ECONOMETRICA, INC.

REPORT

Annual Report 2014

Evaluation and Monitoring of the Bundled Payments for Care Improvement Model 1 Initiative

Contract No.:
HHSM-500-2011-00015I

Order No.:
HHSM-500-T0008

Project No.:
2248-000

Submitted To:

Centers for Medicare & Medicaid Services

Attn.: Arpit Misra
Contracting Officer's Representative
Center for Medicare & Medicaid Innovation
7500 Security Boulevard, Mail Stop 06-05
Baltimore, MD 21244

Submitted By:

Econometrica, Inc.

7475 Wisconsin Avenue
Suite 1000
Bethesda, MD 20814
www.EconometricaInc.com

July 9, 2015

ECONOMETRICA, INC.

July 9, 2015

Mr. Arpit Misra
Contracting Officer's Representative
Centers for Medicare & Medicaid Services
Center for Medicare & Medicaid Innovation
7500 Security Boulevard, Mail Stop 06-05
Baltimore, MD 21244

Reference: Contract No.: HHSM-500-2011-00015I; Order No.: HHSM-500-T0008; "Evaluation and Monitoring of the Bundled Payments for Care Improvement Model 1 Initiative" (Project No.: 2248-000).

Dear Mr. Misra:

Econometrica is pleased to submit this Annual Report to the Centers for Medicare & Medicaid Services, Center for Medicare and Medicaid Innovation, regarding work being conducted under the above-referenced contract.

Appendixes A, B, C, D, and E are being submitted as separate files.

If you wish to discuss any aspect of this submission, please feel free to contact me at (301) 395-2281.

Sincerely,

Econometrica, Inc.

Monique Sheppard, Ph.D.
Project Director

cc: Contract File

Table of Contents

List of Tables

List of Figures

List of Acronyms

ACOs	accountable care organizations
AMS	Applied Medical Services
BPCI	Bundled Payments for Care Improvement
CARE	Continuity Assessment Record and Evaluation
CCW	Chronic Conditions Data Warehouse
CMMI	Center for Medicare & Medicaid Innovation
CMO	Chief Medical Officer
CMS	Centers for Medicare & Medicaid Services
DiD	difference-in-differences
DME	durable medical equipment
ED	emergency department
FC	Facilitator Convener
FFS	fee-for-service
HCAHPS	Hospital Consumer Assessment of Healthcare Providers and Systems
HCC	Hierarchical Condition Category
HHA	home health agency
HHS	Department of Health and Human Services
ICU	intensive care unit
IP	Implementation Protocol
IPPS	Inpatient Prospective Payment System
IQR	Inpatient Quality Reporting
IVIg	intravenous immunoglobulin
MDC	Major Diagnostic Category
MDX	Medical Data Exchange
Model 1	BPCI Model 1
MS-DRG	Medicare Severity Diagnosis Related Group
NJHA	New Jersey Hospital Association
NN	nearest neighbor
NPI	National Provider Identifier
NQF	National Quality Forum
OR	odds ratio
PAC	post-acute care
PC	Program Coordinator
PFS	physician fee schedule
PHC	Physician Hospital Collaboration
PHES	Patient Health and Experience Survey
PIRE	Pacific Institute for Research and Evaluation
PQ	performance quarter
PUF	public use file
SNF	skilled nursing facility

Executive Summary

The Center for Medicare & Medicaid Innovation (CMMI), under the Centers for Medicare & Medicaid Services (CMS), is authorized under Section 1115A of the Social Security Act to test alternative payment and service delivery models[1] that have the potential to reduce health care expenditures while maintaining or improving quality of care for beneficiaries. One such model is the Bundled Payments for Care Improvement (BPCI) Model 1 initiative. Econometrica, Inc., and its partners—IMPAQ International, LLC; Optimity Matrix; and Pacific Institute for Research and Evaluation (PIRE)—are contracted under CMS to evaluate and monitor BPCI Model 1. This Annual Report of BPCI Model 1 presents interim findings for the first five performance quarters (PQs) of this model, beginning on April 1, 2013.

This Executive Summary first identifies BPCI Model 1 roles, provides a high-level model description, and details model Awardee characteristics. It then presents an overview of evaluation and monitoring activities under this contract, a synthesis and discussion of interim findings, and future activities under this contract.

BPCI Model 1 Roles

This Annual Report considers the following BPCI Model 1 roles.

- *Awardee.* Awardees are acute-care inpatient hospitals that submit applications to CMS for enrollment in BPCI Model 1 (Model 1). Once accepted, these hospitals sign an Awardee Agreement with CMS to enroll in BPCI Model 1.
- *Facilitator Conveners (FCs).* FCs are organizations that facilitate Awardee Model 1 participation but do not bear the risks or requirements in the Awardee Agreement.
- *Enrolled Practitioners.* Enrolled practitioners are physician or non-physician practitioners that are medical suppliers who furnish health care services to Model 1 Awardee beneficiaries, receive Medicare payments under the Medicare Physician Fee Schedule (PFS), engage in BPCI Model 1 care redesigns, and have a gainsharing agreement.[2] CMS vets practitioners that Awardees propose for enrollment. Currently, only physicians are enrolled across Awardees.

BPCI Model 1 Description

Model 1 focuses on care received at Awardee hospitals during an acute-care inpatient hospitalization ("episode") for all Medicare Severity Diagnosis Related Groups (MS-DRG), unless excluded[3] by Awardee. Through care redesigns, Awardees attempt to achieve efficiency gains in health care delivery, primarily in the form of reduced health care redundancies, improved care processes, and internal hospital cost savings. These efficiency gains may translate to reduced Medicare costs while maintaining or improving quality of care for Medicare beneficiaries. Awardees are permitted to share internal hospital cost savings engendered under this model with enrolled practitioners ("gainsharing"). Gainsharing is expected to promote

[1] Alternative, for example, to the traditional Medicare fee-for-service (FFS) system.
[2] Explained below.
[3] For example: NJHA Awardees all exclude Psychiatry and Normal Newborns diagnosis related groups.

alignment between Awardees and enrolled practitioners for successful implementation of care redesigns and model requirements.

Awardees face an automatic, predetermined discount to their Medicare Inpatient Prospective Payment System (IPPS) payments for episodes at their hospital.[4] Moreover, Awardees are financially at risk for increases in both aggregate Medicare Part A and Part B costs 30 days after an episode and face quality and activity reporting requirements. These requirements and periods of financial risk in Model 1 are meant to ensure that any additional costs are not passed back to CMS via other points in the health care continuum and that quality of care is maintained or improved.

Model Awardees

The BPCI Model 1 performance period began on April 1, 2013, with 23 Awardee hospitals (all located in New Jersey) under 1 FC, the New Jersey Hospital Association (NJHA). Of the 23 Awardees, 6 participated in the Physician Hospital Collaborative (PHC), a gainsharing demonstration similar to BPCI Model 1 that ended in July 2012. Those same six Awardees opted to participate in an extension of PHC from July 2012 to March 2013 that led up to the BPCI Model 1 performance period.

On January 1, 2014, 1 Awardee hospital located in Kansas initiated an Awardee Agreement with CMS for BPCI Model 1 for a maximum of 24 Awardees; this hospital does not have an FC. As of July 1, 2014, 9 of the 24 hospital Awardees terminated their Awardee Agreement[5] and are no longer active in BPCI Model 1. This report considers all 24 BPCI Model 1 Awardees and groupings of these Awardees.

ES.A. Evaluation and Monitoring Activities

Two overarching evaluation and monitoring questions guided the analytical framework, data collection, and analyses presented in this report.

1. What is the impact of BPCI Model 1 on Medicare costs (payments to hospital providers) and the quality of care provided to patients?
2. What are characteristics of the model, providers, or environment that influenced model impacts?

Findings that address these questions are presented within and across measure domains inspired by CMMI evaluation and monitoring measures:[6] [7]

[4] Physicians at Awardee hospitals are paid separately under the Medicare PFS.

[5] Further, as of March 1, 2015, 12 of the 24 Awardee hospitals terminated their Awardee Agreement..

[6] These domains and measures formed by measure and domain guidance from Rapid Cycle Evaluation Innovation Center. *CMMI Core Measures Version 9*. Baltimore, MD: CMS, CMMI. March 28, 2013.

[7] Other domains include Care Processes, Patient Care Experience, and Hospital-Internal Cost. Measures in these domains will be included in future reports as data become available. Care Process domain measures primarily rely on Inpatient Quality Reporting (IQR) data and are unavailable for the PQs covered in this report. Patient Care Experience domain measures primarily rely on responses from the Patient Health and Experience Survey (PHES)—a beneficiary-level survey developed under this contract that is in its early stage of implementation. Measures within

- *Implementation and organizational responses* (e.g., organization and infrastructure changes in response to BPCI Model 1, physician engagement, care redesign type and implementation).
- *Episode case mix and patient characteristics* (e.g., indicators for severity of patient's condition).
- *Health care outcomes and resource utilization* (e.g., all-cause and condition-specific mortality and length of stay).
- *Medicare payments* (e.g., total episode Medicare payments).

