preferred PAC provider list and assisting patients with their PAC facility decision, whereas others did not provide that education. One interviewee stated, "With patients, all we can do is give them a list and say 'Have at it.'" Interviewees asserted that their care redesign changes related to patient education led to increased patient awareness, assisted with setting patient expectations, and helped establish relationships between the hospital and the patient.

> "From the moment you park your car until you leave the hospital, this is what you are going to do."
>
> – Model 4 Hospital, regarding lessons shared with patients during a pre-operative class

Given that Model 4 providers are financially responsible for related readmissions, some interviewees focused their efforts on enhancing communication and coordination with PAC providers and PCPs. Towards this end, one interviewee reported that they appointed a vice president of PAC to better understand the needs of their PAC providers, while another interviewee said they implemented a new protocol that details the level of care they expect their patients to receive from their HHA partners. Both of these interviewees reported that they offered training and education to staff in the PAC settings to increase their competence in caring for their patients. One interviewee described working with PAC staff to identify the scenarios in which they should call the hospital. Finally, some interviewees reported that they worked to improve their relationships

> "I've been trying to monitor [PAC] length of stay. It's hard because we're not at-risk for PAC so it's hard for me to ask them [our PAC partners] to hand me this data every time."
>
> – Model 4 Hospital

with PCPs. For example, one interviewee noted that they helped their patients schedule post-discharge follow-up appointments with their PCPs.

Patient risk stratification was discussed with a couple of Model 4 interviewees. One interviewee described how they use risk assessment for patient optimization before surgeries, stating, "We now explain to patients that we won't do elective surgery until they improve the control of their diabetes." Another interviewee used risk assessment scores (Zwolle scores) as part of their process for determining when patients could be sent home from the hospital after acute myocardial infarctions.

As in OY1, Model 4 site visit interviewees reported that BPCI helped to accelerate the standardization of their care protocols. One aspect of care where they focused their efforts was their approach to pain management. All site visit interviewees described efforts to reduce the use of narcotics and opioids as part of their pain management protocol. For example, one interviewee reported that they reduced the use of narcotics, which reduced sedation and facilitated earlier patient ambulation. Another interviewee described a system they developed to flag patients based on their history of prescription opioid use, noting that they consult pharmacists, psychiatrists, and substance abuse specialists in the care of patients with prior opioid dependence. In addition to standardizing their pain management protocols, interviewees described their efforts to standardize pre-operative protocols to optimize their patients for surgery. Also similar to OY1, interviewees reported efforts to standardize their discharge protocols.

In addition to the standardization of care protocols, two Model 4 interviewees discussed the standardization of devices and components as a key strategy they employed to achieve cost savings. One interviewee noted that one of their practice groups was initially resistant to device standardization and that achieving cost savings varied by episode type, with orthopedics being the most successful. The other interviewee noted that providing cost data to their surgeons led to the elimination or reduction of items in default surgical trays that were rarely or never used.

Similar to OY1, Model 4 interviewees reported the need to hire additional staff, or to shift current staff responsibilities among existing clinical and non-clinical staff, as a result of their BPCI participation.

6. Implementation challenges

Interview Findings Reported Previously

- Interviewees indicated that they believed that the episode definitions were too broad and could include beneficiaries who did not fit into the new care pathways designed for that bundle.
- Interviewees noted challenges related to short stay exclusions, which precludes them from benefiting from their lower costs.
- Interviewees explained that partnering with physicians to standardize care has been difficult, due to their varying approaches to care provision.
- Another challenge described by interviewees was the inability to partner with every PAC provider to which their patients are discharged. When beneficiaries could not be accounted for during the post-discharge period, they presented a higher likelihood for readmission.
- Interviewees noted that it was difficult to track readmissions to non-affiliated hospitals due to incomplete or delayed data from CMS. It was reported that these challenges inhibited improvement efforts concerning readmissions.
- Interviewees described the need for additional staff, increased responsibilities for current staff, and more staff training related to the BPCI initiative.
- Interviewees explained the importance of educating and supporting physicians early in the gainsharing process to mitigate challenges with physicians who were reluctant to enter into agreements.

Several Model 4 site visit interviewees stated that identifying BPCI patients was a challenge, specifically ensuring that the correct MS-DRG was assigned (e.g., locating proper documentation, updating information if the patient had additional unexpected surgical procedures). Additionally, a couple of interviewees stated that many direct care staff members were not aware of which patients were in BPCI.

A majority of Model 4 site visit interviewees expressed frustration with the exclusion of patients from bundles, particularly patients that were discharged prior to the geometric mean length of stay (LOS). One awardee reported, "We didn't recognize that hundreds of our joint replacement patients would be excluded." They reported that before BPCI, they had a high-performing joint program with a high discharge-to-home rate around 80 to 85%. However, they explained that "under BPCI, if patients go home prior to the geometric mean LOS and receive community nursing services, they are excluded from the bundle. If we waited until three days after discharge to consult home health services, or if we kept patients in the hospital longer, then they would be kept in the bundle, but it wouldn't be the right thing for the patient."

Several Model 4 interviewees noted staffing challenges. In particular, staff recruitment, staff turnover, and staff resistance to change were reported among the primary challenges that participants faced in regards to their efforts to standardize care and successfully implement BPCI initiatives.

Similar to Model 2 and Model 3 interviewees, those in Model 4 experienced challenges related to partnerships. In particular, interviewees reported that PAC partners, especially SNFs, have differing incentives than participants. With the Medicare per diem payment method, PAC providers are not incentivized to decrease the length of stay. Interviewees also found that partner facilities often had differing EMRs, making it more difficult to communicate and share patient data.

7. Episode initiators that withdrew from BPCI

Interview Findings Reported Previously

Awardee interviewees said that difficulty using CMS data was a key reason for terminating their BPCI participation. This challenge resulted in administrative burden (e.g. cost of new employees, additional time of existing staff, paying physicians timely and accurately).

This section discusses characteristics of BPCI Model 4 ACHs that withdrew from BPCI during the first eight quarters (Q4 2013 through Q3 2015) of the initiative.[1] As seen in Exhibit 232, by the third quarter of 2015, 14 BPCI hospitals (61% of all Model 4 BPCI ACHs) had dropped all clinical episodes. On average, withdrawn hospitals had a significantly shorter average tenure in the BPCI initiative than BPCI hospitals that did not withdraw. Further, withdrawn hospitals were significantly more likely to operate under a convener and initiate fewer BPCI clinical episode per quarter on average than hospitals that were still active at the end of Q3 2015.

Exhibit 232: BPCI Characteristics among Active and Withdrawn BPCI EIs, Model 4, Q4 2013 – Q3 2015

	BPCI Hospital EIs that did not Withdraw (N=9)	Withdrawn BPCI Hospitals (N=14)
	Mean	Mean
Tenure in BPCI (months)	15.4	8.4*
Average number of BPCI episodes initiated per quarter, within a given clinical episode	43.6	24.4
Percentage of Hospitals operating under a Convener	11%	71%*

Source: Awardee-submitted data, claims data, and CMS' BPCI database.

Note: *Withdrawn BPCI Hospitals* include all BPCI hospitals that dropped **all** of their clinical episodes from Model 4 at some point during the first eight quarters of the initiative. This includes hospitals that switched models. *BPCI Hospital EIs that did not withdraw* include all hospitals that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative.

*Indicates statistical significance at 5% level.

Exhibits 233a, 233b, and 233c compare BPCI-participating hospitals to those that dropped all of their episodes on a number of provider and market characteristics. Hospitals that withdrew from BPCI were more likely to have an inpatient rehabilitation facility (IRF) in their market (86% vs. 44%, $p=0.04$) and higher Medicare Advantage penetration (39% vs. 27%, $p<0.01$) than hospitals that continued in BPCI. Additionally, withdrawn hospitals had fewer admissions for BPCI MS-DRGs ($p=0.03$). Withdrawn hospitals were located in more competitive markets on average (0.13 vs. 0.34, $p=0.03$) and had a smaller market share (11% vs. 35%, $p=0.02$). Hospitals that withdrew from BPCI Model 4 also had fewer Medicare inpatient days as a share of total inpatient days (25% vs. 40%, $p<0.01$) than hospitals that did not withdraw from BPCI.

[1] Withdrawn hospitals include all BPCI hospitals that dropped **all** of their clinical episodes from Model 4 at some point during the first eight quarters of the initiative. This includes hospitals that switched models. BPCI hospital EIs that did not withdraw include all hospitals that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative.

Exhibits 233a, 233b & 233c: Provider and Market Characteristics among Active and Withdrawn BPCI EIs, Model 4, Q4 2013 – Q3 2015

Hospital Characteristics		BPCI Hospital EIs that have not Withdrawn (N=9)		Withdrawn BPCI Hospitals (N=14)	
		N	%	N	%
Ownership	Non-Profit	8	89%	8	57%
	Government	0	0%	1	7%
	For-Profit	1	11%	5	36%
Urban/Rural	Urban	7	78%	14	100%
Part of Chain	Yes	4	22%	6	0%

Characteristic	BPCI Hospital EIs that have not Withdrawn (N=9)	Withdrawn BPCI Hospitals (N=14)
	Mean	Mean
Bed Count	553	309
Number of Admissions for BPCI Episode MS-DRGs, 2011	5,707	2,015*
Medicare Days Percent	40%	25%*
Resident-to-bed ratio	0.10	0.17
Disproportionate Share Percent	31%	27%

Market Characteristics	BPCI Hospital EIs that have not Withdrawn (N=9)	Withdrawn BPCI Hospitals (N=14)
	Mean	Mean
Hospital-Market Share	35%	11%*
Herfindahl Index	0.34	0.13*
Medicare Advantage Penetration	27%	39%*
PCPs per 10,000	7.3	8.0
SNF beds per 10,000	54.0	41.4
IRF in Market	44%	86%*
Population	4,632,447	3,349,417
Median Household Income	$50,702	$55,297
% Age 65+	14%	12%

Source: Lewin analysis of 2013 POS and 2011 Medicare claims, Model 4 EIs, Q4 2013 – Q3 2015.

Note: *Withdrawn Hospitals* include all BPCI hospitals that dropped **all** of their clinical episodes from Model 4 at some point during the first eight quarters of the initiative. This includes hospitals that switched models. *BPCI Hospital EIs that did not withdraw* include all hospitals that have continued participation in **at least one** clinical episode during the first eight quarters of the initiative.

*Indicates statistical significance at 5% level

B. Model 4 impact of BPCI

This section presents the estimates of the impact of BPCI Model 4 on payments, utilization, and quality of care based on episodes that were initiated by acute care hospitals during the first eight quarters of the BPCI initiative (Q4 2013 through Q3 2015).

1. Payment, utilization, quality

We present results for the key outcomes for MJRLE and CABG episodes because they were the only two Model 4 EI type/episode combinations with sufficient sample size. The exhibits present the estimated differential change in risk adjusted outcomes for patients receiving care from hospitals participating in BPCI between the baseline and the intervention period relative to the same change for patients receiving care from hospitals in a comparison group. See Section II.D.2 for additional details on the statistical approach. All claim and assessment-based results for the two Model 4 clinical episodes are in **Appendix S**.

Across both Model 4 episodes, there were statistically significant changes in key payment outcomes for MJRLE episodes. There was an increase in the HHA standardized allowed amount and a decrease in payments for outpatient therapy relative to a comparison group. There was a statistically significant increase in the percentage of BPCI MJRLE patients discharged to PAC and a decrease in the total number of institutional days relative to the comparison group. There were mixed quality of care results. BPCI hospitals participating in MJRLE saw a statistically significant decline in the readmission rate. Among patient assessment measures, there was a statistically significant increase in the improvement in self-care among IRF patients with MJRLE episodes. Among CABG episodes, there was a statistically significant increase in improvement in ADL measures, relative to a comparison group, for patients sent to home health but significant declines among patients sent to SNFs.

a. Sample characteristics

Before discussing the impact of BPCI on payments, utilization, and quality, we present some basic statistics by clinical episode in Exhibit 234 to better understand the sample included in the analysis. There are nine EIs and 742 episodes included in the CABG analysis and 14 EIs with 3,379 episodes for MJRLE. Because providers were allowed to join BPCI over an extended period, these data represent an average of 4 quarters of participation. In contrast to Models 2 and 3, there were no providers that joined Model 4 of the initiative in Q3 2015 and seven providers discontinued participation in the episode on average across the two clinical episodes analyzed.

Exhibit 234: Characteristics of the Matched BPCI Providers included in the DiD Estimates, Model 4, Q4 2013 – Q3 2015

Clinical Episodes	Matched EIs (#)	Matched Intervention Period Episodes (#)	Average length of participation (quarters)	EIs that joined in Q3 2015 (%)	EIs that joined in Q2 or Q3 2015 (%)	EIs that exited (#)	Intervention Period Episodes from EIs that exited (%)
Coronary artery bypass graft	9	742	4	0%	0%	5	5%
Major joint replacement of lower extremity	14	3,379	4	0%	7%	9	27%

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2013 through Q3 2015 for BPCI providers. This table is limited to the matched BPCI providers used to calculate the difference-in-differences (DiD) results in the remainder of this section.

b. Key payment, utilization, and quality of care outcomes

How have the average standardized allowed amounts (Medicare payments and coinsurance/copayments combined) changed under BPCI?

There were no statistically significant differences in the relative change in total allowed payment amounts within the bundle or through the 90 days following hospital discharge.[2] The only change was a statistically significant increase in the HHA standardized allowed amount for MJRLE episodes (p=0.02) (Exhibit 235). As Model 4 only includes the inpatient stay and related readmissions within the bundle period, the lack of statistically significant differences in PAC spending during the 90-day PDP is expected. While Awardees most frequently reported reducing readmissions as a cost-saving strategy in Model 4, only one of two clinical episodes actually reduced readmissions during the 90 days post discharge relative to the comparison group.

[2] The payment outcomes are based on Medicare's fee-for-service (FFS) payment amounts, not the prospective rate Model 4 participants are paid under BPCI. The analysis uses Medicare FFS payments rather than the discounted, prospective rate for the anchor stay and related Part B services and related readmissions. This allows us to isolate the impact of BPCI on changes in service use and spending.

Exhibit 235: Impact of BPCI on Allowed Payment Outcomes, by Clinical Episode, Model 4, Baseline to Intervention, Q4 2013-Q3 2015

Clinical Episode	Number of Episodes Q4 2013- Q3 2015	Total amount included in bundle definition[1]	Total allowed payment amount, IP through 90-day PDP	Readmissions Standardized Allowed Amount, 90-day PDP	SNF Standardized Allowed Amount, 90-day PDP[2]	HHA Standardized Allowed Amount, 90-day PDP[2]	IRF Standardized Allowed Amount, 90-day PDP[2]
Coronary Artery Bypass Graft	742	-$2	-$1,151	$200	$221	$13	$443
Major Joint Replacement of the Lower Extremity	3,379	-$266	$226	-$82	$20	$316	$99

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively.). Medicare payment outcomes are standardized to remove the effect of geographic and other adjustments and are trended to 2015. PDP = post-discharge period.

[1] The total amount included in bundle definition values are based on only the 90-day episodes.

[2] These payment measures are not conditional upon use of the service.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How have the services changed under BPCI?

The Model 4 bundle includes the inpatient stay and related readmissions within 30 days of discharge from the anchor hospital. There was an increase in the proportion of MJRLE patients discharged to PAC ($p<0.01$) and a decline in institutional days post-discharge ($p=0.08$) (Exhibit 236).

Exhibit 236: Impact of BPCI on Utilization Outcomes, by Clinical Episode, Model 4, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of Episodes Q4 2013- Q3 2015	% discharged to an institution out of those who received any PAC	% discharged to post-acute care	Acute Inpatient Care LOS	Number of HH Visits, 90-day PDP[1]	Number of SNF Days, 90-day PDP[2]	Number of Institutional Days, 90-day PDP[3]
Coronary Artery Bypass Graft	742	3.5	1.3	0.0	0.6	4.0	1.26
Major Joint Replacement of the Lower Extremity	3,379	4.0	8.9	-0.1	0.8	-1.4	-1.56

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively.. PDP = post-discharge period.

[1] Beneficiaries must have spent a minimum of one day in HH care during the 90-day PDP to be included in the DiD estimate.

[2] Beneficiaries must have spent a minimum of one day in a SNF during the 90-day PDP to be included in the DiD estimate.

[3] Beneficiaries must have spent a minimum of one day in a SNF, IRF, or LTCH during the 90-day PDP to be included in the DiD estimate.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

There were no statistically significant changes in pre-bundle payments, post-bundle payments, or total payments included in the bundle definition, for CABG or MJRLE episodes under Model 4 (Exhibit 237).

Exhibit 237: Impact of BPCI on Allowed Payment Outcomes outside of the bundle, by Clinical Episode, Model 4, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of Episodes Q4 2013- Q3 2015	Total amount not included in bundle definition[1]	Part B, 30-day pre-bundle period	Part A & Part B, inpatient hospital, days 1-30 PBP	Other Part A & Part B, days 1-30 PBP	Part A & Part B, inpatient hospital, days 1-90 PBP[2]	Other Part A & Part B, days 1-90 PBP[2]
Coronary Artery Bypass Graft	742	$717	-$268		$233	-$394	$344
Major Joint Replacement of the Lower Extremity	3,379	$421	$38	-$4	-$139	$18	-$97

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. 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. PBP = Post-Bundle Period.

[1] Other Part A and Part B not included in bundle definition (total not included in the bundle definition minus excluded readmissions).

[2] Please note these outcomes are reported with a one-quarter delay (Q4 2013 through Q2 2015).

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

How has quality of care changed under BPCI?

There were no statistically significant changes in claims based quality measures for CABG or MJRLE episodes except for a decline in the unplanned readmission rate for MJRLE (-2.5 pp, p=0.03) (Exhibit 238). Among the approximately three-quarters of MJRLE patients discharged to a SNF or HHA who had two patient assessments, there was no statistically significant change in the proportion of patients with improved ADLs, though there was a statistically significant improvement in the self-care score among patients discharged to an IRF (1.7 pp, p<0.01). The ADL results for CABG patients discharged to SNF or HHA were mixed, with increased improvements in both lower body dressing (8.6 pp, p=0.01) among those discharged to an HHA and decreased improvements in self-care function (-24.4 pp, p=0.06) among those discharged to a SNF. We also found that among patients discharged to an HHA there was a relative increase of 7.5 percentage points (p=0.02) of BPCI patients who had an improvement in upper body dressing, although this estimate may be biased as we rejected the null hypothesis that the BPCI and comparison groups had parallel trends during the baseline time period.

Exhibit 238: Impact of BPCI on Claim-based and Assessment-based Quality Outcomes, by Clinical Episode, Model 4, Baseline to Intervention, Q4 2013 - Q3 2015

Clinical Episode	Number of Episodes Q4 2013-Q3 2015	All-cause mortality rate, 90-day PDP	Emergency Department Use, 90-day PDP	Unplanned readmission rate, 90-day PDP	ADL, HHA, improved ambulation	ADL, HHA, improved bathing	ADL, HHA, improved bed transferring	ADL, HHA, improved lower body dressing	ADL, HHA, improved upper body dressing	ADL, SNF, improved mobility function	ADL, SNF, improved overall function	ADL, SNF, improved self-care function	ADL, IRF, change in self-care score[1]	ADL, IRF, change in mobility score[1]
Coronary Artery Bypass Graft	742	-0.2 pp**	-2.0 pp	0.7 pp	-0.4 pp	-6.2 pp	-3.5 pp	8.6 pp	7.5* pp	-11.3 pp	-21.5 pp	-24.4 pp	-0.2 pp	1.2 pp
Major Joint Replacement of the Lower Extremity	3,379	-0.1 pp		-2.5 pp	-2.1 pp	-1.0 pp	4.3 pp	0.3 pp	-2.4 pp	-3.2 pp	0.8 pp	-4.8 pp	1.7 pp	0.6 pp

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. PDP = post-discharge period.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate. Equal trends test was conducted for ED, readmission, and mortality outcomes.

**There was insufficient sample during the baseline period to test the null hypothesis that the BPCI and comparison providers had parallel trends for this outcome.

[1] Change in IRF ADL is measured from admission to discharge, and a positive value represents an improvement in self-care or mobility.

2. Patient functional status, health status, and health care experience

Do patients treated under BPCI report differences in changes of functional status, health status, and health care experience compared with similar patients who were not in BPCI?

We surveyed beneficiaries treated by Model 4 ACH EIs in Waves 4 and 5, along with matched comparison beneficiaries. We combined responses from Waves 4 and 5 to achieve the number of complete responses (310) that power analyses indicated would be required for precise analysis. Due to the small number of Model 4 EIs and episodes, we pooled all 16 clinical episode types in which Model 4 participants were enrolled. Respondents with a MJRLE episode represented nearly half (48.9%) of all Model 4 respondents. Two other episodes comprised more than 10% of the Model 4 respondents: non-cervical spinal fusion (13.4%) and CABG (11.0%).

Respondents were asked to report on their functional status before and after the clinical episode, as well as their overall health status, and their health care experiences. We present risk–adjusted results reflecting differences in outcomes between respondents receiving care from a BPCI EI relative to respondents from a matched comparison group. Estimates for the BPCI and the comparison group, along with confidence intervals for the estimated differences, are available in **Appendix P**.

Exhibits 239-241 present point estimates of the differences in survey outcomes between respondents treated at BPCI Model 4 EIs and those treated at matched comparison hospitals. Measures of interest are listed in the first row. Changes in functional status (Exhibit 239) are further classified as rates of improvement and rates of decline. Results are summarized below.

Overall, compared with comparison respondents, BPCI respondents with episodes initiated by Model 4 EIs reported similar changes in five out of seven functional status measures from before to after the episode, similar mental and physical health outcomes, and similar health care experiences.

a. Improvement and decline in functional status

BPCI respondents with episodes initiated by Model 4 EIs reported similar changes in functional status before and after their episode as the matched comparison group, for five out of seven functional status measures (Exhibit 239). BPCI respondents were significantly less likely to report increased dependence on a mobility device (-9.5 percentage points, p=0.05) but were significantly more likely to report a decline in ability to independently bathe, dress, use the toilet, or eat (4.7 percentage points, p=0.09).

b. Patient reported health status

Estimated differences in depression symptoms, overall mental health and physical health between Model 4 BPCI and comparison respondents were small and statistically insignificant at the 5% and 10% level (Exhibit 240).

c. Health care experience

Model 4 BPCI respondents reported similar care experiences relative to the matched comparison group (Exhibit 241). Estimates were small in magnitude and statistically insignificant.

Exhibit 239: Improvement and decline in functional status, difference between BPCI and comparison survey respondents, Model 4, 2015

Clinical episode	Wave	N BPCI responses*	Bathing, dressing, using the toilet, or eating		Planning regular tasks		Use of mobility device		Walking without rest		Using stairs		Physical/ emotional problems limit social activities		Pain limiting regular activities	
			Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.	Imp.	Dec.
Model 4 – all episodes	5	391	-4.29	4.71	0.31	0.60	6.00		8.85	-2.09	-4.21	-3.44	0.06	-2.56	1.61	2.80

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P**.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance. *Abbreviations:* Imp., improvement in functional outcomes or stay at the best status; Dec., decline in functional outcomes or stay at the worst status.

Source: Lewin analysis of Model 4 BPCI and comparison beneficiary surveys collected in 2015.

Exhibit 240: Patient reported health outcomes, difference between BPCI and comparison survey respondents, Model 4, 2015

Clinical episode	Wave	N BPCI responses*	Composite depression binary indicator†	Self-reported physical health binary indicator‡	Self-reported mental health binary indicator‡
Model 4 – all episodes	5	391	2.39	6.61	1.79

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P**.

†The composite depression indicator is a binary measure equal to one when respondents reported a score of 3 or more on the Patient Health Questionnaire-2 (PHQ-2) and otherwise equals zero.

‡ The self-reported physical and health are binary measures equal one when respondents reported that their health was excellent, very good or good, and equal zero when respondents reported fair or poor health.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance.

Source: Lewin analysis of Model 4 BPCI and comparison beneficiary surveys collected in 2015.

Exhibit 241: Health care experience, difference between BPCI and comparison survey respondents, Model 4, 2015

Clinical episode	Wave	N BPCI responses*	Never received Conflicting advice	Always received Appropriate care	Staff always used preferred language	Discharged at right time	Preferences considered for services after discharge†	Understand care of self†	Medication instructions clearly explained†	Follow-up explained†	Able to manage health†	Satis-faction with recovery ‡
Model 4 – all episodes	5	391	3.13	3.33	-2.42	-0.77	1.87	0.15	0.91	2.49	-2.20	4.14

*The number of overall BPCI survey responses does not reflect item non-response (i.e. missing responses on a particular question) or comparison group Ns. In general, item non-response was low and the number of comparison respondents was similar to the number of BPCI respondents. Full sets of item-specific Ns can be found in **Appendix P**.

†Measure reflects that respondents either agreed or strongly agreed with this statement.

‡ Measure reflects that respondents were either "quite a bit" satisfied or "extremely" satisfied with their recovery.

Notes: Results are presented in terms of percentage point differences between BPCI and comparison respondents. Statistically significant results indicating better outcomes among BPCI respondents are shaded green, while significant results indicating worse outcomes among BPCI respondents are shaded orange. The lighter shade indicates 10% significance, and the darker shade indicates 5% significance.

Source: Lewin analysis of Model 3 BPCI and comparison beneficiary surveys collected in 2015.

3. Change in patient mix

To examine changes in patient mix, we investigated selected patient characteristics in the baseline and intervention periods for BPCI-participating hospitals compared to the matched comparison group of hospitals.

There were several statistically significant changes in patient characteristics that suggest CABG patients at Model 4 participating hospitals were healthier during the intervention period than in the baseline relative to the change in patients for comparison hospitals (Exhibit 242). BPCI-participating hospitals had a relative decline in the proportion of CABG patients who were eligible for Medicaid or who had prior use of either HH or IRF compared to the change for comparison providers. The BPCI-participating hospital CABG patients also had a relative decline in count of HCC indicators. Note, however, that the risk adjusted payment, quality, and utilization measures control for these variables.

There is evidence that the MJRLE patient population of BPCI-participating hospitals was less healthy after the introduction of BPCI relative to the comparison group. In the intervention period, BPCI MJRLE patients used more health care services prior to the anchor hospitalization, an indication that their health status was worse relative to the comparison group. There were statistically significant increases in the percentage of BPCI patients who had hospital, home health, IRF, or SNF care in the six months prior to their index hospitalization, and a statistically significant decrease in the percentage using no institutional care and using no PAC in the six months prior to their index hospitalization compared to the matched comparison hospitals (Exhibit 242).