As an example, evaluation analyses considered the effect of gainsharing between Awardees and enrolled practitioners on care redesign implementation and how implementation of care redesigns may have translated Medicare payments to Awardees (i.e., costs to Medicare) or other measures. Section ES.A.1 reviews the data, methods, and analyses that informed results in Section ES.B.

ES.A.1. Data and Methods

Quantitative and qualitative analyses of BPCI Model 1 employed data from various sources, including Awardee interviews, focus groups, enrollment data, and Medicare claims across Model 1 Awardee hospitals and similar non-Awardee hospitals.

Awardee and enrolled practitioner interview and focus group data enriched understanding of care redesign implementation processes, successes, and issues. These data further provided insight into unanticipated BPCI- and non-BPCI related issues that Awardees faced during model implementation. These data also contextualized the quantitative findings of BPCI Model 1 impacts on measures assessed. Generally, these data were triangulated over multiple time periods and are summarized below:

- *Telephone interview data* analyzed in this report were collected over two time periods ("waves"). The first wave occurred from April 1, 2013, to December 23, 2013, and included 23 Awardee hospitals; the second wave occurred from January 3, 2014, to May 31, 2014, and included 14 hospitals.[8]
- *Focus group data* analyzed in this report were collected during visits to distinct Awardee hospitals ("site visits") at four points in time: July 2013, one hospital (pilot visit); October 2013, four hospitals; March 2014, two hospitals; and October 2014, two hospitals.
- *Medicare claims data* from the 100-percent research identifiable Medicare Claims and Enrollment Database from the Chronic Conditions Data Warehouse (CCW) covering a

these domains are expected by Quarter 2 of 2015. Completed data for hospital internal cost savings over the first three PQs are expected within Quarter 1 of 2015; these data come from BPCI Model 1 Awardee hospitals.

[8] Nine hospitals terminated their Awardee Agreement with CMS before or during the second data collection wave. One hospital entered the BPCI Model 1 demonstration on January 1, 2014.

baseline period of January 1, 2011, through March 31, 2013 ("Baseline") and a model performance period of April 1, 2013, through June 30, 2014 ("Since BPCI Inception").[9]

- *Other data* included hospital characteristics data (e.g., Provider of Service file) or hospital and physician/practitioner enrollment data provided by CMS.

Quantitative BPCI Model 1 impacts on measures in this report come from comparative analyses of Model 1 Awardees to *similar* non-Awardee hospitals ("comparison hospitals" or "comparison group"). By construction, comparison hospitals allow for inferences of a counterfactual scenario: What would Model 1 Awardee measure performance have been had they not participated in BPCI Model 1?

Construction of the comparison hospital group included identifying a pool of non-Awardee hospitals (hospitals that did not participate in BPCI Model 1) and then determining the similarity of non-Awardee hospitals to Awardee hospitals. In this construction, similarity was determined by statistical examination of observable characteristics between each Awardee hospital and all non-Awardee hospitals. Observable characteristics were examined over the Baseline period of this report—January 1, 2011, through March 31, 2013—and included the number of hospital beds, patient case mix, and Baseline values of select measures (e.g., average episode length). The resulting comparison hospital group was composed of non-Awardee hospitals deemed statistically similar in their Baseline period to Awardees across these characteristics. Multiple statistically similar non-Awardee hospitals were *matched* to each Awardee to minimize anomalous performance of any singular non-Awardee hospital over time.

Comparison hospitals were included in Awardee-level and Awardee cohort-level comparative analyses. Awardee-level analyses involved analyzing each Awardee relative to its matched comparison hospitals across observed and (risk) adjusted measure statistics over time. Further, these analyses assessed BPCI Model 1 Awardee-level impacts on measures by utilizing quasi-experimental difference-in-differences (DiD) regression models. These regression models leverage the aforementioned counterfactual scenario directly by comparing Awardee measure differences before and after BPCI Model 1 inception (i.e., Baseline vs. Since BPCI Inception) to like differences for comparison hospitals. Additionally, these models controlled for both episode and patient characteristics to account for residual differences between Awardee and comparison hospitals that were not otherwise captured by the comparison hospital selection process.

Awardee *cohort*-level analyses also included analyses of observed and adjusted measure statistics and DiD measure impacts. Full cohort ("program-wide") analyses compared all Awardees to all comparison hospitals, while subcohort analyses compared Awardees classified in a subcohort to comparison hospitals matched to those Awardees. Three subcohort types were considered:

1. *Awardee enrollment status cohorts.* This cohort type included *Active* and *Exiting* cohorts. Designation into either cohort was determined by Awardee BPCI Model 1 participation

[9] Medicare claims data were pulled November 14, 2014, to allow for a minimum of 4 months of maturation for claims data pulled through June 2014. Medicare data are expected to be 80 to 96 percent complete for a 3- to 6-month run-out period.

status as of July 1, 2014. Awardees that terminated their Awardee Agreement with CMS prior to July 1, 2014 (nine Awardees) were classified in the Exiting cohort. Awardees that remained active in BPCI Model 1 through July 1, 2014 (15 Awardees) were classified in the Active cohort.

2. *Care Redesign cohorts.* Care redesigns vary across Awardees in number pursued, implementation methods, and objectives. Thus, to infer impacts from care redesigns, Awardees were classified by their overall approach to affect Model 1 goals through their care redesigns. Awardees that, on average, had care redesigns focused on larger patient populations or a broad spectrum of the in-hospital care continuums (e.g., from admissions to discharge) were classified in the Expansive Care Redesign cohort (eight Awardees). Conversely, Awardees that, on average, had care redesigns focused on targeted patient populations (e.g., cardiac patients) or specific in-hospital process points (e.g., patient discharge) were classified in the Targeted Care Redesign cohort (16 Awardees).

3. *Recent gainsharing experience.* Awardees in the PHC demonstration that lead up to BPCI Model 1 were classified into the PHC cohort (six of nine Awardees that terminated before July 1, 2014).

Section ES.B focuses on Full, Active, and Exiting cohort-level analyses and uses Awardee-level analyses to identify outlier Awardee performance. Section III of this report exhibits findings for all cohorts.

ES.B. Results

The following results focus on program-wide (i.e., Full cohort) assessments of Awardee care redesign implementation and participation in BPCI Model 1, and DiD measure *impact* estimates. Impact estimates come from comparative analyses of Awardee measure performance changes from Baseline to the Since BPCI Inception period to like changes for comparison hospitals. Differential impact estimates between Active and Exiting cohort Awardees are also noted.

ES.B.1. Awardee Activities and Engagement Under Model 1

At the inception of Model 1, Awardees included 23 hospitals located in New Jersey that voluntarily participated in the program with 1 FC. One additional hospital, located in Kansas, initiated an Awardee Agreement with CMS for the model in January 2014. Over the first 5 PQs, 9 Awardee hospitals terminated their Awardee Agreement, bringing the number of active Awardees down to 15. Exit interviews conducted with these Awardees indicated that they terminated their Awardee Agreements with CMS for the following reasons:

1. Hospital-physician employment structure already required/motivated physician adherence to protocols and participation in care redesigns through existing employment/contractual relationships.
2. Inability to associate perceived or realized internal hospital cost savings from care redesigns. Or, expecting relatively long timescales involved in realizing potential internal hospital cost savings from care redesigns.

3. Lack of physician enrollment and engagement resulting in insufficient critical mass to generate the change in practice required to deliver internal hospital cost savings from care redesigns.

Awardee hospitals enrolled and engaged physicians to assist in their Model 1 activities (e.g., care redesigns) through various recruiting techniques. These recruitment efforts, however, have been hindered, in part due to physician skepticism of gainsharing, misunderstanding of model components, or the added effort required for model (reporting) requirements. Increased physician enrollment and engagement was noted after an initial gainsharing distribution (end of PQ 3 through PQ 5).

Each Awardee implemented between two and nine care redesigns. Across Awardees, these care redesigns varied in type (e.g., in-hospital patient flow improvement versus fall prevention programs), ease of implementation, timing of implementation, definition of "fully implemented," and desired outcome(s) from implementation. Reports on care redesign implementation generally stated care redesign implementation timeframes spanning from 3 to 6 (or even 9) months. In primary data collected early in the first year of BPCI Model 1, administrative and clinical interviewees emphasized that care redesigns aimed to affect quality of care first and cost of providing care second, as efficient and efficacious care should reduce cost. Recently, in PQs 4 and 5, administrators have appeared more cost conscious by discussing whether the pursued care redesigns have induced sufficient cost savings to validate a business case for continuing in Model 1, where they face an across-the-board IPPS discount.

ES.B.2. Changes in Episode Case Mix and Patient Characteristics

Generally, the age of patients treated across Awardee and comparison hospitals decreased ($p < 0.01$) from Baseline to Since BPCI Inception. The average patient age across Awardees was 76.53 years in Baseline and 76.42 years Since BPCI Inception, while the average patient age across comparison hospitals was 75.18 years in Baseline and 74.87 years Since BPCI Inception. Average CMS Hierarchical Condition Category (CMS-HCC) scores also decreased (less than 0.09 points, $p < 0.01$) from Baseline for both Awardee and comparison hospitals.