Exhibit 242: Relative Changes in Claim-based Characteristics of BPCI and Matched Comparison Beneficiaries, by Clinical Episode, Baseline to Intervention, Model 4, Q4 2013– Q3 2015

	Demographics					Prior Health Status		Utilization Six Months Prior to Index Hospitalization									
Clinical Episode	Age: 20-64 years	Age: 65- 79 years	Age: 80+ years	Female	Medicaid Eligibility	Disabled, no ESRD	Count of HCC indicators	Inpatient Acute Care Hospital	Emergency Room Admission	Home Health	Inpatient Rehabilitation Facility	Skilled Nursing Facility	Psychiatric Hospital	Long-term Care Hospital	Institutional Nursing Facility	No Institutional Care	No Post-acute Care
Coronary artery bypass graft	2.4	-1.8	-0.6	2.9	-8.6	3.2	-0.3	-3.6	-1.1	-3.1	-1.2	-1.6	0.1	n/a	-1.5	3.3	2.8
Major joint replacement of lower extremity	0.0	-1.2	1.2	-2.0	0.7	0.6	0.1	2.7	-0.5	3.0	1.3	1.9	-0.2	0.0	1.3	-2.7	-3.0

Note: The estimates in this table are the results of a difference-in-differences (DiD) model. Positive DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light green shaded cells, respectively. Negative DiD estimates that are significant at the 5% or 10% significance level are indicated by dark and light orange shaded cells, respectively. DiD = difference-in-differences. ESRD = end stage renal disease.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

C. Discussion

There were 23 hospitals under 16 Awardees that participated in BPCI Model 4 during the first two years of the initiative. By September 2015, nine hospitals remained in Model 4; one EI transitioned to Model 2 and the remainder withdrew from the BPCI initiative. As of Q2 2017, only two hospitals remain in Model 4. Participants that withdrew from the initiative identified administrative burden as the key reason for withdrawal. Under Model 4, Awardees are paid a prospectively determined amount and they, in turn, pay the providers caring for beneficiaries in the episode. Hospital representatives talked about problems receiving the correct billing information from CMS and data collection burden. Hospitals also stated that excluding short stay patients, defined as an admission with a length of stay less than the geometric mean, from episodes limited their ability to succeed under this Model.

It is difficult to draw conclusions about the impact of BPCI Model 4 due to the small number of participating hospitals and episodes. There was only sufficient sample to analyze coronary artery bypass graft (CABG) and major joint replacement of the lower extremity (MJRLE) episode payment and quality outcomes and to compare self-reported satisfaction and functional status for all Model 4 clinical episodes combined. In general, BPCI Model 4 resulted in few statistically significant changes in payments, utilization, quality of care, or patient satisfaction.

Model 4 is intended to broaden hospital incentives to improve coordination of care by including physician services during the hospitalization and to reduce readmissions. BPCI provides additional tools, in particular the ability to share gains due to reductions in internal hospital costs, with physicians in order to achieve further efficiencies. Even without BPCI, however, hospitals were reducing readmissions because of the Hospital Readmissions Reduction Program and they already had incentives to control internal hospital costs because of the per discharge MS-DRG payment. The administrative requirements that hospitals essentially become the payer of physician claims for services provided to Model 4 patients, however, have been difficult to implement.

D. Clinical episode issue brief chapters

1. Coronary Artery Bypass Graft

During the first eight quarters (October 2013 through September 2015) of the Bundled Payments for Care Improvement (BPCI) initiative, there were 798 coronary artery bypass graft (CABG) episodes of care for Medicare beneficiaries in nine BPCI-participating acute care hospitals (ACH) under Model 4. Five of the Model 4-participating ACHs terminated their participation in CABG episodes after only two intervention quarters. Average Medicare payments for the hospitalization plus the 90-day post-discharge period (PDP) and payments included in the bundle did not differ between the treatment and comparison groups. Utilization measures also did not differ between the treatment and comparison groups. While quality of care was largely maintained, BPCI beneficiaries discharged with home health agency (HHA) care demonstrated relative improvements in functional status, but BPCI beneficiaries discharged to skilled nursing facilities (SNF) demonstrated relative declines in functional status. It should be noted that the assessment-based outcomes are based on the approximately three-quarters of the patients who had two patient assessments and may not be representative of all patients discharged to PAC. These changes occurred even though Model 4 hospitals have no financial incentives with respect to the post-discharge period, except to minimize readmissions.

a. Participants

During the first eight quarters of the BPCI initiative, nine hospitals (39% of all Model 4 BPCI hospitals) participated in the CABG episode. Five of them stopped participating in CABG episodes after two intervention quarters. Because participants were allowed to join BPCI over an extended period, the CABG results are based on an average of four quarters of experience. During the BPCI initiative, Model 4 participants initiated 798 CABG episodes (approximately 10% of episodes across the 48 BPCI clinical episodes in Model 4).

Compared with hospitals that did not participate in BPCI, hospitals that participated in the CABG episode under Model 4 were larger, for-profit hospitals in urban areas with larger teaching programs (measured by resident-to-bed ratios) (Exhibits 243a & 243b). Additionally, prior to joining BPCI, participating hospitals had more CABG discharges and a lower percentage of Medicare days. Model 4 hospitals participating in the CABG episode had higher average 2011standardized Part A payments during the anchor hospitalization plus the 90-day PDP.

Matched comparison hospitals were identified for seven of the nine BPCI-participating hospitals in the sample.[3] While matching reduced some of the differences, the matched BPCI-participating hospitals remained significantly different from comparison hospitals. Matched BPCI hospitals were more likely to be large, for-profit hospitals that were less likely to be part of a chain, and had a lower percentage of total inpatient days from Medicare patients. Due to the small number of Model 4 CABG participants, we focused on matching utilization and spending patterns of CABG patients. Hence, there are more likely to be differences between BPCI Model 4 CABG and comparison providers relative to BPCI Model 2 or Model 3 providers and comparison providers along the domains of provider and market characteristics. We do not know how these differences affected our results.

[3] There were two participants that were not included in the analysis because they had fewer than five relevant discharges in 2011 and 2012. Coarsened exact matching was used to match each BPCI-participating hospital with comparison hospitals, and therefore no caliper was used.

Exhibits 243a & 243b: Characteristics of BPCI-participating Hospitals with Coronary Artery Bypass Graft Episodes, Compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals
Model 4, Q4 2013 - Q3 2015

Characteristic		All CABG BPCI Hospitals (N=9)**		Non-participating Hospitals (N=2,774)		All BPCI Hospital (N=23)		Matched CABG BPCI hospitals (N=7)		Comparison Hospitals (N=39+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	5	56%	1,594	57%	16	70%*	5	71%	36	92%*
	Government	0	0%	542	20%*	1	4%*	0	0%	0	0%
	For-Profit	4	44%	638	23%*	6	26%*	2	29%	3	8%*
Urban/Rural	Urban	9	100%	1,902	69%*	21	91%*	7	100%	39	100%
Part of Chain	Yes	4	44%	1,469	53%	10	43%	4	57%	29	74%*

Characteristic	All CABG BPCI Hospitals (N=9)**	Non-participating Hospitals (N=2,774)	All BPCI Hospital (N=23)	Matched CABG BPCI hospitals (N=7)	Comparison Hospitals (N=39+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	529	175*	405*	633	520*
Medicare Days Percent	31%	42%*	31%	31%	38%*
Resident-to-bed ratio	0.11	0.05*	0.14*	0.12	0.13
Disproportionate Share Percent	31%	29%	29%	31%	30%
Number of CABG Discharges, 2011	69	14*	48*	84	78
Standardized Part A Allowed Payment inpatient stay plus 90-day post-discharge period, CABG, 2011	**$38,773**	**$37,174***	**$35,576***	**$38,257**	**$38,162**

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 4 EIs, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 2 hospitals participating in BPCI during the first two years.

*Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All CABG BPCI" hospitals. "Comparison" hospitals are compared to "Matched CABG BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for the fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 24.

b. Medicare Payments and Utilization

There was no statistically significant difference in the change in total Medicare payments ($2015) between BPCI CABG patients and comparison patients from the baseline to intervention periods

(Exhibit 244).[4,5] There were no differences in the change in total payments for the Model 4 episode or total payments for the anchor hospitalization plus all care delivered during the 90-day PDP for BPCI providers relative to the comparison group. None of the individual payment measures, including measures capturing payments during and outside of the bundle, or utilization measures differed between the treatment and comparison groups at a 10% significance level.

Exhibit 244: Impact of BPCI on Total Allowed Payment Amount for Inpatient Hospitalization plus 90-day PDP and Payments under the Bundle, Coronary Artery Bypass Graft Episodes, Model 4, Q4 2011-Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total payment, inpatient stay plus 90-day post-discharge period	658	657	$52,693	$50,081	$54,734	$50,970	$1,151	-$4,251	$6,552	0.68
Total payment under the bundle	675	669	$41,939	$39,419	$40,065	$37,546	-$2	-$1,926	$1,921	1.00

DiD = difference-in-differences, LCI = lower 95% confidence interval; UCI = upper 95% confidence interval.

These estimates are developed using a difference-in-differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. 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 2015 for BPCI and comparison hospitals.

c. *Change in Patient Mix*

Characteristics of the CABG patients in BPCI-participating hospitals and CABG patients at comparison hospitals changed from the baseline to the intervention period (Exhibits 245a & 245b). The percent of CABG patients that were eligible for Medicaid (dual eligible) ($p<0.01$) and the percent that had prior HHA utilization ($p=0.10$) declined more among BPCI episodes than comparison episodes. There was also a relative decrease in the average number of HCC indicators for BPCI patients relative to the change in the comparison group ($p=0.07$). The changes in these characteristics suggest that the BPCI patients became relatively healthier compared to the comparison group from baseline to intervention. It is not clear whether this was due to the initiative.

[4] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

[5] The payment outcomes are based on Medicare's fee-for-service (FFS) payment amounts, not the prospective rate Model 4 participants are paid under BPCI. The analysis uses Medicare FFS payments rather than the discounted, prospective rate for the anchor stay and related Part B services and related readmissions. This allows us to isolate the impact of BPCI on changes in service use and spending.

Exhibits 245a & 245b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Coronary Artery Bypass Graft, Model 4, Q4 2011 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	33	7%	67	9%	65	14%	99	13%	2.4	-2.6	7.4	0.35
	65-79	329	69%	514	69%	325	68%	521	70%	-1.8	-9.3	5.7	0.64
	80+	115	24%	161	22%	87	18%	122	16%	-0.6	-7.1	5.9	0.85
Gender	Female	126	26%	198	27%	145	30%	206	28%	2.9	-4.4	10.2	0.43
Medicaid and Disability	% Eligible for Medicaid	48	10%	69	9%	55	12%	144	19%	-8.6	-14.1	-3.2	0.00
	% Disability, no ESRD	45	9%	80	11%	84	18%	117	16%	3.2	-2.3	8.7	0.26
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	164	34%	236	32%	120	25%	194	26%	-3.6	-11.0	3.8	0.34
	Emergency Room Admission	93	19%	167	23%	113	24%	206	28%	-1.1	-8.0	5.9	0.76
	Home Health	33	7%	44	6%	15	3%	39	5%	-3.1	-6.8	0.6	0.10
	Inpatient Rehabilitation Facility	5	1%	2	0%	1	0%	5	1%	-1.2	-2.4	-0.1	0.04
	Skilled Nursing Facility	9	2%	13	2%	2	0%	14	2%	-1.6	-3.6	0.4	0.12
	Psychiatric Hospital	1	0%	2	0%	1	0%	1	0%	0.1	-0.6	0.9	0.72
	Long-term Care Hospital	0		0		0		0					
	Institutional Nursing Facility[+]	11	2%	15	2%	4	1%	15	2%	-1.5	-3.7	0.7	0.19
	No Institutional Care	312	65%	503	68%	354	74%	544	73%	3.3	-4.1	10.7	0.39
	No Post-acute Care	259	54%	399	54%	280	59%	411	55%	2.8	-5.3	10.9	0.50

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		P-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Lower	Upper	
Average Count of HCCs	2.07	1.77	2.04	1.62	1.87	1.57	2.09	1.87	-0.26	-0.54	0.02	0.07

DiD = Difference-in-Differences.

[+]Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

Notes: "Count" represents the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

d. Quality of Care

Quality of care for BPCI patients was largely maintained relative to comparison patients. The change in all claim-based quality of care measures and the majority of assessment-based quality of care measures did not differ between the BPCI and comparison patients. The remaining outcomes paint a mixed picture with one improvement and one decline in quality of care. Change in beneficiary function was calculated using HHA and SNF patient assessment data, which indicated that BPCI patients discharged to SNF declined in one out of three ADL measures relative to the comparison patients. We also found that BPCI patients discharged to HHA improved in two out of five ADL measures relative to the comparison patients, although one estimate may be biased as we rejected the null hypothesis of the DiD estimate that the BPCI and comparison group had parallel trends during the baseline period (Exhibit 246).

The proportion of BPCI patients discharged to HHA who received two patient assessments and experienced an improvement in upper body dressing increased 7.5 percentage points (p=0.02) and the proportion with an improvement in lower body dressing increased 8.6 percentage points (p=0.01) relative to the comparison group (Exhibit 246).[6] These improvements may be related to activities that three BPCI-participating hospitals discussed with us in telephone interviews in which they specified their expectations for HHA care, including how soon after discharge home visits should occur, the number of physical therapy sessions, and amount of time spent walking.

The proportion of BPCI patients discharged to a SNF who received two patient assessments who experienced an improvement in self-care function decreased by 24.4 percentage points (p=0.06) relative to the comparison group. During the baseline, 61.0% of BPCI beneficiaries discharged to a SNF improved in self-care function, which declined to 44.5% in the intervention period. In contrast, 45.8% of comparison group beneficiaries discharged to a SNF improved in self-care function, which increased to 53.7% in the intervention period.

[6] The estimate on the relative change in proportion of patients with improvement in upper body dressing might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

Exhibit 246: Impact of BPCI on Select Functional Status Outcomes, Coronary Artery Bypass Graft Patients Discharged to HHA or SNF, Model 4, Q4 2011-Q3 2015

Measure+	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
HHA Beneficiaries With Improved Upper Body Dressing	199	160	91.7%	96.5%	93.8%	91.1%	7.5*	1.2	13.8	0.02
HHA Beneficiaries With Improved Lower Body Dressing	199	160	90.4%	95.8%	92.9%	89.8%	8.6	2.0	15.2	0.01
SNF Beneficiaries With Improved Self-care Function	104	65	61.0%	44.5%	45.8%	53.7%	-24.4	-50.1	1.4	0.06

DiD = Difference-in-differences; LCI = lower 95% confidence interval; UCI = upper 95% confidence interval; HHA = Home Health Agency; SNF = Skilled Nursing Facility.

Source: Lewin analysis of patient assessment data for episodes that began Q4 2011 through Q3 2015 for BPCI and comparison providers.