Average episode MS-DRG weight, an indicator of resource intensity of diagnostic classifications, increased by 0.04 points ($p < 0.01$) for both Awardee and comparison hospitals from Baseline.

Taken together, when compared to average episode and patient population characteristics of comparison hospitals, BPCI Model 1 Awardee patients were older throughout the analysis period (average difference of 1 year), had a slightly higher CMS-HCC score (average difference of 0.03 points), and had a lower MS-DRG weight (–0.04 points).

ES.B.3. Changes in Medicare Payments

The Medicare episode payment measure included actual payments for services provided during episode. These payments included adjustments to providers for other CMS initiatives (e.g., Hospital Readmission Reduction Act payment adjustments) that occurred during the Baseline and Since BPCI Inception periods. Consequently, impacts on these payment measures may not be wholly attributable to BPCI Model 1; they are presented only for preliminary insight and thus

should be interpreted with caution. Future work will examine standardized allowed Medicare payment amounts that will provide a truer assessment of BPCI Model 1's impact on Medicare payments.

On average, total Medicare episode payments for Awardee and comparison hospitals increased from Baseline to the Since BPCI Inception period. These increases were greater for comparison hospitals than for Awardees. Program-wide DiD impact estimates indicate an average statistically significant decrease of $123 ($p < 0.01$) in total Medicare episode payments over the model performance period relative to Baseline and like changes for comparison hospitals. Further analysis of this impact estimate shows that a decrease of $83 ($p < 0.1$) was attributable to Medicare episode payments to hospitals. The average reduction of inpatient episodes was approximately $30[10] from the IPPS discount on Medicare payments to Awardee hospitals and may account for a portion of the –$83 impact estimate. The remaining $40 ($p < 0.01$) of the impact estimate of –$123 stemmed from decreased physician (or other provider) billing during the episode.

The impact estimate of –$123 appears to be driven by Exiting cohort Awardees; the Exiting cohort Medicare episode payment impact estimate was –$219 ($p < 0.01$). Indeed, the hospital-only Medicare episode payment impact for Active cohort Awardees was not statistically significant at +$11, while the Exiting cohort Awardee impact was –$171 ($p < 0.05$).[11] Non-hospital Medicare episode payment impacts (e.g., those made to physicians for services during the episode) were relatively similar across Exiting and Active cohort Awardees (~$40, $p < 0.01$).

The total post-episode Medicare payment measure is a *combination* of Medicare payments to providers that rendered services to patients up to 30 days after that patient's episode discharge. Generally, post-episode Medicare payments increased for Awardee and comparison hospitals since Baseline. These increases translated to an average (of total) post-episode Medicare payment impact estimate of +$95 ($p < 0.05$), relative to Baseline and comparison hospitals. Active cohort Awardees influenced this program-wide result heavily with a post-episode Medicare payment impact estimate of +$248 ($p < 0.01$). Preliminary examination of unadjusted average Medicare payments across post-episode provider/service types for Awardee and comparison hospitals indicates that these increases stem from payments to home health, skilled nursing, and other inpatient services/facilities over the model performance period, relative to Baseline and comparison hospitals. These relative payment increases to facilities may be indicative of increased utilization or care intensity (and corresponding cost) or geographical payment differences (e.g., wage index differentials). Examination of unadjusted utilization rates for these facilities indicates no noteworthy changes from Baseline or relative to comparison hospital patients. Analysis of this measure is not complete and should be interpreted with caution. This measure will be explored in more detail for the 2015 Annual Report.

[10] The IPPS discount is applied to adjusted portions of IPPS operating payments.
[11] This estimate of –$171 can be interpreted as the portion of the –$219 impact estimate for payments made to hospitals, as opposed to other providers (e.g., physicians).

ES.B.4. Changes in Health Care Outcomes and Resource Utilization

In this domain, episode length of stay, the number of days a patient spent in an acute-care hospital, and whether a patient had intensive care unit (ICU) services during their episode were assessed. A patient's likelihood for experiencing a mortality event within 30 days of episode admission or readmission/rehospitalization event for any reason within 30 days of episode discharge was also assessed.

As with payment measures, these assessments compared episode or post-episode level data over the Model 1 performance period in this report to Baseline and comparison hospitals. Likelihood impacts are presented in multiplicative odds ratios (OR)[12], where an OR above 1.0 indicates increased odds of patients from Awardee hospitals experiencing an event relative to Baseline and comparison hospitals and an OR below 1.0 indicates a decreased likelihood.

Generally, there were no statistically significant changes in mortality rates from Baseline to the Since BPCI Inception period for Awardee and comparison hospital cohorts. Program-wide DiD impact analyses exhibited no program-wide statistically significant changes in the likelihood of a patient experiencing a mortality event within 30 days over the model performance period, relative to Baseline and patient discharged from comparison hospitals. Subcohort analyses, however, exhibited an increased impact estimate for Active cohort Awardees (1.03 OR, $p < 0.05$), stemming from an elevated PQ 4 likelihood estimate (1.09 OR, $p < 0.05$). Examination of hospital-level impact estimates[13] indicated that two-thirds of the Active cohort Awardees had elevated mortality impact estimates (OR above 1.0) but only two Awardee estimates were statistically significant at the 10-percent level. Additional analyses indicated that these elevated mortality impact estimates were driven by post-episode mortality, not in-hospital (i.e., episode) mortality.

Changes in ICU use from Baseline to the Since BPCI Inception period varied in direction and magnitude across Awardee and comparison hospital cohorts. Program-wide DiD impact analyses exhibited a statistically significant increase in the likelihood of a patient having ICU services at an Awardee hospital (relative to Baseline and comparison hospitals, 1.10 OR, $p < 0.01$). This result appears to be driven by Active cohort Awardees, with an OR impact estimate of 1.20 ($p < 0.01$). This cohort was noted to have a higher average (CMS-HCC) risk profile relative to Exiting cohort Awardees. Hospital-level DiD impact analyses indicated that all but three of the Active Awardees exhibited increased likelihoods of their patients having an ICU stay during their episode (OR above 1.0). However, the attribution of these elevated ICU estimates to Model 1 must be qualified. Placebo DiD tests[14] for the Full and Active cohort Awardees indicated that these elevated ICU likelihood estimates were found *before* BPCI Model 1 implementation and may consequently not be attributable to Model 1. Unadjusted ICU statistics do exhibit a 0.76 percentage point increase across Awardees and a 0.10 percentage point increase across

[12] Future work will also assess marginal impacts for these estimations.
[13] An individual Awardee's performance compared to its four matched comparison hospitals.
[14] Placebo DiD Tests test the reliability of associating DiD impact estimates with an intervention (e.g., BPCI Model 1) by assuming placebo/pseudo start dates *before* the intervention's *actual* start date. Statistically significant estimates in the placebo periods can indicate an inappropriateness of impact estimate attribution to BPCI Model 1.

comparison hospitals. This yields an unadjusted DiD result of 0.66 percentage points—indicative of increased ICU utilization—for Awardees relative to Baseline and comparison hospitals.

Changes in episode length of stay from Baseline to the Since BPCI Inception period varied in direction and magnitude across Awardee and comparison hospital cohorts but were generally minimal with differences of less than 0.5 days. There were no program-wide DiD impact estimates in this measure. Awardee-level impact analyses indicated that 11 of the 24 Awardees did achieve statistically significant decreases in episode length of stay ranging from decreases of 0.22 to 0.62 days (all statistically significant, $p < 0.01$).

There were no statistically significant impact findings for program-wide analysis of readmission likelihood. Exiting Awardees did, however, exhibit a statistically significant decrease in this measure (0.97 OR, $p < 0.1$).

ES.C. Discussion

The longevity of BPCI Model 1 relies on an acceptable level of Awardee participation, and its success relies on in-hospital clinician support. As previously noted, nine Awardee hospitals terminated their Awardee Agreement for the period of analysis covered in this report, through June 30, 2014. As of March 2015, 3 additional Awardees have terminated their Awardee Agreements with CMS for a total of 12 of 24 Awardee withdrawals. Terminating Awardees repeatedly noted the financial burden of the IPPS discount as a motivating factor for termination. Some of these Awardees also noted difficulties in attributing internal hospital cost savings (or even realizing any) to BPCI Model 1 such that the IPPS discount was offset. Many of the initial terminating Awardees did believe they had achieved sufficient progress in their care redesign implementation to proceed without model incentive components.

Model 1 does presume that enrolled practitioners—those with the potential to receive gainsharing from Awardees—are incentivized by gainsharing to carry out care redesigns and help affect efficiency gains at a hospital. Care redesigns, however, are not solely implemented by enrolled practitioners and are typically hospital-wide or multi-departmental endeavors. Thus, despite the lackluster physician enrollment or engagement reported by some Awardees (in Exiting and Active cohorts), non-enrolled physicians and non-physician clinical staff may aid the implementation of care redesigns and their effect on model goals. The extent to which this may occur is currently unknown. Future analyses will attempt to discern the extent by stratifying Model 1 impact effects by enrolled and non-enrolled physicians.
Impact estimates indicated that Medicare payment increases were *muted* (i.e., increased less than comparisons) for Awardees over the primary period of focus under this model, the inpatient stay (i.e., episode). Medicare payments to other providers after the episode period (e.g., physicians, nursing facilities, and rehabilitation hospitals) did increase relative to baseline and comparison hospitals. These Medicare payment findings provide interim insight on *potential* Model 1 effects. Payments over this post-episode period are monitored and compared to historical baselines by another CMS contractor, and information from CMS has indicated that Awardees did not exceed predetermined thresholds for payment increases over the 30-day post-episode period.