Note: Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

* This might be a biased estimate because we rejected the null hypothesis that BPCI and matched comparison providers had parallel trends for this outcome (with 90% confidence), which is required for an unbiased estimate.

+ Improvement means improvement or staying in the best possible status.

e. Conclusions

Results from the first eight quarters indicate that the BPCI initiative did not have a statistically significant impact on Medicare payments or utilization for CABG episodes. The quality of care was largely maintained. It should be noted, however, that these findings are based on nine hospitals that initiated a CABG episode under the BPCI initiative, with only four remaining in the initiative after eight intervention quarters.

2. Major Joint Replacement of the Lower Extremity (MJRLE)

During the first eight quarters (October 2013 through September 2015) of the Bundled Payments for Care Improvement (BPCI) initiative, there were 3,591 major joint replacement of the lower extremity (MJRLE) episodes of care for Medicare beneficiaries in 15 BPCI-participating acute care hospitals (ACH) under Model 4. Two-thirds of Model 4-participating ACHs terminated their participation in MJRLE episodes after only two intervention quarters. Average Medicare payments for the hospitalization plus the 90 day post-discharge period (PDP) and payments included in the bundle did not differ between the treatment and comparison groups. There were, however, statistically significant changes in utilization of care. BPCI beneficiaries were more likely to receive post-acute care (PAC) services than comparison group beneficiaries. In addition, BPCI beneficiaries experienced relative declines in both 30-day and 90-day unplanned readmission rates. There were no differences in the changes in other quality of care measures. Model 4-participating

hospitals have incentives to reduce readmissions, but not other costs during the 90 day PDP. Their response to these incentives may be to increase PAC use to reduce readmissions.

a. Participants

During the first eight quarters of the BPCI initiative, 15 hospitals (65% of all Model 4 BPCI hospitals) participated in the MJRLE episode. However, by Q3 2015 only five ACHs remained: two-thirds of the ACHs stopped participating in MJRLE after two intervention quarters. One ACH joined in Q2 2015. As a result, the MJRLE results are based on an average of four quarters of experience. During the BPCI initiative, these providers initiated 3,591 MJRLE episodes (approximately 47% of Model 4 episodes across the 48 BPCI clinical episodes).

Compared with hospitals *not* participating in BPCI, hospitals that participated in the MJRLE episode under Model 4 were more likely to be larger, urban, for-profit hospitals, with higher teaching participation (measured by resident-to-bed ratios). They were less likely to be part of a chain and had lower shares of Medicare days. Further, prior to joining BPCI, participating hospitals had more MJRLE discharges and had lower standardized Part A Medicare payments for MJRLE episodes in 2011 compared with non-participating hospitals (Exhibits 247a & 247b).

Matched comparison hospitals were identified for 14 of the 15 BPCI-participating hospitals in the sample.[7] Due to the small number of Model 4 MJRLE participants, we focused on matching utilization and spending patterns of MJRLE patients. Hence, there are more likely to be differences between BPCI Model 4 MJRLE and comparison providers relative to BPCI Model 2 or Model 3 providers and comparison providers along the domains of provider and market characteristics. The 14 matched BPCI-participating hospitals were quite different from the comparison group hospitals. The matched BPCI-participating hospitals were more likely to be larger, for-profit hospitals, with higher resident-to-bed ratios, and to have higher disproportionate share percentages. They also were less likely to be part of a chain and had lower shares of Medicare patients (Exhibits 247a & 247b). We do not know how these differences affected our results.

[7] Each BPCI-participating hospital was matched with up to 15 comparison hospitals based on a propensity score model that considered market and hospital characteristics and baseline outcomes. One participant was not included in the analysis because it had fewer than five relevant discharges in either 2011 or 2012.

Exhibits 247a & 247b: Characteristics of BPCI-participating Hospitals with Major Joint Replacement of the Lower Extremity Episodes, Compared with Non-participating Hospitals, All BPCI-participating Hospitals, and Comparison Hospitals Model 4, Q4 2013 – Q3 2015

Characteristic		All MJRLE BPCI Hospitals (N=15)**		Non-participating Hospitals (N=2,774)		All BPCI Hospitals (N=23)		Matched MJRLE BPCI hospitals (N=14)		Comparison Hospitals (N=208+)	
		N	%	N	%	N	%	N	%	N	%
Ownership	Non-Profit	10	67%	1,594	57%	16	70%	10	71%	180	87%*
	Government	0	0%	542	20%*	1	4%*	0	0%	0	0%
	For-Profit	5	33%	638	23%*	6	26%	4	29%	28	13%*
Urban/Rural	Urban	14	93%	1,902	69%*	21	91%	13	93%	193	93%
Part of Chain	Yes	5	33%	1,469	53%*	10	43%*	5	36%	109	53%*

Characteristic	All MJRLE BPCI Hospitals (N=15)**	Non-participating Hospitals (N=2,774)	All BPCI Hospitals (N=23)	Matched MJRLE BPCI hospitals (N=14)	Comparison Hospitals (N=208+)
	Mean	Mean	Mean	Mean	Mean
Bed Count	322	175*	405*	334	258*
Medicare Days Percent	27%	42%*	31%*	28%	39%*
Resident-to-bed ratio	0.09	0.05*	0.14*	0.10	0.06*
Disproportionate Share Percent	31%	29%	29%	31%	26%*
Number of MJRLE Discharges, 2011	181	105*	205	193	220
Standardized Part A Allowed Payment inpatient stay plus 90 day PDP, MJRLE, 2011	**$22,717**	**$23,836***	**$23,738***	**$22,483**	**$22,477**

PDP=post-discharge period; MJRLE=major joint replacement of the lower extremity.

Source: Lewin analysis of 2013 POS and 2011 Medicare claims. BPCI participating hospitals are defined as Model 4 Eis, Q4 2013 – Q3 2015. Non-participating hospitals are all other hospitals and exclude Model 2 hospitals participating in BPCI.

*Indicates the standardized mean difference between two groups is greater than 0.2 (means are standardized by dividing them by the pooled standard deviation). "Non-participating" and "All BPCI" hospitals are compared to "All MJRLE BPCI" hospitals. "Comparison" hospitals are compared to "Matched MJRLE BPCI" hospitals.

** Please note that BPCI-participating hospitals that received Medicare certification after 2011 are not included in this table.

+ This column represents the weighted number of comparison providers to account for fact that some comparison providers are matched to multiple BPCI hospitals. The unique number of matched comparison providers is 197.

b. Utilization

The share of BPCI MJRLE patients receiving PAC increased 8.9 percentage points relative to the comparison group ($p<0.01$).[8] The share of BPCI patients receiving PAC increased from 67.3% to 72.5%, while the share declined from 80.9% to 77.2% among comparison group patients. Among

[8] These estimates are developed using a Difference-in-Differences (DiD) framework. The DiD framework quantifies the impact of an intervention by comparing changes in outcomes of the intervention group to changes of a control group. The changes are observed before and after the intervention began.

those patients receiving institutional PAC services, BPCI patients experienced a relative decline of 1.6 days (p=0.08). Average institutional days decreased from 22.9 to 22.4 among BPCI patients, while the average days increased from 22.4 to 23.5 among comparison patients (Exhibit 248). The relative reduction in institutional PAC days may be due to the increased share of BPCI patients discharged to PAC because presumably, the additional patients receiving PAC due to BPCI would be healthier and need less PAC than patients who would receive PAC absent BPCI.

Exhibit 248: Impact of BPCI on Proportion of MJRLE Episodes discharged to any Post-Acute Care and Average Number of PAC Days for MJRLE Episodes, Model 4, Q4 2011-Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	p-value
Proportion of MJRLE Episodes discharged to any PAC+	3,341	3,344	67.3%	72.5%	80.9%	77.2%	8.9	3.5	14.4	<0.01
Average Number of PAC Days	1,508	1,521	22.9	22.4	22.4	23.5	-1.6	-3.3	0.2	0.08

DiD = Difference-in-Differences; MJRLE= major joint replacement of the lower extremity. LCI = lower confidence interval at the 5% level; UCI = upper confidence interval at the 5% level. PAC = post-acute care.

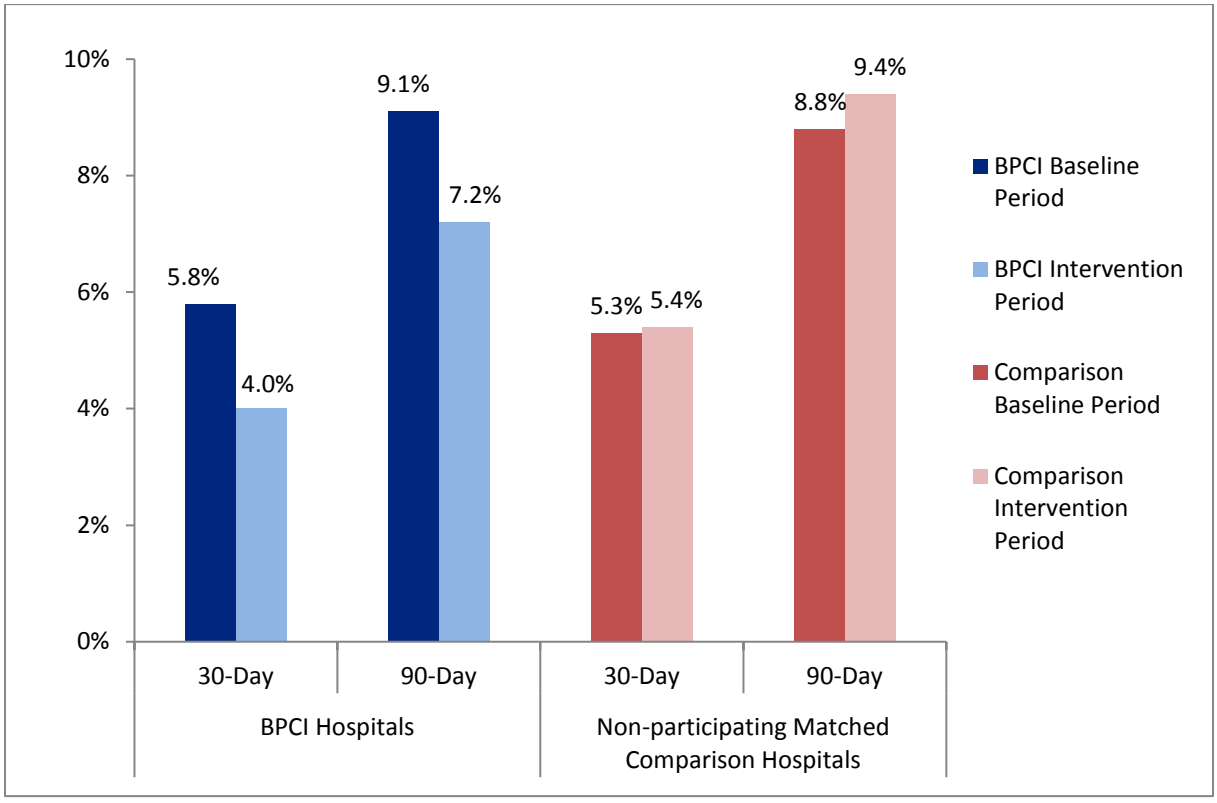
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

+ The proportion of episodes that were discharged to any PAC setting, skilled nursing facility (SNF), inpatient rehabilitation facility (IRF), long term care hospital (LTCH), or home health agency (HHA).

c. *Quality of Care*

There were few relative differences in quality measures between BPCI and comparison episodes. There were not statistically significant differences in the change in emergency department (ED) use or mortality rates or assessment-based outcomes for patients who received care from HHAs or SNFs. BPCI patients experienced relative declines in 30-day and 90-day unplanned readmission rates, which is consistent with the incentives under Model 4 (Exhibit 249). The 30-day readmission rate declined from 5.8% to 4.0% for BPCI patients and it increased from 5.3% to 5.4% among comparison patients for a difference of 1.8 percentage points (p=0.02). The BPCI 90-day readmission rate declined from 9.1% to 7.2%, while it increased from 8.8% to 9.4% in the comparison group for a 2.5 percentage point difference (p=0.03).

Exhibit 249: Impact of BPCI on 30-Day and 90-Day Unplanned Readmission Rate for MJRLE Episodes, Model 4, Q4 2011-Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

It is noteworthy that Model 4 BPCI hospitals significantly reduced their unplanned readmission rates for MJRLE, but Model 2 hospitals did not. The baseline 90-day unplanned readmission rate for Model 4 BPCI MJRLE patients was 9.1% and declined to 7.2% by Q3 2015. For Model 2 BPCI hospitals, the baseline 90-day MJRLE readmission rate was 10.0% and 9.4% by Q3 2015. At the same time, institutional PAC use increased for Model 4 patients and declined for Model 2 patients.

One of the two activities of daily living (ADL) measures for BPCI patients discharged to an IRF improved relative to the comparison group (Exhibit 250). The average self-care score among BPCI IRF beneficiaries improved from 11.6 to 13.7, while the average self-care score among comparison group IRF beneficiaries improved from 11.7 to 12.1, resulting in a relative improvement of 1.7 ($p<0.01$).[9]

[9] Changes in self-care for IRF patients are measured as the change in a self-care related index score between IRF admission and IRF discharge. The self-care index ranges from a score of six (total assistance) to 42 (fully independent).

Exhibit 250: Functional Status of BPCI and Comparison MJRLE Patients with IRF Stays, Model 4, Q4 2011-Q3 2015

Measure	BPCI (N= 378)		Comparison (N= 239)		DiD estimate			
	Baseline	Intervention	Baseline	Intervention	DiD	Lower	Upper	p-value
Average Change in Self-Care Score for IRF Episodes	11.6	13.7	11.7	12.1	1.7	0.5	2.9	<0.01

DiD = Difference-in-Differences; IRF= Inpatient Rehabilitation Facility; Lower = lower confidence interval and Upper = upper confidence interval at the 5% level.