Resource utilization impact findings were mixed as evidenced by increased likelihoods of ICU services and some decreases in episode length of stay across Awardees. Among Active cohort

Awardees, increased mortality likelihoods for patients discharged from Awardee hospitals were noted. The elevated mortality impact estimate appeared to be attributable to outlier Awardee performance on this measure. Taken together, these impacts and increases in post-episode Medicare payments may be indicative of (unintended) adverse consequences of Model 1. However, such conclusions would be premature without additional analyses that incorporate patient health and experience data and other quality indicators. These data sources are expected in the 2015 Annual Report.

Aforementioned impact differentials between Active and Exiting cohort Awardees were not unexpected. As previously noted, exit interview data indicated that Awardee Agreement terminations were due, in part, to Awardee beliefs that their care redesigns achieved sufficient progress to continue without model incentive components or the Model 1 IPPS payment discount. Six of the nine Exiting cohort Awardees had recent experience in implementing care redesigns and similar incentive mechanisms from another CMS model, the PHC. These six Awardees were among the earliest Awardees to terminate their agreement. Additional quarters of data will allow for inferences as to whether BPCI Model 1 had sustained impact on Exiting cohort Awardees. Conversely, Active Awardees still believed they had enough to gain from BPCI Model 1 in terms of improving clinician and hospital alignment towards care redesigns. This sentiment may be indicative of a delay in the translation of Active cohort Awardee care redesign implementation effects to Model 1 goals.

ES.D. Forthcoming Analyses

ES.D.1. Domains and Measures

Table 1 presents measures for future analyses, expected in the 2015 Annual Report. These measures will provide a more comprehensive view of beneficiary outcomes and Awardee ability to achieve internal hospital cost savings.

Table 1: Domains and Measures in Future Reports

Domain	Measure
Care Processes	• Care coordination at discharge (PHES)
Medicare Payments and Internal Hospital Costs	• Internal hospital costs savings from enrolled physicians
Health Care Outcomes/ Patient Experience	• Functional status – mobility (PHES) • Functional status – pain intensity (PHES) • PAC assessment (MDS, HHA-OASIS, IRF-PAI)

ES.D.2. Types of Analyses

In addition to measures noted above, forthcoming analyses aim to examine the following:

- *Enrolled physicians.* Differences across measures for physicians enrolled in BPCI Model 1 will be examined. Further, these differences will be compared to non-enrolled physician measure performance at Awardee hospitals. These comparisons will allow inference into whether impact results for an Awardee hospital reflect an "overpowering" of non-enrolled physician measure performance, or vice-versa.
- *Alternative groupings of care redesigns across Awardees.* The initial grouping of Awardees by care redesign types for this report might not have great specificity due to

the variety of care redesigns pursued across Awardees. Further, Awardees may have changed their pursued care redesigns. Alternative models for care redesign classification will be assessed in 2015.

- *Testing for care shifting and other methods of "gaming" the system.* Physicians often have admissions privileges at multiple hospitals or can rely on transfers to keep lower-cost patients at BPCI hospitals. The availability of physician-level data will allow testing for various methods that could potentially bias results.

Section I. Introduction

The Bundled Payments for Care Improvement (BPCI) Model 1 initiative aims to reduce Medicare costs while maintaining or improving quality of care.[15] BPCI Model 1 is a 4-year program that was launched on April 1, 2013. Econometrica, Inc., and its partners—IMPAQ International, LLC; Optimity Matrix; and Pacific Institute for Research and Evaluation (PIRE)—are contracted by the Center for Medicare & Medicaid Innovation (CMMI), under the Centers for Medicare & Medicaid Services (CMS), to evaluate and monitor of BPCI Model 1 (Model 1). This Annual Report is an interim synthesis of findings that aims to provide insight on Awardee activities[16] and model impacts on Medicare costs, quality of care, outcomes, and resource utilization through performance quarter (PQ) 5 (June 30, 2014).

This section first identifies BPCI Model 1 roles and provides a high-level model description. It then presents a brief overview of evaluation and monitoring activity motivation under this contract; background information on prior, related models; in-depth detail on BPCI Model 1; and information on hospitals participating in Model 1. Section I concludes with an outline of the remainder of this report.

BPCI Model 1 Roles

This Annual Report considers the following BPCI Model 1 roles.

- *Awardee.* Awardees are acute-care inpatient hospitals that submit applications to CMS for enrollment in BPCI Model 1 (Model 1). Once accepted, these hospitals sign an Awardee Agreement with CMS to enroll in BPCI Model 1.
- *Facilitator Conveners (FCs).* FCs are organizations that facilitate Awardee Model 1 participation but do not bear the risks or requirements in the Awardee Agreement.
- *Enrolled Practitioners.* Enrolled practitioners are physician or non-physician practitioners that are medical suppliers who furnish health care services to Model 1 Awardee beneficiaries, receive Medicare payments under the Medicare Physician Fee Schedule (PFS), engage in BPCI Model 1 care redesigns, and have a gainsharing agreement.[17] CMS vets practitioners that Awardees propose for enrollment. Currently, only physicians are enrolled across Awardees.

BPCI Model 1 Description

Model 1 focuses on care received at Awardee hospitals during an acute-care inpatient hospitalization ("episode") for all Medicare Severity Diagnosis Related Groups (MS-DRGs), unless excluded[18] by Awardee. Through care redesigns, Awardees attempt to achieve efficiency gains in health care delivery, primarily in the form of reduced health care redundancies, improved care processes, and internal hospital cost savings. These efficiency gains may translate

[15] Similarly, while preventing adverse changes to quality of care.
[16] Awardee activities include the implementation of care redesigns, methods to increase physician enrollment and engagement, process for and effects of the physician incentive mechanisms, and successes and challenges encountered.
[17] Explained below.
[18] For example: NJHA Awardees all exclude Psychiatry and Normal Newborns diagnosis related groups.

to reduced Medicare costs while maintaining or improving quality of care for Medicare beneficiaries. Awardees are permitted to share internal hospital cost savings engendered under this model with enrolled practitioners ("gainsharing"). Gainsharing is expected to promote alignment between Awardees and enrolled practitioners for successful implementation of care redesigns and model requirements.

Awardees face an automatic, predetermined discount to their Medicare Inpatient Prospective Payment System (IPPS) payments for episodes at their hospital.[19] Moreover, Awardees are financially at risk for increases in both aggregate Medicare Part A and Part B costs 30 days after an episode and face quality and activity reporting requirements. These requirements and periods of financial risk in Model 1 are meant to ensure that any additional costs are not passed back to CMS via other points in the health care continuum and that quality of care is maintained or improved.

Evaluation and Monitoring Analytical Motivation

Two overarching evaluation and monitoring questions guided the analytical framework, data collection, and analyses presented in this report.

1. What is the impact of BPCI Model 1 on Medicare costs (payments to hospital providers) and the quality of care provided to patients?
2. What are characteristics of the model, providers, or environmental factors that influenced model impacts?

These questions are addressed for the first five program quarters through a study of Awardee hospital characteristics (e.g., financial stability, clinical staff mix and longevity, and patient populations); care redesigns implemented; clinical staff engagement; and impacts on Medicare costs, quality of care, outcomes, and resource utilization. These characteristics, activities, and quantifiable measures are grouped into domains inspired by CMMI evaluation and monitoring measures,[20] defined later.

The remainder of Section I provides motivation for BPCI Model 1, brief overviews of similar models, design features and requirements of Model 1, information on Awardee hospitals, and an outline of this report.

I.A. Background

Changes in health care policies are transforming the health care industry and placing more accountability on hospitals, health care systems, physicians, and payers to contain health care costs while improving access to and quality of care. The Department of Health and Human Services (HHS) recently released its timetable for transitioning Medicare away from its fee-for-service (FFS) model. HHS is aiming to tie 30 percent of traditional Medicare payments to care quality through Accountable Care Organizations (ACOs) or bundled payment arrangements by

[19] Physicians at Awardee hospitals are paid separately under the Medicare PFS.
[20] These domains and measures formed by measure and domain guidance from Rapid Cycle Evaluation Innovation Center. *CMMI Core Measures Version 9*. Baltimore, MD: CMS, CMMI. March 28, 2013.

the end of 2016, with 50 percent tied to care quality by the end of 2018. Altogether, the targets represent a 50-percent increase in health care reimbursement more concerned with the *value* of care than its *volume* by 2016.[21]

Currently, in Medicare's FFS payment system, hospitals are paid under the IPPS for Medicare Part A services, while physicians are paid separately under the PFS for Medicare Part B services. IPPS payments to hospitals are fixed amounts per patient visit and are determined by patient MS-DRGs. These payments are meant to cover the primary costs of a hospital's Medicare patients, including hospital-based services incurred or ordered by physicians during a patient's hospital stay.