Note: The baseline period is from Q4 2011 through Q3 2012. The intervention period is from Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

Changes in self-care for IRF patients are measured as changes in a self-care related index score between IRF admission and IRF discharge. The self-care index ranges from a score of six (total assistance) to 42 (fully independent).

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

d. Total Medicare Payments

The changes in unplanned readmission rates and PAC use did not result in statistically significant changes in Medicare payments (Exhibit 251). Neither the total Medicare payments ($2015) for the anchor hospitalization plus all care delivered during the 90-day PDP nor the total payments during the bundle changed significantly for BPCI providers relative to the comparison group.[10]

[10] The payment outcomes are based on Medicare's fee-for-service (FFS) payment amounts, not the prospective rate Model 4 participants are paid under BPCI. The analysis uses Medicare FFS payments rather than the discounted, prospective rate for the anchor stay and related Part B services and related readmissions. This allows us to isolate the impact of BPCI on changes in service use and spending.

Exhibit 251: Impact of BPCI on Total Payments ($2015) for Inpatient Hospitalization plus 90-day PDP and Payments under the Bundle, Major Joint Replacement of the Lower Extremity Episodes, Model 4, Q4 2011-Q3 2015

Measure	Number of Intervention Episodes		BPCI		Comparison		DiD estimate			
	BPCI	Comparison	Baseline	Intervention	Baseline	Intervention	DiD	LCI	UCI	P-value
Total allowed payment amount, inpatient stay plus 90-day PDP	3,301	3,334	$27,735	$26,931	$29,582	$28,551	$226	-$1,000	$1,453	0.72
Total Payments Included in the Bundle	3,346	3,351	$16,317	$15,491	$16,244	$15,684	-$266	-$584	$52	0.10

DiD = Difference-in-Differences; LCI = lower confidence interval at 5% significance level; UCI = upper confidence interval at 5% significance level; PDP = Post discharge Period.

Note: Medicare payments are risk adjusted and standardized to remove the effect of geographic differences in wages, extra amounts to account for teaching programs and other policy factors, and inflated to 2015. Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U. Baseline is defined as episodes that began Q4 2011 through Q3 2012. Intervention is defined as episodes that began Q4 2013 through Q3 2015. Sample sizes reflect number of episodes initiated during the intervention period that met inclusion criteria for the outcome.

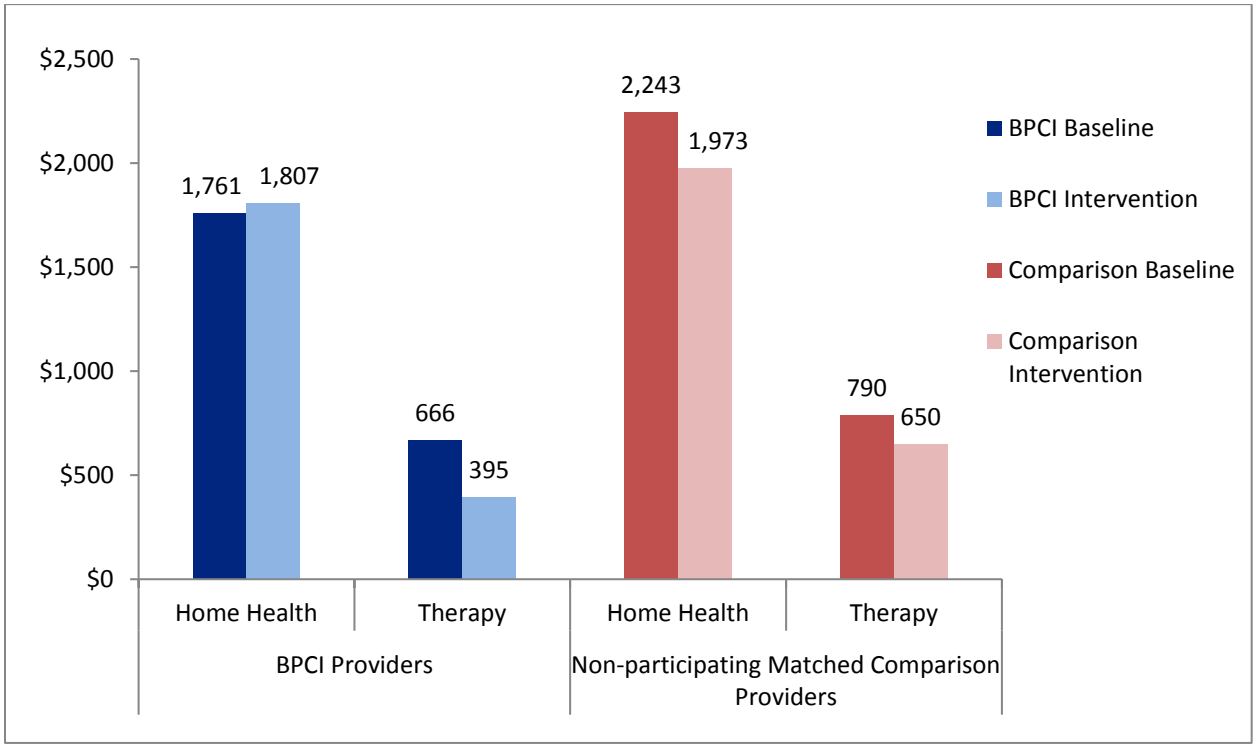
Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison hospitals.

e. Payments Related to PAC Utilization

There was a significant increase in HHA payments and a significant decline in outpatient therapy payments for BPCI relative to comparison patients (Exhibit 252). While there were no statistically significant changes in the average number of HH visits among patients receiving HHA services, 90-day post-discharge period HH payments for BPCI beneficiaries increased by $316 more than payments for comparison group beneficiaries (p=0.02). In the intervention period, HHA payments made up 13% of overall 90-day PDP payments. The HHA payments for BPCI beneficiaries increased from $1,761 to $1,807 while HHA payments decreased for comparison beneficiaries from $2,243 to $1,973.

We also saw a reduction in provision of therapy services (Exhibit 252). Therapy payments made up 3% of overall 90-day PDP payments. Therapy payments during the 90-day PDP declined by $130 (p<0.01). The payments for BPCI beneficiaries declined from $666 to $395 for BPCI episodes while payments decreased from $790 to $650 for comparison group episodes.

Exhibit 252: Impact of BPCI on Home Health and Therapy Payments for Days 1-90 Post-Discharge Period for MJRLE Episodes, Model 4, Q4 2011-Q3 2015



Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

f. Change in Patient Mix

While the impact analysis on payment, utilization, and quality presented above controls for changes in these patient characteristics, it does not directly examine any changes in patient mix. There is evidence that the MJRLE patient population of BPCI-participating hospitals was less healthy after the introduction of BPCI relative to the comparison group. In the intervention period, the BPCI patient population used more health care services prior to the anchor hospitalization, an indication that their health status was worse relative to the comparison group. There were statistically significant increases in the percentage of BPCI participants' MJRLE patients who used an inpatient acute care hospital ($p=0.03$), home health ($p=0.01$), an IRF ($p<0.01$), or a SNF ($p=0.01$) in the six months prior to their index hospitalization, and a statistically significant decrease in the percentage using no institutional care ($p=0.04$) and in the percentage using no PAC ($p=0.09$) in the six months prior to their index hospitalization compared to the matched comparison hospitals (Exhibits 253a & 253b).

Exhibits 253a & 253b: Characteristics of BPCI and Comparison Beneficiaries with a Hospitalization for a Major Joint Replacement of the Lower Extremity, Model 4, Q4 2011 - Q3 2015

Characteristic		BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
		Baseline Period		Intervention Period		Baseline Period		Intervention Period					
		Count	%	Count	%	Count	%	Count	%		Lower	Upper	
Age	20-64	150	6%	228	7%	210	9%	316	9%	0.0	-2.1	2.0	0.97
	65-79	1,448	62%	2,175	64%	1,450	62%	2,218	66%	-1.2	-4.8	2.4	0.52
	80+	745	32%	976	29%	683	29%	845	25%	1.2	-2.1	4.6	0.47
Gender	Female	1,622	69%	2,214	66%	1,571	67%	2,209	65%	-2.0	-5.5	1.5	0.26
Medicaid and Disability	% Eligible for Medicaid	314	13%	502	15%	294	13%	449	13%	0.7	-1.8	3.3	0.58
	% Disability, no ESRD	180	8%	270	8%	260	11%	365	11%	0.6	-1.6	2.8	0.59
Utilization Six Months Prior to Index Hospitalization	Inpatient Acute Care Hospital	278	12%	470	14%	304	13%	415	12%	2.7	0.2	5.2	0.03
	Emergency Room Admission	377	16%	505	15%	400	17%	556	16%	-0.5	-3.3	2.2	0.71
	Home Health	230	10%	404	12%	230	10%	303	9%	3.0	0.7	5.2	0.01
	Inpatient Rehabilitation Facility	25	1%	52	2%	37	2%	25	1%	1.3	0.5	2.1	<0.01
	Skilled Nursing Facility	78	3%	177	5%	109	5%	156	5%	1.9	0.4	3.5	0.01
	Psychiatric Hospital	7	0%	5	0%	8	0%	12	0%	-0.2	-0.6	0.2	0.41
	Long-term Care Hospital	3	0%	3	0%	3	0%	4	0%	0.0	-0.3	0.2	0.82
	Institutional Nursing Facility+	113	5%	202	6%	129	6%	182	5%	1.3	-0.4	3.0	0.14
	No Institutional Care	2,036	87%	2872	85%	2,010	86%	2,925	87%	-2.7	-5.3	-0.1	0.04
	No Post-acute Care	1,691	72%	2,401	71%	1,653	71%	2446	72%	-3.0	-6.3	0.4	0.09

Characteristic	BPCI Hospitals				Comparison Hospitals				DiD	95% Confidence Interval		p-value
	Baseline Period		Intervention Period		Baseline Period		Intervention Period					
	Mean	SD	Mean	SD	Mean	SD	Mean	SD		Lower	Upper	
Average Count of HCCs	1.16	1.38	1.16	1.51	1.16	1.41	1.08	1.41	0.08	-0.03	0.18	0.16

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

DiD = Difference-in-Differences; "Count" represent the numerator for the given characteristic. The % is the numerator divided by the total number of episodes with non-missing values.

+ Institutional Nursing Facility is defined as any days in a nursing facility regardless of payer (Medicare, Medicaid, beneficiary) based on the Minimum Data Set (MDS) assessment data. All other measures are based on Medicare claims.

g. Conclusions

Results from the first eight quarters indicate that the Model 4 MJRLE BPCI initiative increased the utilization of PAC, and reduced unplanned readmission rates. Increased PAC use led to relative increases in HH payments and relative declines in therapy payments. Overall payments and payments covered during the bundle did not significantly change.

It should be noted that these findings are based on 14 matched hospitals that initiated a MJRLE episode under the BPCI initiative, with only five remaining in the initiative after eight intervention quarters.

VI. Cross Model Episode Comparison

A. Congestive Heart Failure: An Overview of Findings under Models 2 and 3

There was sufficient episode volume to analyze congestive heart failure (CHF) episodes across three strata: Model 2 acute care hospitals (ACHs or hospitals), Model 3 skilled nursing facilities (SNFs), and Model 3 home health agencies (HHAs). During the first eight quarters of the initiative, there was evidence of two broad conclusions about the impact of BPCI: (1) quality of care was maintained across strata, and (2) only Model 3 HHAs significantly decreased total Medicare payments for BPCI patients relative to the comparison group. There were also some indications that Model 2 hospitals and Model 3 HHA BPCI participants may have treated a healthier mix of patients post-BPCI relative to the change for the comparison group. However, these changes do not explain the impact estimates on payment, utilization, and quality outcomes as those analyses controlled for claim-based patient characteristics. In addition, although some providers were able to achieve positive net payment reconciliation amounts (NPRA) under BPCI, interviewees in all three strata discussed similar challenges with the CHF episodes, such as difficulty identifying CHF patients, high incidence of comorbidities, and financial uncertainty about the episode due to difficulties controlling costs after discharge.

Total Medicare payments ($2015)[1] during the qualifying hospital stay and the 90-day post-discharge period (PDP) were $970 (4% of BPCI baseline dollars) lower for episodes initiated at BPCI-participating HHAs relative to matched comparison HHAs from the baseline to the intervention period (p=0.05). This was due to declines in SNF and HHA payments for Model 3 BPCI HHA episodes relative to the comparison group. In addition, HHA visits decreased for Model 3 BPCI HHA patients relative to the comparison group. Although there was no statistically significant change in total payments for episodes initiated by Model 2 hospitals, there was some evidence that hospitals substituted costlier inpatient rehabilitation facility (IRF) care with home health services relative to comparison hospitals. There were no statistically significant changes in payments or PAC utilization among Model 3 BPCI SNFs relative to comparison SNFs.

While the impact analysis on payment, utilization, and quality controls for changes in claim-based patient characteristics, it does not directly examine changes in patient mix. There was some evidence that Model 2 hospital and Model 3 HHA BPCI participants treated healthier patients during the intervention than the baseline period relative to the comparison group. Among Model 2 hospitals, there were statistically significant declines from the baseline to intervention period in the proportion of BPCI patients who used an emergency room, IRF, or LTCH during the six months prior to the index hospitalization relative to the comparison group. For Model 3 HHAs, the patient assessment conducted at the beginning of the episode showed improvements in five out of 19 measures, including a 17.4 percentage point decline in the share of BPCI HHA patients with poor overall status relative to the comparison group, indicating a healthier patient mix. The BPCI-participating HHAs also saw a larger decline in the HHA case-mix index than the comparison group. We conducted sensitivity analyses to determine if the changes observed in the patient assessment data may explain the significant BPCI impacts on payment for Model 3 HHA CHF episodes and found that the impact estimate remained

[1] Medicare payments are expressed in 2015 dollars. The 2015 dollars are the result of adjusting actual dollars based on changes in the medical component of the CPI-U.

statistically significant. We observed no consistent pattern of changes in patient characteristics among CHF patients treated by Model 3 BPCI SNFs.

Interviewees in all three strata discussed similar challenges to participating in BPCI. They indicated that it is difficult to identify CHF patients, patients are complex to manage due to multiple comorbidities, and the CHF episode is financially risky due to the difficulties controlling costs after discharge. However, some providers were able to achieve positive NPRA under BPCI. Analysis of top and bottom financial performers in Model 2, as measured by standardized NPRA[2], points to certain factors that may have contributed to the variation in NPRA among Model 2 CHF BPCI providers. Top performers may have achieved positive NPRA by decreasing unplanned readmissions and decreasing the share of patients discharged to institutional PAC, while increasing the proportion of patients discharged to HHA. Decreasing the share of patients discharged to institutional PAC may have been facilitated by a higher use of institutional PAC among patients treated during the baseline period. In addition, a selection of top and bottom performers that were interviewed indicated that the ability to identify CHF patients before discharge and working with preferred PAC providers were important factors in their success in the CHF episode under BPCI.

B. Major Joint Replacement of the Lower Extremity: An Overview of Findings under Models 2, 3, and 4

There was sufficient episode volume to analyze major joint replacement of the lower extremity (MJRLE) episodes initiated by Model 2 hospitals, Model 3 skilled nursing facilities (SNFs), Model 3 home health agencies (HHAs), and Model 4 hospitals. In addition, we had sufficient sample size to stratify episodes initiated by Model 2 hospitals and Model 3 SNFs by fracture status. Early results from the evaluation suggest that linking the payments for all providers involved in delivering an MJRLE episode of care holds promise for reducing total payments for the inpatient stay and the 90-day post-discharge period (PDP). We observed statistically significant decreases in Medicare payments for the inpatient hospital stay and all care delivered during the 90-day PDP for MJRLE episodes initiated by Model 2 hospitals and Model 3 SNFs, which were driven by reductions in the use of institutional post-acute care (PAC). Quality of care, as measured by mortality, readmissions, and emergency department use rates, was maintained or improved for BPCI beneficiaries across Models, but we observed some declines in quality as measured by functional status outcomes from patient assessments. There is some evidence that BPCI participants in Models 2 and 3 may have treated a healthier mix of MJRLE patients during the intervention relative to the comparison group. However, these changes do not explain the impact estimates on payment, utilization, and quality outcomes as those analyses controlled for claim-based patient characteristics.

1. Reductions in Institutional PAC Use Drove Decreases in Total Payments

Model 2 hospitals and Model 3 SNFs significantly reduced total payments for the inpatient hospital stay and all care delivered during the 90-day PDP ($2015) for both non-fracture and

[2] To adjust for differences in baseline payments and thus target prices among BPCI providers, we standardize NPRA by dividing it by the provider's average target price.

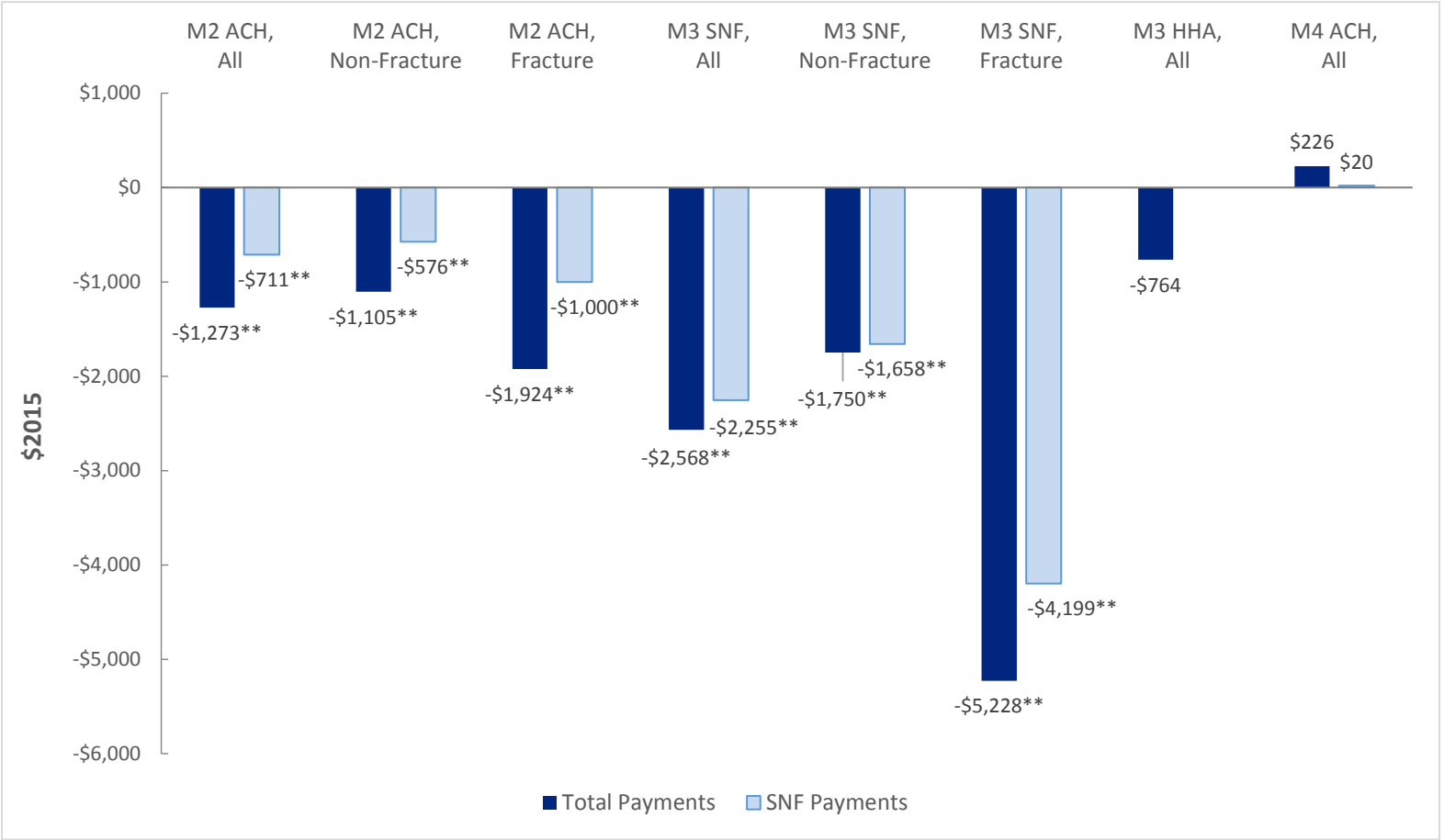
fracture MJRLE episodes (see Exhibits 254 and 255).[3] Model 2 hospitals and Model 3 SNFs successfully reduced total payments by at least $1,105 per episode between the baseline and intervention periods relative to comparison episodes ($p<0.01$). As shown in Exhibits 254 and 255, across Models and episode types, the largest reduction in total payments occurred for fracture episodes initiated by Model 3 SNFs (a relative decrease of $5,228 or 10.6%, $p<0.01$). There was no statistically significant change in total payments among MJRLE episodes initiated by Model 3 HHAs or Model 4 hospitals.

The decrease in total payments among Model 2 hospitals and Model 3 SNFs was driven by a reduction in institutional PAC payments. A decrease in SNF payments explains more than 50% of the decline in total payments among Model 2 hospitals and more than 80% of the decline in total payments among Model 3 SNFs (see Exhibit 254). There were also significant decreases in inpatient rehabilitation facility (IRF) payments among Model 2 hospitals. The observed reductions in institutional PAC payments are consistent with information we received through quarterly interviews and site visits with BPCI participants.

Model 2 hospitals and Model 3 SNFs had more of an opportunity and incentive to reduce the number of SNF days than Model 3 HHAs or Model 4 hospitals. The share of patients that have any SNF days is substantially lower for episodes initiated by HHAs than hospitals or SNFs. Further, prior to BPCI, patients in MJRLE episodes initiated by Model 2 hospitals and Model 3 SNFs spent more days in a SNF during the 90-day PDP than patients in episodes initiated by HHAs (approximately 24.0 days for hospital- and SNF-initiated episodes vs. 15.5 days for HHA-initiated episodes). There was a decline of 1.5 SNF days among episodes initiated by Model 2 hospitals and 4.0 SNF days among episodes initiated by Model 3 SNFs relative to comparison episodes ($p<0.01$). There was no statistically significant change in the number of days spent in a SNF among episodes initiated by Model 3 HHAs. Although beneficiaries in MJRLE episodes initiated by Model 4 hospitals had a similar number of SNF days prior to BPCI (23.9 days) as beneficiaries in episodes initiated by Model 2 hospitals and Model 3 SNFs, there was no significant change in the number of SNF days among episodes initiated by Model 4 hospitals. Model 4 hospitals likely faced less incentive to reduce the number of SNF days because post-acute services are not included in the Model 4 bundle.

[3] To facilitate the comparison across models, this section compares total payments for the inpatient hospital stay and a 90-day post-discharge period across Models 2, 3, and 4. Under Model 2, an episode includes the anchor hospitalization plus 30, 60, or 90 days of care post-discharge. Under Model 3, an episode includes post-discharge services for 30, 60, or 90 days. Under Model 4, an episode includes the anchor hospitalization and readmissions within 30 days of discharge.

Exhibit 254: Impact of BPCI on Total Payments during the Inpatient Hospitalization and the 90-day PDP and SNF Payments during the 90-day PDP ($2015) for MJRLE Episodes, by Model and EI Type, Q4 2013 – Q3 2015[+]



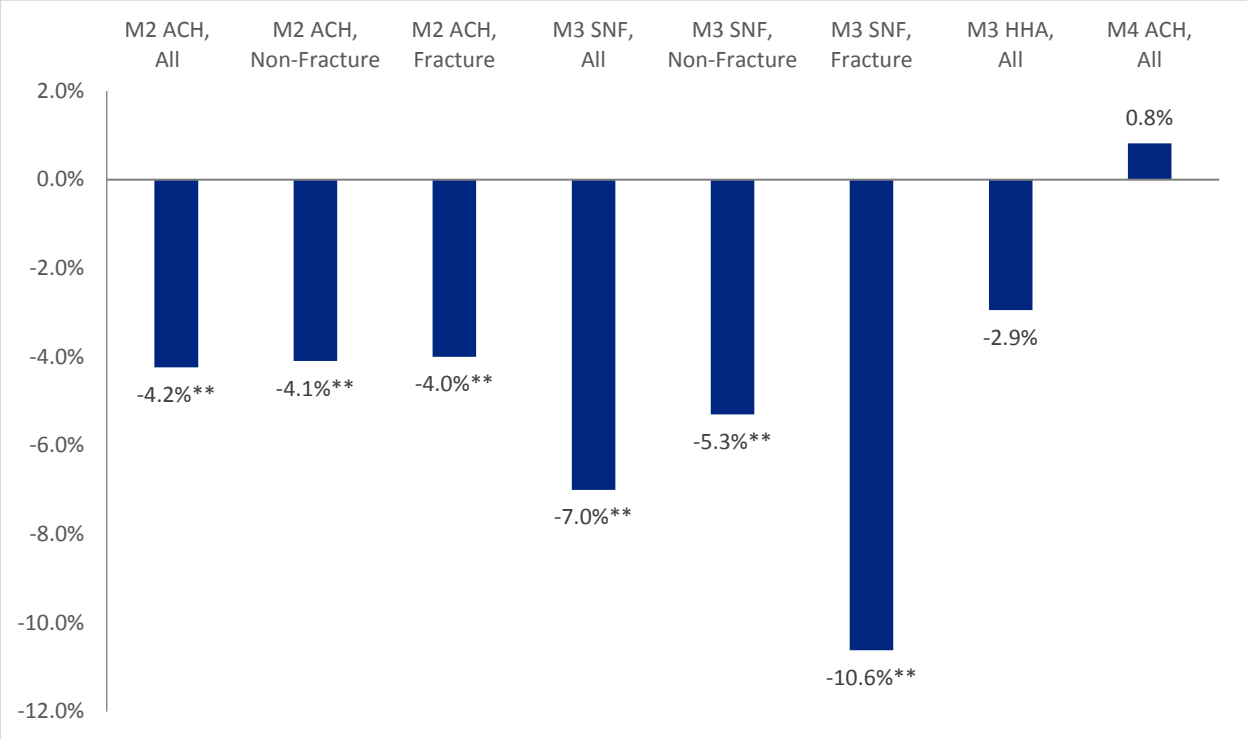
ACH = acute-care hospital; HHA = home health agency; SNF = skilled nursing facility; MJRLE = Major Joint Replacement of the Lower Extremity; M2 = Model 2; M3 = Model 3; M4 = Model 4; PDP = post-discharge period

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

[+]There is not sufficient volume to calculate SNF payments for HHA initiated episodes.

Note: ** Indicates $p < 0.05$

Exhibit 255: Percent Change in Total Payments during the Inpatient Hospitalization and the 90-day PDP for BPCI MJRLE Episodes from Baseline to Intervention Relative to Comparison Episodes, by Model and EI Type, Q4 2013 – Q3 2015



ACH = acute-care hospital; HHA = home health agency; SNF = skilled nursing facility; MJRLE = Major Joint Replacement of the Lower Extremity; M2 = Model 2; M3 = Model 3; M4 = Model 4; PDP = post-discharge period

Source: Lewin analysis of Medicare claims and enrollment data for episodes that began Q4 2011 through Q3 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI EIs and the matched comparison providers.