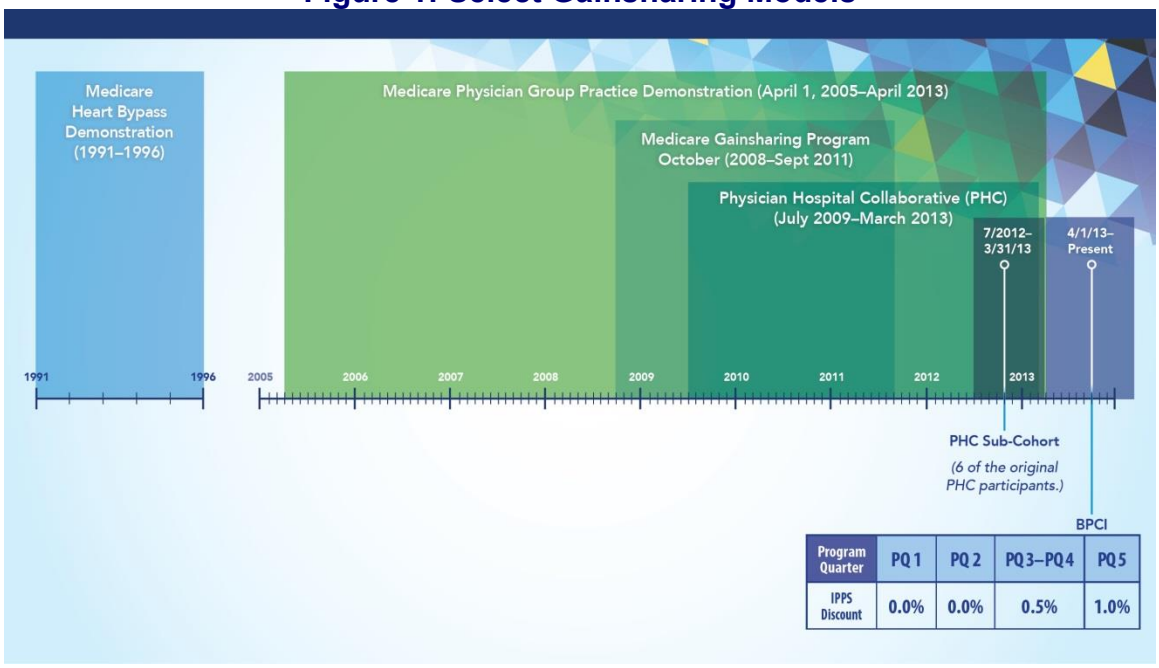
The aforementioned increased accountability for the quality of care provided to patients and health care costs come from a variety of payment and reporting models. Some of these models, such as bundled payments, have mechanisms meant to foster focus across multiple providers to align and engage health care redesigns toward CMS' goals. Within BPCI Model 1, gainsharing is one such alignment mechanism. With this mechanism, hospitals provide physicians with a financial stake in managing and reducing hospital costs by offering physicians a share of any internal hospital cost savings achieved from efficiency or productivity gains in delivering health care. Typically, the share that physicians receive depends on regulatory bounds and physician performance across Awardee-determined quality measures (e.g., the percentage of patients on antibiotics following coronary artery bypass surgery). Section I.A.1 below provides an overview of some previously completed similar gainsharing models.

I.A.1. Previous Gainsharing Models

CMS has used gainsharing mechanisms in demonstrations before (Figure 1). In the Medicare Heart Bypass Demonstration conducted between 1991 and 1996, physicians received fixed payment amounts that were included within hospital payments. Reportedly, hospital costs were reduced, quality of care improved, and shifts to post-acute care (PAC) settings yielded no negative offsets to Medicare savings.[22] At a group practice level, the Medicare Physician Group Practice demonstration (April 1, 2005, to December 31, 2013) used a shared-savings mechanism similar to gainsharing, where physicians in a group practice could receive shares of savings if they met predefined targets.[23]

[21] Department of Health and Human Services. (2015, January 26). *Better, Smarter, Healthier: In a historic announcement, HHS sets clear goals and timeline for shifting Medicare reimbursements from volume to value.* Press Release. Available at: http://www.hhs.gov/news/press/2015pres/01/20150126a.html.
[22] Detailed, numerical impacts are not available at this time.
[23] This was initially a 5-year demonstration (through March 31, 2010) that led to a 2-year extension program. The PGP Transition Demonstration ran from January 1, 2011, through December 31, 2013.

Figure 1: Select Gainsharing Models



CMS also conducted two similar gainsharing demonstrations that were hospital-based: the Medicare Gainsharing Program (October 2008 to September 2011) and the Physician Hospital Collaboration (PHC; July 2009 to July 2012/March 2013). The primary difference between these two programs was the 90-day PAC discharge period included in PHC versus the 30-day discharge period in the Medicare Gainsharing Program. The additional 60 days assessed in PHC emphasized community and integrated delivery systems for that demonstration. An evaluation of the Medicare Gainsharing Program found evidence of internal hospital savings, found no evidence of statistically significant changes to Medicare inpatient payments, and could not associate statistically significant changes in quality of care.

I.B. BPCI Model 1

Under Model 1, Awardee hospitals redesign existing care processes to achieve efficiency gains in health care that may potentially translate into decreases in Medicare costs (Medicare payments to Awardees) while not adversely affecting patient quality and experience of care. In pursuit of efficiency gains from care redesigns, Awardees may engender reductions to internal hospital costs (cost savings). Waivers under Model 1 allow Awardees to distribute a portion of these internal cost savings with enrolled practitioners/physicians (i.e., gainsharing) as a financial incentive to encourage enrolled physician participation,[24] implementation of care redesigns, and achievement of BPCI Model 1 goals.

BPCI Model 1 Awardees face an automatic, predetermined discount to their Medicare IPPS payments[25] for inpatient stays, defined as an episode of care for each MS-DRG.[26] In addition to

[24] Non-physician clinical staff may also be engaged in BPCI Model 1 implementation of care redesigns.
[25] Physicians at Awardee hospitals are paid separately under the Medicare PFS.

this episode, Awardees are financially at risk for increases in Medicare payments, both aggregate Medicare Part A and Part B expenditures, occurring 30 days after discharge. Section I.B.1 details these and other model features.

I.B.1. Features of BPCI Model 1

I.B.1.1. IPPS Discount and Other Requirements

Medicare pays Awardees a discounted amount for acute hospital care based on the payment rates established under the IPPS across all MS-DRGs (unless excluded[27] by Awardee). The financial implications for Awardees will vary, in part, by their volume of Medicare patients. Physician payments under this model are unaffected and are reimbursed separately under the PFS for Medicare Part B services. The discount to the IPPS operating payment[28] was 0.0 percent for the first two PQs, 0.5 percent at the start of PQ 3 (October 1, 2013) to PQ 4, and 1.0 percent in PQ 5. The IPPS discount percentage was set to rise to 2.0 percent on April 1, 2015, but is currently frozen at 1.0 percent indefinitely.[29]

Additional payment requirements include the following:

- *Episode payments.* IPPS payments will be discounted prospectively.
- *Episode monitoring.* Medicare Part A and Part B payment for the inpatient hospital stay that exceeds trended historical aggregate Part A and Part B payment beyond a risk threshold (taking the discount into consideration) must be paid by the Awardee to Medicare.
- *Post-episode monitoring.* Medicare Part A and Part B payment during the post-episode monitoring period that exceeds trended historical aggregate Part A and Part B payment beyond a risk threshold must be paid by the Awardee to Medicare.

I.B.1.2. Enrolled Practitioner Requirements

Each Awardee provides CMS with a list of practitioners for enrollment in BPCI Model 1 under the Awardee hospital. CMS accepts physicians based on a series of requirements (e.g., eligible to furnish Medicare services, valid National Provider Identifier (NPI)). Awardee-specific requirements for practitioners may vary. For example, Awardee hospitals in New Jersey mandate that physicians have at least 10 admissions in the prior year, be on staff for a minimum of 12 months, and have maintained their credentials in good standing at Awardee hospitals.

I.B.1.3. Care Redesigns

Awardees implement care redesigns to achieve cost savings—and potentially improve quality of care—through improved efficiency or productivity of health care provided at hospitals. The care redesigns Awardees pursue under this model typically affect processes of care. These redesigns

---

[26] Unless excluded by the Awardee. For example: NJHA Awardees all exclude Psychiatry and Normal Newborns diagnosis related groups.
[27] For example: NJHA Awardees all exclude Psychiatry and Normal Newborns diagnosis related groups.
[28] The IPPS operating payment excludes disproportionate share hospital, indirect medical education, and outlier payments.
[29] The IPPS discount freezing came from NJHA discussions with CMS. Rationale underlying these discussions is discussed in Section III.A.

are hospital-driven changes that require physician and non-physician clinical staff participation. Gainsharing, discussed later, is the incentive mechanism that aims to align physician practice behavior with an Awardee's care redesigns to achieve BPCI Model 1 goals.

The following are care redesign examples under this model:

- *Efficient use of Emergency Department (ED)*. Enhancing process efficiency to resolve or prevent bottlenecks of patients waiting to be admitted.
- *Discharge planning.* Engaging social workers or outpatient therapy earlier for beneficiaries in need of post-discharge assistance.
- *Computerized physician/provider order entry.* Requiring physicians to engage in electronic entry of patient treatment instructions to potentially decrease errors in patient care orders (e.g., related to instruction transcription from handwritten instructions) and increase hospital efficiency (e.g., billing).
- *Redesign of clinical practice*. Ensuring implementation of clinical "best practices."

I.B.1.4. Quality Monitoring Requirements

CMS requires BPCI Model 1 Awardees to report on all Hospital Inpatient Quality Reporting (IQR) program measures and other measures agreed upon between CMS and Awardees. In addition, Awardee hospitals may adhere to additional requirements. New Jersey Awardees, for example, agreed to:[30]

- Comply with accreditation standards from The Joint Commission.
- Submit data related to the Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).
- Review nationally recognized quality indicators such as mortality, readmissions, and length of stay.
- Condition receipt of gainsharing payments to enrolled practitioners based on achievement of meeting agreed upon quality measures; these conditions may vary across Awardees.

Awardees were originally required to implement the B-CARE Tool, a streamlined version of the Continuity Assessment Record and Evaluation (CARE) tool. This requirement was dropped from BPCI Model 1 in PQ 4 after feasibility discussions between CMS and model Awardees regarding the resource and time burden to meet the requirement.

I.B.1.5. Gainsharing and Incentive Payments

Guidelines for gainsharing to physicians and non-physician practitioners providing Medicare services are broad; specific calculations are proposed by BPCI Model 1 Awardees in their Implementation Protocol (IP) and accepted by CMS. The following are restrictions listed in the BPCI Model 1 Awardee Agreement:

[30] The following are requirements of the Facilitator Convener, New Jersey Hospital Association, as part of its implementation protocol. These requirements are not those of CMMI or BPCI Model 1.

- Total incentive payments to an individual Enrolled Practitioner must be limited to 50 percent[31] of the aggregate Medicare payment amount determined under the PFS paid to the Enrolled Practitioner for services furnished to Medicare beneficiaries at BPCI Model 1 Awardee beneficiaries in the Performance Year.
- An incentive payment to an Enrolled Practitioner or BPCI Entity must be derived solely from internal hospital cost savings the Awardee generated during a time period when the Enrolled Practitioner or BPCI Entity has been deemed eligible by CMS to receive Incentive Payments.
- Model 1 Awardee is responsible for monitoring and enforcing payment restrictions.

Each Awardee maintains a Steering Committee that develops its own conditions for receiving gainsharing payments, but employs consistent methodologies to calculate that amount using cost-to-charge ratios and medical claims data. CMS requires that Awardees outline their Care Redesign, Management and Staffing, Gainsharing, and Beneficiary Notification in their Implementation Protocols, which are submitted prior to executing an Awardee Agreement with CMS and must be updated prior to the period of performance if the Awardee would like to change their Protocol.

I.B.2. BPCI Model 1 Awardee Hospitals

The BPCI Model 1 performance period began on April 1, 2013, with 23 Awardee hospitals (all located in New Jersey) under 1 FC, the New Jersey Hospital Association (NJHA). Of the 23 Awardees, 6 participated in the PHC, a gainsharing demonstration similar to BPCI Model 1 that ended in July 2012. Those same six Awardees opted to participate in an extension of PHC from July 2012 to March 2013 that led up to the BPCI Model 1 performance period.

On January 1, 2014, 1 Awardee hospital located in Kansas initiated an Awardee Agreement with CMS for BPCI Model 1 for a maximum of 24 Awardees; this hospital does not have an FC. As of July 1, 2014, 9 of the 24 hospital Awardees terminated their Awardee Agreement[32] and are no longer active in BPCI Model 1. This report considers all 24 BPCI Model 1 Awardees and groupings of these Awardees.

The New Jersey FC, NJHA, works with Applied Medical Services (AMS) to provide New Jersey Awardees with internal hospital cost savings and select measure data (e.g., length of stay).[33] These data track two domains on physician performance: (1) prior year's performance compared with current year's performance and (2) performance compared with peers (statewide). Administrators may receive dashboard-like reports on individual enrolled physicians, which include their utilization patterns of medical supplies and hospital cost centers (e.g., emergency rooms, labs, radiology), and compare these costs with prior incurred costs and with a "Best

[31] This is an increase from 25 percent under prior demonstrations.
[32] Further, as of March 1, 2015, 12 of the 24 Awardee hospitals terminated their Awardee Agreement.
[33] More information is available online at http://www.appliedmedicalsoftware.com/for-hospitals-health-systems/.

Practice Norms" as a benchmark.[34] Eleven of the Awardees increase this data specificity with additional data tools/platforms.[35] Section III.A notes limitations of these tools.

Table 2 identifies BPCI Model 1 Awardees and notes the quarter that exiting Awardees terminated their Awardee Agreement with CMS. Table 2 also identifies those Awardees that previously participated in the PHC demonstration. Note that all PHC Awardee hospitals terminated their Awardee Agreement with Model 1. Table 2 also exhibits select Awardee hospital characteristics. It shows that the majority of Awardees have non-employed (i.e., voluntary) physician staff and that average bed size for an Awardee hospital is 300 beds, with the smallest hospital having 24 beds and the largest having more than 650 beds. The average percentage of Medicare admissions for these Awardees is approximately 45 percent, indicating a fair share of revenue may be subjected to the IPPS discount in this model.

[34] Best practice norms are determined using local (statewide) practice data and minimum case volume standards. These data adjust for physician specialty and patient severity of illness to ensure comparable norms and rates.
[35] Crimson, owned by The Advisory Board Company, is a provider of data, analytics, and business intelligence software to hospitals, health systems, and physicians.

Table 2: BPCI Model 1 Hospitals' Enrollment, Termination, and Select Characteristics*

Hospital Name (Abbreviated Name)	Active in BPCI Model 1	Exiting Quarter	Prior PHC Participation	Physician Staff Mix	Beds	Hospital Size	Percent of Medicare Admissions
Capital Health Medical Center – Regional ("Capital Health")	✓			66% Voluntary; 33% Employed	176	Medium	46.25
Capital Health Medical Center – Hopewell ("Capital Health Hopewell")	✓			66% Voluntary; 33% Employed	211	Medium	30.10
CentraState Medical Center ("CentraState")		PQ 3	✓	95% Voluntary; 5% Employed	245	Medium	48.73
Cooper Hospital/University Medical Center ("Cooper")		PQ 3		95% Voluntary; 5% Employed	488	Medium	26.54
Deborah Heart and Lung ("Deborah")		PQ 5		100% Employed	89	Small	55.98
Hunterdon Medical Center ("Hunterdon")		PQ 5	✓	Not Specified	170	Medium	47.13
Jersey Shore University Medical Center ("Jersey Shore")		PQ 3	✓	50% Voluntary; 50% Employed	513	Large	44.22
JFK Medical Center ("JFK")		PQ 5	✓	Mostly Voluntary; Some Employed	351	Medium	46.27
Morristown Medical Center ("Morristown")		PQ 5		80% Voluntary; 20% Employed	531	Large	41.14
Overlook Medical Center ("Overlook")		PQ 5	✓	Mostly Voluntary	423	Large	48.91

Hospital Name (Abbreviated Name)	Active in BPCI Model 1	Exiting Quarter	Prior PHC Participation	Physician Staff Mix	Beds	Hospital Size	Percent of Medicare Admissions
Robert Wood Johnson University Hospital ("RWJU")	✓			50% Voluntary; 50% Employed (Faculty)	610	Large	38.34
Robert Wood Johnson University Hospital – Hamilton ("RWJ Hamilton")	✓			Mostly Voluntary	234	Medium	49.47
Robert Wood Johnson University Hospital – Rahway ("RWJ Rahway")	✓			100% Voluntary	141	Medium	61.43
St. Joseph's Regional Medical Center ("St. Joseph's")	✓			75% Voluntary; 25% Employed	670	Large	34.08
Saint Clare's Hospital ("Saint Clare's")	✓			Mostly Voluntary	337	Medium	39.55
Saint Michael's Medical Center ("Saint Michael's")	✓			90% Voluntary; 10% Employed	217	Medium	45.46
Saint Peter's University Hospital ("Saint Peter's")	✓			50% Voluntary; 50% Employed	383	Medium	26.27
Inspira Medical Center – Elmer ("Inspira Elmer")	✓			70% Voluntary; 30% Employed	88	Small	55.60
Inspira Medical Center – Vineland ("Inspira Vineland")	✓			70% Voluntary; 30% Employed	284	Medium	45.09

Hospital Name (Abbreviated Name)	Active in BPCI Model 1	Exiting Quarter	Prior PHC Participation	Physician Staff Mix	Beds	Hospital Size	Percent of Medicare Admissions
Inspira Medical Center – Woodbury ("Inspira Woodbury")	✓			80% Voluntary; 20% Employed	219	Medium	52.11
St. Mary's Hospital Passaic ("St. Mary's")	✓			100% Voluntary	210	Medium	51.70
The Valley Hospital ("Valley")		PQ 5	✓	Mostly Voluntary; Some Employed	423	Medium	52.18
University Medical Center of Princeton at Plainsboro ("UMC Princeton")	✓			95% Voluntary; 5% Employed	206	Medium	43.62
Kansas Surgery and Recovery Center ("KSRC")	✓			100% Voluntary	24	ND	41.75

* Only terminations that occurred prior to the writing of this report (October 31, 2014) are noted. PQ = PQ starting from April 1, 2013.
Data source: American Hospital Association (AHA) Annual Survey 2012 and Awardee hospitals.