Note: ** Indicates $p < 0.05$

2. The Impact of BPCI on Quality

Quality of care as measured by claims data (i.e., mortality, readmissions, and emergency department use rates) was maintained or improved for BPCI patients across Models. From the baseline to the intervention period, there were no statistically significant changes in claim-based quality measures for Model 2 episodes and there were improvements in quality across episodes in Models 3 and 4 (see Exhibit 256). Among beneficiaries in non-fracture episodes initiated by SNFs, there was a statistically significant reduction in the 90-day mortality and readmission rate. There were also statistically significant improvements in the 30- and 90-day readmission rates among episodes initiated by hospitals in Model 4. This result aligns with the design of Model 4, which provides what may be the strongest incentive across Models to reduce readmissions. Under Model 4, PAC use is not included in the prospective payment amount but readmissions are included, so hospitals may be incentivized to increase the use of PAC to reduce hospital readmissions. Indeed, the proportion of Model 4 MJRLE patients discharged to any PAC increased by 8.9 percentage points relative to patients treated by comparison hospitals ($p<0.01$).

Changes in quality as measured by the change between two functional assessments were less favorable, especially among non-fracture patients who received SNF care in Models 2 and 3.[4] Among this group, which is approximately three-quarters of the episodes, the share of patients treated by Model 2 hospitals who experienced an improvement in each of three activities of daily living (ADL) measures decreased between the baseline and intervention periods relative to comparison episodes.[5] The proportion of non-fracture Model 3 SNF patients who experienced an improvement also declined for one of the three ADL measures. Consistent with these results, there was a relative statistically significant decline in one of two quality measures (the self-care score) among Model 2 non-fracture patients who were discharged to an IRF, though this was due to a smaller improvement among BPCI patients (Exhibit 256). There were otherwise few changes in functional status among the approximately three-quarters of patients with two assessments discharged to other PAC facilities, and the changes were not consistent across strata.

One possible explanation for the relative declines in quality among beneficiaries discharged to SNF or IRF is that patients who were discharged to an institutional PAC setting spent fewer days there. As the ADL measures capture changes from the beginning to the end of the first PAC stay, a reduction in the length of time spent in an institutional PAC setting could help explain the relative declines observed in ADL measures among beneficiaries discharged to a SNF or IRF. Under BPCI, providers are incentivized to move patients to appropriate and less costly settings more quickly. The earlier discharge from institutional PAC facilities is also consistent with research quoted by orthopedic surgeons during case studies, which indicated that MJRLE patients recover better and faster at home. It is not possible to know if patients with shorter stays achieved the same level of function after leaving the institutional PAC setting as they would have had they remained longer. It should also be noted that these patients may not be representative of all patients discharged to PAC due to the reasons they do not have two patient assessments. Patients were not included in the assessment-based quality outcomes if the patient was readmitted to the hospital, died, had little or no cognitive function, or if the episode occurred later during our measurement period and therefore a second assessment was not yet available in the data.

[4] Functional status measures related to the SNF setting are calculated for patients discharged to a SNF in Models 2 and 4 and patients who began their episode in a SNF in Model 3. The patients had to remain in the SNF long enough to have two assessments in order to be included in this measure.

[5] "Improvement" means moving to a higher functional status or maintaining the best possible status.

Exhibit 256: Impact of BPCI on Claim-based and Assessment-based Quality Outcomes, MJRLE, Models 2, 3, and 4, Q4 2013 – Q3 2015

Measure	Model 2			Model 3				Model 4
	ACH			SNF			HHA	ACH
	All	Non-Fracture	Fracture	All	Non-Fracture	Fracture	All	All
All-cause mortality rate, 30-day PDP	0.01	-0.07	0.52	-0.35	-0.21	-0.84		-0.16
All-cause mortality rate, 90-day PDP	0.08	0.04	0.37	-1.25	-1.01	-1.82		-0.09
Emergency Department Use, 30-day PDP	0.22	0.15	0.78	-0.31	-0.29	-0.42	-1.72	-0.74
Emergency Department Use, 90-day PDP	0.38	0.22	1.52	-0.28	-0.99	2.08	-3.22	-0.29
Unplanned readmission rate, 30 day PDP	-0.11	0.03	-0.71	-1.44	-1.01	-2.77	-0.61	-1.85
Unplanned readmission rate, 90-day PDP	-0.34	-0.16	-1.03	-2.67	-2.44	-3.58	-2.13	-2.47
ADL, HHA, improved bathing	-0.19	-0.11	-2.57				0.72	-1.03
ADL, HHA, improved ambulation	-0.96	-0.79	-6.14				0.67	-2.06
ADL, HHA, improved upper body dressing	-0.01	-0.09	0.87				1.17	-2.40
ADL, HHA, improved lower body dressing	-0.72	-0.62	-4.88				-1.95	0.30
ADL, HHA, improved bed transferring	-1.19	-1.03	-9.93				7.91	4.25
ADL, SNF, improved overall function	-2.24	-2.87	-0.37	-5.04	-7.25	1.60		0.80
ADL, SNF, improved self-care function	-3.94	-5.19	-0.17	-9.18	-11.99	0.19		-4.79
ADL, SNF, improved mobility function	-2.86	-3.62	-0.67	-4.61	-6.27	0.97		-3.15
ADL, IRF, average change in mobility score	-0.13	-0.34	0.30					0.64
ADL, IRF, average change in self-care score	-0.31	-0.67	0.33					1.70

ACH = acute-care hospital; HHA = home health agency; SNF = skilled nursing facility; MJRLE = Major Joint Replacement of the Lower Extremity; PDP = post-discharge period; ADL=activities of daily living

Note: Statistical significance at the 0.10 level is indicated by orange and green shaded cells. Dark Orange indicates the DiD estimate was negative and statistically significant at 5% level; light orange indicates the DiD estimate was negative and statistically significant at 10% level; dark green indicates the DiD estimate was positive and statistically significant at 5% level; and light green indicates the DiD estimate was positive and statistically significant at 10% level. 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 2012 (baseline) and Q4 2013 through Q3 2015 (intervention period) for BPCI hospitals and the matched comparison providers.

3. Changes in Functional Status and Patient Experience from Beneficiary Surveys

Results from the beneficiary survey, which are available for patients in MJRLE episodes initiated by Model 2 hospitals and Model 3 SNFs, indicate that BPCI survey respondents consistently reported equal or greater changes in functional status (e.g., improvements in walking without rest and using stairs) from before to after their surgical episode relative to comparison respondents.[6] This finding was consistent across Models and survey waves.

Survey results for patient experience were less consistent. In both Models 2 and 3, BPCI and comparison respondents reported similar levels of overall satisfaction with recovery. Model 2 BPCI respondents generally reported the same or better experiences relative to comparison respondents, though there were two measures (out of ten) in which BPCI respondents reported worse experiences in at least one survey wave.[7] In Model 3, BPCI respondents reported worse care experience relative to comparison respondents for three measures. However, this result was limited to the single wave for which we have Model 3 MJRLE results.[8]

4. Change in Patient Mix

Episode initiators have incentives to select a healthier mix of patients, or avoid potentially high cost ones, to reduce their episode payments below their target. While the impact analysis on payment, utilization, and quality controls for changes in the claim-based patient characteristics, it does not directly examine any changes in patient mix. It is more likely that we would observe changes in patient characteristics indicative of a healthier population among the non-fracture patient population than the fracture patient population at Model 2 hospitals because non-fracture MJRLE procedures are usually planned whereas fracture procedures typically are not. There is evidence that suggests the non-fracture patient populations became healthier at Model 2 hospitals, which may be an indication that participants are selecting healthier patients after joining BPCI. Model 2 hospitals experienced declines in the proportion of non-fracture patients

[6] Model 2 survey results are available in the "Functional Status and Care Experiences" section of the Model 2 MJRLE issue brief. Survey results for Model 3 SNF episodes are available in the "Quality of Care" section of the Model 3 SNF MJRLE issue brief.

[7] Subsequent results from the Wave 6 (Spring 2016) and Wave 7 (Summer 2017) surveys indicated a reversal of the positive changes in functional status among Model 2 BPCI respondents with MJRLE episodes. In Wave 6, BPCI respondents did not report any significantly higher rates of improvement or lower rates of decline on any measure relative to comparison respondents. Rather they indicated significantly lower rates of improvement or higher rates of decline for 4 of 7 functional status measures, including planning regular tasks, use of a mobility device, using stairs, and the degree to which pain limited their regular activities. In Wave 7, BPCI respondents did report significantly higher rates of improvement in pain limiting everyday activities, but still reported lower rates of improvement or higher rates of decline for 3 functional status measures, including planning regular tasks, walking without rest, and using stairs.

[8] Subsequent surveys covering beneficiaries participating in the BPCI Model in the spring (Wave 6) and summer (Wave 7) of 2016 were analyzed together at the conclusion of Wave 7. Results indicated worse care experiences among BPCI respondents relative to comparison respondents for seven of ten measures, although only two differences were statistically significant. However, none of the three negative outcomes that were significant in Wave 5 remained significant in Wave 7. Moreover, unlike Wave 5, where BPCI respondents were 3.0 percentage points less likely to be satisfied with their overall recovery, BPCI respondents in Wave 7 were 3.0 percentage points more likely to be satisfied with their overall recovery. Thus, while there seems to be some persisting negative perception of care experience across waves, it is not specific to any particular outcome, nor does it have a continued impact on overall satisfaction.

with health care use in the six months prior to the anchor hospitalization. Among fracture episodes initiated by Model 2 hospitals, we observed potential increases in the severity of the patient population, as evidenced by increases in the proportion of patients with health care use in the six months prior to the anchor hospitalization.

SNFs have considerable discretion to determine which patients they admit to their facilities among those with fracture and non-fracture joint replacements. There is some evidence of a shift to a healthier patient mix among Model 3 SNF non-fracture and fracture episodes, which may be an indication that participants selected a healthier patient mix after joining BPCI. Model 3 SNFs experienced declines in the proportion of non-fracture patients with health care use in the six months prior to the index hospitalization and admitted a greater proportion of younger non-fracture patients between the baseline and intervention period relative to comparison group. Additionally, data from the initial SNF patient assessment indicate BPCI SNFs experienced declines in the proportion of non-fracture patients with several characteristics indicative of severity (e.g., proportion of patients who need extensive assistance or are totally dependent walking in room) relative to the change for the comparison group. While there were no statistically significant changes in claim-based patient characteristics among fracture patients, data from the initial SNF patient assessment indicate BPCI SNFs experienced declines in the proportion of fracture patients with several characteristics indicative of severity relative to comparison episodes. As presented earlier, there was a significant decline in payment, utilization, and quality outcomes for SNF non-fracture MJRLE. We tested if the statistical significance and direction of the impact estimate changed after adding patient assessment characteristics to the risk adjustment model. We found that even when we include patient characteristics from the initial SNF patient assessment, the impact estimates remained statistically significant.

We also analyzed changes in patient mix for episodes initiated by Model 3 HHAs and Model 4 hospitals. We observed no consistent pattern of changes in patient characteristics among patients admitted to Model 3 HHAs. Among patients in HHA-initiated episodes, claim-based characteristics suggest there was an increase in the share of resource-intensive patients. Several assessment-based characteristics also indicated patients in HHA-initiated episodes were less healthy, while other characteristics indicated there was a relative increase in the share of healthier patients. We also observed potential increases in the severity of MJRLE episodes initiated by Model 4 hospitals, as evidenced by increases in the proportion of patients with health care use in the six months prior to the index hospitalization.

5. Market Share of MJRLE Episodes

We explored whether BPCI participation impacted the market share of MJRLE episodes for Model 2 hospitals and Model 3 SNFs. The changes that we observed do not suggest that BPCI participants increased their market share. Among Model 2 hospitals, the magnitude of change in market share was small and the direction of change was not consistent over time. Among Model 3 SNFs, the market share of MJRLE episodes varied by cohort; the market share of SNFs that joined earlier in the initiative tended to decrease during the intervention period, while the market share of SNFs that joined later tended to stay the same.

We did detect changes in the number and quality of PAC providers to which Model 2 hospitals discharged their patients. Model 2 hospitals tended to discharge the majority (75%) of MJRLE patients to a slightly larger number of SNFs (i.e., lower concentration) and slightly fewer HHAs

(i.e., higher concentration) during the intervention than the baseline. Additionally, among Model 2 patients discharged to a SNF or HHA, slightly greater shares of patients were discharged to highly-rated SNFs and HHAs relative to the baseline period.

VII. Discussion and Conclusion

A. Discussion

In this third annual report on the evaluation of Models 2, 3, and 4 of the BPCI initiative, we are beginning to see the impact of changes providers are making in response to the BPCI incentives based on the first two years of the initiative. The effects are far from clear or straightforward, however. The lack of consistent or significant results may be partly due to the short average tenure of participants in the initiative. The data in this report are based on episode initiators (EIs) with an average of three quarters of BPCI experience, which may not be enough time to see results on payments and quality from care redesign. Even so, patterns across Models and episodes begin to tell a tale of less intensive use of PAC.

The providers that chose to participate in BPCI are a self-selected group that are generally larger, urban, and in more affluent areas. Providers also were allowed to choose among alternative ways to participate in BPCI. Their choices of specific Model, episode length, and clinical episodes were based on their assessment of where their organization could make the most advantageous changes and have the most impact. In particular, there is a consistent pattern of participants entering into clinical episodes in which they had higher than average baseline costs, which may indicate that they had the most inefficient patterns of care that would be easier to change. The implication of these choices is that the impacts of BPCI are likely to be in part limited to higher cost, less efficient episodes.

It is not surprising that the initiative has seen the most notable impacts on episode payments for MJRLE under Models 2 and 3. MJRLE is a relatively high volume procedure with substantial PAC use. It is usually an elective surgery, so providers can plan the episode of care and prepare the patient. In addition, MJRLE patients can generally be identified at the beginning of or prior to the hospital admission. This is important to ensuring that care protocols can be implemented in a timely manner, which has been a challenge for other clinical episodes because providers indicated that they often did not know for days after admission whether a patient was in a BPCI episode or not. Interestingly, reductions in episode payments were even more pronounced for MJRLE due to fracture, which is a higher cost episode with greater PAC use, but is generally not an elective surgery.

The BPCI initiative has had its greatest impact on PAC payments and use. Medicare spending for PAC is quite variable, even across beneficiaries in the same clinical episode. The variability may be due to inefficiency in PAC use and the lack of clinical consensus over what constitutes appropriate care for rehabilitation or recovery following a hospitalization. Therefore, reducing PAC spending may be the most viable approach to reducing total episode payments. We have seen reductions in hospital discharges to SNF and particularly IRF, both of which typically have higher Medicare payments than HHA. We have also seen declines in SNF length of stay. Under Model 2, reducing PAC spending is particularly important in achieving positive net payment reconciliation amount (NPRA) because the hospital payment, which is often the largest component of the episode payment, is a per discharge amount. Reducing resources used during the hospital stay can contribute to internal cost savings for the hospital, but is unlikely to affect Medicare's payment (unless the hospital length of stay falls below a limit that triggers a per diem payment). The same is true for HHA payments. Because Medicare pays HHAs for a 60-day

episode of care, reducing HHA visits may generate internal cost savings for the agency, but is unlikely to reduce Medicare payments (unless the number of visits falls below a low volume episode limit). SNFs, by contrast, are paid a daily amount, so the decline in SNF length of stay that we have observed in many clinical episodes translates into lower Medicare payments.

While our analysis shows some significant changes in quality of care outcomes for a few Model 2 clinical episodes, the lack of consistency for any clinical episode across the claims, assessment, and patient survey measures reduces concerns about systematic problems. In addition, there have been few statistically significant changes in service use among clinical episodes with declines in quality measures, so it is unclear how these outcomes could be due to BPCI. That said, MJRLE clinical episodes under Model 2, which showed the most significant changes in service use, also showed significant relative improvements in patient-reported outcomes.[1]

For several SNF Model 3 clinical episodes, readmissions, mortality, and emergency department use increased relative to the comparison group. It is unclear what is causing these statistically significant changes in quality outcomes, but it is important to keep in mind that the results are based on a small number of EIs and patient episodes and they were not consistently attributed to the same clinical episodes. There were also large differences in the baseline period rates for many of the clinical episodes, despite our attempts to find a good comparison group, which contributed to these findings. We will continue to monitor these outcomes, track whether they are consistent over time and across clinical episodes, and evaluate them in the next annual report which will include an additional four quarters of BPCI participation experience. At the same time, however, among the subgroup of patients who remained in the SNF long enough to receive two patient assessments, the functional assessments administered in the SNFs indicated improvements for patients in BPCI episodes, relative to the change for comparison patients.

Differences between hospital EIs that achieved the highest and the lowest NPRA for their MJRLE episodes provide further insights into how participants respond to BPCI. Hospitals that received the highest relative NPRA had higher baseline institutional PAC use in their MJRLE episodes, which may have facilitated greater PAC payment reductions under the initiative. This is consistent with evidence from site visits and interviews in which representatives of hospitals with the highest NPRA discussed their strategies for reducing SNF use by substituting HHA care. Further, from the baseline to the intervention period, MJRLE patients in hospitals with the highest NPRA became less severe across several measures, and the share of fracture patients in these hospitals declined compared with hospital EIs with the lowest relative NPRA. This may have been because their volume of MJRLE discharges increased, primarily because of an increase in non-fracture patients who are typically less severe than fracture patients.

B. Limitations

There are several limitations with this evaluation. First, even though this is the third Annual Report, most participants still have not had much experience under the BPCI initiative. The outcomes we report reflect up to the first two years of the initiative; however, the average length of participation among EIs is three quarters, primarily because most participants did not join the

[1] Subsequent results indicated a reversal of the positive changes in functional status among Model 2 BPCI respondents with MJRLE episodes.

initiative until April or July 2015. The care process changes that are expected to affect episode costs and quality are often difficult to implement and may require additional time before achieving results that we can measure.

The primary analytic approach for this evaluation relies on the differential change in claim-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 is therefore 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 that are expected to affect provider responses to BPCI and are available primarily in Medicare administrative data. Some drivers of success, however, cannot be captured through administrative data. This limits our ability to match on all factors that may have influenced participation in BPCI and performance under the initiative. It is important to keep in mind that providers are responding to multiple, often conflicting incentives in the changing health care environment. This context makes it difficult to isolate the effects of BPCI.

Because the difference-in-differences (DiD) estimate attributes differences in trends between BPCI and the comparison group during the intervention period to BPCI, it is essential that the comparison and BPCI providers have parallel trends during the baseline period. With this in mind, we matched BPCI participants with other providers on key outcomes such as emergency department visits, readmissions, mortality, and total payments in both 2011 unadjusted levels and changes from 2011 to 2012. Despite this approach, we rejected the null hypothesis that there were parallel trends in the baseline period among the final matched episodes for 34 of 287 (12%) DiD estimates. Thus, for these estimates, the underlying assumptions of the DiD method were violated. Many of these cases were associated with relatively low incidence events for which year to year fluctuations in provider-level unadjusted outcomes may be poor proxies for quarterly trends in the baseline period. Furthermore, the baseline levels differed for many of the Model 3 SNF quality outcomes in particular.

The estimates of the BPCI impact on payment, utilization, and quality of care account for differences in provider and market characteristics, as well as patient mix that is measurable with claims data. However, the change in patient mix for Model 3 BPCI SNF and HHA participants relative to the comparison group suggest that claim-based characteristics may not be sufficient to measure changes in patient mix. For example, the claim-based characteristics among Model 3 SNF non-fracture MJRLE episodes did not suggest there was as change in patient mix relative to the comparison group, while there were indications from assessment-based measures (e.g. functional status measures) that BPCI SNFs served healthier patients during the intervention period relative to the change for the comparison group. To assess if the current risk adjustment models were not appropriately capturing changes in patient mix, we added assessment-based measures to the risk adjustment model. We found that for most outcomes, the statistical significance and direction of the Model 3 impact estimate did not change. This suggests that the claim-based patient characteristics may also capture changes in patient mix observed in assessment-based measures. As with all risk adjustment models, however, we were unable to control for unobservable characteristics.

The evaluation estimates the average impact of BPCI on payment, utilization, and quality of care for BPCI providers, relative to a comparison group. Thus, our evaluation does not measure the

variation in impact of BPCI across providers. However, we did examine the variation in NPRA among Model 2 participating hospitals and the factors associated with the variation, including change in patient mix, PAC utilization, and readmissions. The regression-based impact analyses also did not test if any change in one outcome was associated with changes in another outcome (e.g. a decrease in number of SNF days and the change in unplanned readmission rates). Other analyses did, however, examine the relationship between two outcomes. For example, we examined if greater standardized NPRA was associated with changes in readmission rates by BPCI hospital. In addition, we also gathered data on the relationships between outcomes in our site visits and Awardee interviews. For example, Model 2 hospitals participating in MJRLE stated that quality improved with a reduction in PAC. We will examine the correlation between utilization and quality outcomes where there is sufficient sample size in future analyses.

The beneficiary survey analysis is also dependent on a comparison group, where patients in a BPCI episode were matched to patients with the same clinical condition, although the matching process was not as rigorous as that of the claim-based analysis. In addition, because there was no baseline survey, the analysis relies on the comparison of episodes from BPCI and non-BPCI patients at a given point in time and does not control for time invariant differences. Finally, there is recall bias as beneficiaries are asked to remember their condition prior to the start of episode and the care they received from a hospitalization that may have occurred up to 90 days prior to completing the survey. The recall bias would most greatly impact the pre/post improvement measures that ask for information on functional status prior to the qualifying hospitalization. However, there is no reason to assume the recall bias is any different between BPCI and the comparison group.

The market share analysis requires accurate markets definitions. While core-based statistical areas (CBSAs) are an appropriate boundary for many markets, CBSAs may not accurately define the local health care markets in large urban areas, such as New York City or Chicago. In an attempt to mitigate this limitation, we excluded CBSAs that we identified as too large to accurately define a local health care market. These exclusions may dropped BPCI providers from the analysis that had a response to the initiative not represented in the analysis samples.

The qualitative analytic techniques employed for the site visits, quarterly interviews, and technical expert panels (TEPs) can provide descriptions of themes, patterns, or taxonomies, but cannot provide representative data on the impact of BPCI. Data collected through interviews reflect the opinions of those who were sampled and interviewed and may not necessarily be reflective of the experiences of other BPCI participants.

Two of the new analyses contained in this report are limited to select situations where there was sufficient sample size. We were able to calculate market share and to examine provider characteristics and factors associated with doing well or not under BPCI (based on standardized NPRA) for only three clinical episodes under Model 2 for hospital EIs.

The changes in assessment-based quality outcomes are only available for approximately three-quarters of patients discharged to PAC who remained in the setting long enough to have two assessments. These patients may not be representative of all patients discharged to PAC due to the reasons they do not have two patient assessments. Patients were not included in the assessment-based quality outcomes if the patient was readmitted to the hospital, died, had little

or no cognitive function, or if the episode occurred later during our measurement period and therefore a second assessment was not yet available in the data.

The DiD analysis estimates the impact of BPCI on Medicare payments by comparing the change in payments for BPCI participants relative to the change for a matched comparison group. It should be noted that the observed statistically significant declines in Medicare payments for a few clinical episodes based on the DiD does not necessarily indicate savings to the Medicare program. In order to calculate the impact to the Medicare program, the analysis would need to also incorporate any changes in volume that could be associated with BPCI as well as the NPRA.

The BPCI initiative tests a wide range of configurations, including the three bundled payment Models and multiple options for providers and other organizations to participate in up to 48 clinical episodes. The breadth of the initiative allows CMS to relatively quickly assess responses to payment incentives across a range of situations. This strength of the initiative, however, contributes to the main limitation of its evaluation. Because of the vast range of situations encompassed under the initiative, including the selective and heterogeneous group of participants and limited, varied experience of participants, it is challenging to reach conclusions about the impact of BPCI overall. We can draw conclusions in some situations, such as how participants differ from providers that did not choose to participate, when there is sufficient information and consistent findings. Limitations related to sample size and tenure will continue to be mitigated as the BPCI participants have more time under the initiative, which will increase the number of episodes.

Of equal concern is that because we are measuring multiple outcomes across the range of Model, participant, and clinical episode combinations, by chance alone some results will appear significant, although in reality these are not true effects. This issue compounds as the number of outcomes increases. Fortunately, the strength of our mixed methods approach is the ability to triangulate results across analyses. Taken together, quantitative results from claims, patient assessments, and beneficiary surveys, combined with information gleaned from site visits, interviews, and insights from the TEPs, provide a strong evaluation of BPCI. Consistency across findings lends strength to our conclusions, while inconsistencies raise questions for further inquiry. This report compares results across outcomes, across similar episodes, and across Models for the same clinical episodes. These comparisons likewise add strength to conclusions and illustrate the variations in impact.

C. Future Analyses

One of the most important advances in this evaluation over the next year will be analyzing the impact of BPCI among the BPCI-participating PGPs. We were unable to include the PGP episodes due to inaccuracies in the list of participating physicians for the BPCI PGPs at the time of our analyses for this evaluation report.[2] As of Q3 2015, PGPs accounted for approximately 40% of Model 2 EIs and 13% of Model 3 EIs. The final annual report will include the experience of PGPs and the impact of BPCI on PGP-initiated episode costs and quality. This will help complete the picture of the impact of BPCI on multiple outcomes across all types of EIs.

[2] The lists of BPCI-participating physicians by PGP from Q1 2016 onward were corrected in Q1 2017.

We will also estimate the Medicare program savings for the clinical episodes with a statistically significant decline in total Medicare spending per episode. The estimate will incorporate any BPCI impact on volume of episodes.

We will also strengthen and expand several analyses in the final annual report with the larger volume of episodes due to more BPCI experience. A larger sample will allow us to examine the factors that contribute to whether a participant does well or not under BPCI for a broader set of clinical episodes. Additional waves of the beneficiary survey will allow us to assess the impact on beneficiary satisfaction and quality of care for additional clinical episodes and EI types. Finally, we will further investigate any changes in patient mix, focusing particularly on Model 3 because PAC EIs have the greatest discretion over admission decisions.

D. Conclusion

The mixed methods evaluation we have employed indicates that BPCI participants have responded to BPCI incentives, but there are relatively few instances in which these responses significantly changed key outcomes. Because of the vast range of situations encompassed under the initiative, including the selective and heterogeneous group of participants and limited and varied experience of participants, it is challenging to reach conclusions about the overall impact of BPCI. It is also important to keep in mind that the kind of changes envisioned under the initiative often need to occur within complex organizations and require collaboration across organizations that may have differing objectives.

The evaluation of the BPCI initiative is far from complete, but this third annual report identifies BPCI design features that affect initiative results and conclusions. One key feature is the target price used in calculating NPRA. Because the target price is fixed and based on historical payments, providers tended to choose clinical episodes in which they have historically exhibited higher than average costs and, therefore, have the greatest potential for efficiency gains. A second feature is that quality of care is not directly tied to financial incentives under BPCI. While changes in quality of care have not been correlated with reduced services or payments, there have not been systematic improvements as might be expected with incentives to coordinate care across an episode. A third consideration has to do with the hospital infrastructure needed under Model 4 for managing physician payment, which limited its success. Another issue is that the reliance on MS-DRG assignment to determine the clinical episode has hampered some providers in Model 2 and Model 3 in determining which patients are in BPCI episodes, particularly patients who present with multiple diagnoses that complicate the ultimate assignment of the MS-DRG for the anchor/qualifying hospitalization. The delay in identifying the patients in BPCI episodes, particularly because they might benefit the most from episode-based care coordination, may diminish the impact of BPCI.

In future annual reports there will be more information about the care redesign participants are able to achieve and its impact on the cost and quality of care across clinical episodes. The final annual report will also include information on PGP-initiated episodes under Models 2 and 3. PGPs are significant participants in BPCI and their results may differ from results achieved by institutional providers. Future evaluation work will focus on clarifying the effects of the BPCI initiative with additional data and experience and understanding the provider strategies that are most effective in achieving success under the initiative.