I.C. Report Outline

The purpose of this Annual Report is to provide an interim evaluation of BPCI Model 1 impacts and insight into their impetus through the first five PQs, from April 1, 2013, through June 30, 2014. The remainder of this report is structured as follows: Section II presents detailed methodology for the evaluation/monitoring analyses. Section III presents results and implications for domains assessed in this report, including Awardee responses to BPCI Model 1 policies, design components, and quantifiable outcomes. Section IV discusses findings, and Section V concludes with information on forthcoming measures and analyses.

Section II. Methodology

The methodology employed for analyses in this report allows for the evaluation of whether and why BPCI Model 1 goals have been met. This evaluation occurs through assessment of model design components, how they may trigger changes in Awardee and clinical staff behavior, and how these changes translate to quantifiable outcomes.

Primary data (e.g., from interviews and focus groups) enrich the assessment of BPCI Model 1. This information provides an in-depth understanding of successes, issues, and unanticipated results that may arise during model implementation and contextualizes the comparative, quantitative findings of BPCI Model 1 impacts on measures assessed in this report.

Section II.A details measures assessed in this report and their domains. Section II.B details primary data sources and collection methods for qualitative analyses in this report. Section II.C details secondary data sources, analytical construction, and analyses for quantitative impact analyses presented in this report.

II.A. Domains and Measures for Model Evaluation

The current evaluation approach combines Awardee characteristic and model feature information with that of Awardee activities (e.g., care redesigns) obtained from stakeholder interviews and focus groups. Quantifiable data (e.g., Medicare claims, clinical staff enrollment) were incorporated to assess BPCI Model 1 impacts, and their impetuses, across the following domains:[36]

- *Implementation and Organizational Responses of Awardees.* This domain includes health care organizational features such as health IT infrastructure, provider capacity, systems, and other health care infrastructure support measures. Unlike other domains, data used to generate the measures in the structure domain are collected through primary data collection tools, including surveys;[37] telephone interviews, and focus groups with Awardee stakeholders. This domain also includes insights on the number and types of care redesigns pursued across Awardees and further vary in measurement of progress, success, and even objective. Section III.A presents findings under this domain.
- *Episode Case Mix and Patient Characteristics.* This domain focuses on demographic and health characteristics of patients and/or populations that may affect quality of care or payments. Case mix includes patient severity, patient risk scores, and number of chronic conditions. Patient characteristics include patient age, race/ethnicity, and Medicare-Medicaid dual-eligibility status. The data sources for these measures are available through inpatient claims data, inpatient community risk scores, master beneficiary summary files, and patient experience data. These characteristics are also accounted for in other domains (e.g., Health Care Outcomes and Resource Utilization). Section III.B.1 presents findings under this domain.

[36] These domains and measures formed by measure and domain guidance from Rapid Cycle Evaluation Innovation Center. *CMMI Core Measures Version 9*. Baltimore, MD: CMS, CMMI. March 28, 2013.
[37] Excluded from this report due to lack of follow-up data, which are expected within the next PQ.

- *Care Process.* This domain will include measures related to the process used by the Awardee to provide care to patients. Examples include care coordination and adherence to evidence-based protocols. The selection of specific measures will be guided by those endorsed by the National Quality Forum (NQF) and recommended by CMMI as core measures for evaluation (e.g., IQR measures). The Patient Health and Experience Survey (PHES), developed under this monitoring and evaluation contract, will also incorporate care coordination data, collected from samples of beneficiaries who receive care at Awardee hospitals.[38] This domain is excluded from this report, as data are not yet available. See Section IV for information on when data for this domain are expected.
- *Health Care Outcomes and Resource Utilization.* This domain includes measures of the quantity of services provided, whether or not a particular service was delivered during a specified time interval, and outcome measures. This includes pre- and post-discharge utilization and readmission measures. Data sources for these measures may come from IQR and claims data. The domain may also include mortality measures, patient safety measures, and functional status and health status change measures. The data sources for these measures include inpatient claims and other Medicare administrative data, IQR data, PHES data, and post-acute assessment data files. In this report, measures included in this domain currently come from Medicare administrative and claims data. Section III.B.2 presents findings under this domain.
- *Medicare Payments and Internal Costs.* This domain includes both payments made by Medicare for services provided and costs incurred by providers during care delivery. The data sources for Medicare payments are Medicare claims data files. Data on internal cost savings are obtained from Awardees. Medicare payment measures were considered in this report; however, internal cost measures require additional quarters of data before they can be made available. Section III.B.3 presents findings under this domain.
- *Patient Care Experience.* This domain examines patients' experience/satisfaction with the care they received during their hospital admission. More specifically, this domain includes communication/shared decision-making measures, pain management measures, and satisfaction with care measures. Measures for this domain will include HCAHPS survey measures, such as patients' perception of physician communication and hospital cleanliness. This domain is excluded in this report, as data are not yet available.

Table 3 lists quantitative measure specifications (also provided in Appendix A) for Health Care Outcomes and Resource Utilization and Medicare Payment domains.

[38] The PHES is a beneficiary survey tool developed specifically for BPCI Model 1 evaluation.

Table 3: Domains and Measures

Measure	Definition/Description	Technical Definition	Eligible Sample	Measurement Period
Episode Case Mix and Patient Characteristics				
Episode MS-DRG weight	Average weight of MS-DRGs of episode. Presented at cohort levels for Awardees and comparison hospitals.	MS-DRG weights episodes divided by the number of episodes.		NA
Patient age	Average age of beneficiaries across episodes. Presented at cohort levels for Awardees and comparison hospitals in years.	The average beneficiary age across all episodes calculated as the sum of beneficiary ages for all episodes divided by the number of episodes. Beneficiary age comes from the Master Beneficiary Summary File.		NA
CMS Hierarchical Condition Category (CMS-HCC) score	Average CMS-HCC risk score of beneficiaries across episodes. Presented at cohort levels for Awardees and comparison hospitals.	The sum of CMS-HCC risk scores across episodes divided by the total number of episodes.		NA
Health Care Outcomes and Resource Utilization				
30-day all-cause mortality	Mortality rate for beneficiaries within 30 days of episode admission. Presented at cohort levels for Awardees and comparison hospitals as percentages or odds ratios.	The number of Medicare beneficiaries who die within 30 days from episode admission date divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	30 days from episode admission.
30-day condition-specific mortality (Acute myocardial infarction (AMI), Pneumonia (PN), Heart failure (HF))	Condition-specific mortality rate for beneficiaries within 30 days of episode admission. Presented at cohort levels for Awardees and comparison hospitals as percentages or odds ratios.	The number of Medicare beneficiaries who die within 30 days from AMI-, PN-, or HF-related episode admission date divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	30 days from episode admission.

Measure	Definition/Description	Technical Definition	Eligible Sample	Measurement Period
30-day all-cause readmissions	Readmission rate for beneficiaries within 30 days after episode discharge. Presented at cohort levels for Awardees and comparison hospitals as percentages or odds ratios.	The number of Medicare beneficiaries who are admitted to acute-care inpatient or critical access care hospitals within 30 days after episode discharge divided by the number of episodes. These admissions are attributed to the closest episode instance.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	30 days after episode discharge.
30-day condition-specific readmissions (AMI, PN, HF)	Condition-specific readmission rate for beneficiaries within 30 days after episode discharge. Presented at cohort levels for Awardees and comparison hospitals as percentages or odds ratios.	The number of Medicare beneficiaries who are admitted to acute-care inpatient or critical access care hospitals within 30 days after AMI-, PN-, or HF-related episode discharge divided by the number of episodes. These admissions are attributed to the closest episode instance.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	30 days after episode discharge.
ICU use during episode	Average beneficiary ICU use/stay during an episode. Presented at cohort levels for Awardees and comparison hospitals as percentages or odds ratios.	The number of episodes with revenue center codes 020X (except 0206) divided by the number of episodes. Multiple ICU stays during an episode are counted as a singular ICU stay for that episode.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	Episode.
Length of inpatient episode	Average length of stay of an episode. Presented at cohort levels for Awardees and comparison hospitals in days.	The sum of episode length of stay (= Claim Through Date – Claim From Date + 1) for all episodes divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	Episode.

Measure	Definition/Description	Technical Definition	Eligible Sample	Measurement Period
Medicare Payments				
30-day Medicare payments post-episode	Average Medicare payments to providers/suppliers up to 30 days after episode discharge. Presented at cohort levels for Awardees and comparison hospitals in dollars.	The summed total Medicare payments of non-episode inpatient, carrier, outpatient, SNF, home health, hospice, and DME claims up to 30 days after episode discharge divided by the number of episodes. These post-episode payments are associated with their nearest episode discharge within a 30-day window.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	30 days after episode discharge.
Total episode payments	Average Medicare payments to providers/suppliers during the episode period, including payment for episode, physician services, and durable medical equipment. Presented at cohort levels for Awardees and comparison hospitals in dollars.	The summed total Medicare payment for inpatient, carrier, outpatient, and DME claims during the episode period. This is effectively the sum of hospital and non-hospital Medicare payments during the episode divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	Episode.
Non-hospital episode payments	Average Medicare payments to providers/suppliers during the episode period excluding payments for episodes. Presented at cohort levels for Awardees and comparison hospitals in dollars.	The summed total Medicare payment for carrier, outpatient, and DME claims during the episode period. This is effectively the sum of and non-hospital Medicare payments during the episode period divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	Episode.
Hospital-only episode payments	Average Medicare payments to Awardee or comparison hospitals during the episode period. Presented at cohort levels for Awardees and comparison hospitals in dollars.	The summed total Medicare payment for episodes divided by the number of episodes.	(1) Beneficiaries that maintained FFS A & B enrollment without HMO enrollment during the measurement period or 30 days prior to episode. (2) Episode lengths of stay not exceeding 1 year. (3) Beneficiaries that did not have ESRD entitlement. (4) Episodes with Medicare as a primary payer. (5) Episodes that are at the beginning or middle of a transfer sequence to/from another facility.	Episode.

The next sections review the data and methods employed in this report.

II.B. Primary Data Collection Methods and Analyses

Primary data collected provides valuable information of the status of the model at each Awardee hospital. This qualitative information includes Awardee characteristics, staffing structure, patient population, organizational culture, and competing initiatives. Telephone interviews and focus groups were conducted to gather the desired information. Telephone interviews were conducted with Awardee hospital administrators prior to in-person focus groups with hospital staff ("site visits") to obtain multiple stakeholder perspectives. The site visits added validity to the data collected during telephone interviews and provided a well-rounded picture of model experience at Awardee hospitals.

II.B.1. Primary Data Collection Methods

Table 4 presents an overview of primary data collection methods and frequency. Table 5 details stakeholder participation in telephone interviews and site visits.

Table 4: Telephone Interview and Site Visit Characteristics and Participation

Qualitative Data Collection Methods			
Characteristic	**Telephone Interview**	**Site Visit (Focus Group)[39]**	**Exit Interview**
Frequency	Semiannually	Average of 4 per year	As needed
Instrument	Semi-structured	Semi-structured guides	Semi-structured
Duration	45 minutes	60 to 90 minutes per session	45 minutes
Content	Physician Engagement, Progress on Care Redesigns, Physician Incentive Mechanisms	Physician Engagement, Progress on Care Redesigns, Physician Incentive Mechanisms	Physician Engagement, Progress on Care Redesigns, Physician Incentive Mechanisms
Stakeholder Participation	Awardee Hospital Administrators[40] (1–3)	Awardee Hospital Administrators (5–7) Physicians (4–6) Non-physician clinical[41] (7–10)	Awardee Hospital Administrators (1–3)

Table 5: Telephone Interview and Site Visit Participation

Telephone Interview and Site Visit Participation							
Awardee Name	**Telephone Interview**		**Site Visit (Focus Group)[42]**				**Exit Interview**
	Wave 1	Wave 2	July 2013	October 2013	March 2014	October 2014	As Needed
Capital Health and Capital Health Hopewell	✓	✓					

[39] In situations where three or fewer participants attend a session, the moderator switched to an in-person interview format.
[40] Awardee Hospital administrators include BPCI Model 1 Program Coordinators (PCs) and/or CMOs.
[41] Non-physician clinical staff includes registered nurses, care coordinators, patient navigators, and discharge planners.
[42] In situations where three or fewer participants attend a session, the moderator switched to an in-person interview format.

Telephone Interview and Site Visit Participation							
Awardee Name	**Telephone Interview**			**Site Visit (Focus Group)[42]**			**Exit Interview**
	Wave 1	Wave 2	July 2013	October 2013	March 2014	October 2014	As Needed
CentraState	✓		✓ (Pilot)				✓
Cooper	✓						✓
Deborah	✓						✓
Hunterdon	✓						✓
Jersey Shore	✓						✓
JFK	✓			✓			✓
Morristown	✓						✓
Overlook	✓				✓		✓
RWJU	✓	✓		✓			
RWJ Hamilton	✓	✓		✓			
RWJ Rahway	✓	✓				✓	
St. Joseph's	✓	✓		✓			
Saint Clare's	✓	✓					
Saint Michael's	✓	✓			✓		✓
Saint Peter's	✓	✓				✓	
Inspira Elmer	✓	✓					
Inspira Vineland	✓	✓					
Inspira Woodbury	✓	✓					
St. Mary's	✓	✓					
The Valley	✓						
UMC Plainsboro	✓	✓					
KSRC		✓					

II.B.1.1. Telephone Interviews

This report includes data collected and analyzed from telephone interviews over two waves. Twenty-three Awardee hospitals were interviewed during the first wave, which occurred from April 1, 2013, to December 23, 2013. Fifteen Awardees were interviewed during the second wave, from January 3, 2014, to May 31, 2014.[43]

Telephone interviews were conducted with Awardee hospital administrators semiannually to obtain information on implementation and organizational activities surrounding BPCI Model 1. Awardee hospital administrators include Program Coordinators (PCs), who coordinate Model 1 activities at a hospital, and Chief Medical Officers (CMOs). The purpose of the semiannual telephone interviews is to obtain a comprehensive account of model activities such as the implementation of care redesigns, methods to increase physician enrollment and engagement, process for and effects of the physician incentive mechanisms, and successes and challenges encountered.

[43] Nine Awardees terminated their Awardee Agreement with CMS prior to or during the second data collection wave. KSRC entered the BPCI Model 1 demonstration on January 1, 2014.

Hospital administrators were contacted via email 2 weeks prior to the desired period with potential availabilities. Hospital administrators (e.g., CMOs and BPCI Model 1 PCs) who possess an in-depth knowledge of both the hospital and the model were targeted and could respond to model activities. Telephone interviews were scheduled for approximately 45 minutes, and follow-ups were conducted if all relevant respondents could not attend. Two follow-ups were scheduled during this data collection period, one from each wave. Follow-up interviews with a hospital were included in the wave during which they were conducted. The telephone interview data collection intentionally preceded the site visits in order to inform the focus group guides.

In order to assess the overall progress of the model, semi-structured guides included questions across three key domains: physician engagement, implementation of care redesigns, and physician incentive mechanisms and perspectives. These guides were developed from Awardee-submitted IPs that were submitted to CMS for acceptance prior to the model start date. The IPs detail hospital mechanisms for implementing Model 1, including responsibilities of hospital personnel and care redesign implementation. The guides were piloted at four hospitals (Jersey Shore, Morristown, Inspira Elmer, and Inspira Woodbury) and adjusted based on data collected from the telephone interviews conducted during the first wave. These guides were refined over time, based on new information gathered from Awardees regarding care redesign updates and/or evolved foci of the program from CMS (e.g., physician gainsharing, model requirement changes).

II.B.1.2. Site Visits

This report includes data collected during nine site visits over four points in time: July 2013, one hospital; October 2013, four hospitals; March 2014, two hospitals; and October 2014, two hospitals (see Table 5). Awardee hospitals were targeted for site visits based on unique successes and challenges reported during telephone interviews.

For site visits, a broader hospital administrator group was targeted than for telephone interviews, including Chief Financial Officers and department heads. Focus group sessions conducted during site visits were comprised of three types of hospital stakeholders:

- *Hospital Administrators.* This group contained hospital administrators involved in managing or implementing the BPCI Model 1 initiative, such as the BPCI Model 1 PCs and members of the BPCI Model 1 Steering Committee.[44]
- *BPCI Model 1 Enrolled Practitioners.* This group included physicians currently enrolled in BPCI Model 1 and potentially eligible for gainsharing payments who did not hold an administrative position. Physician respondents would be able to provide insight on their understanding of the model, its goals, and their perspective on model activities and features such as gainsharing.
- *Care Redesign Team Leaders.* This included non-physician clinical staff members involved in care redesign activities[45] (e.g., registered nurses, care coordinators, patient

[44] Steering Committees in this model are primarily responsible for overseeing quality initiative progress, physician enrollment, and gainsharing payment distributions.

navigators, discharge planners). Non-physician team leaders can respond to experiences implementing care redesigns for the model and staff engagement for these redesigns.

Semi-structured guides were tailored for each of the three focus group types.[46] Combined, these site visit guides paralleled telephone interview guides in their thematic framework; the open-question format was used to better triangulate data collected across these tools and Awardees.

Site visit and focus group participant agendas and logistics were coordinated with the PC at that hospital. The PC was responsible for recruiting BPCI Model 1 administrator, physician, and non-physician staff for each focus group.

Similar to the telephone interview process, focus group guides were piloted at CentraState in July 2013. This hospital was selected due to its previous experience with gainsharing (PHC). The pilot site visit consisted of three focus groups with three types of personnel: administrators, physicians, and non-physician clinicians. Since the pilot site visit, the number of focus groups has increased to obtain perspectives from medical physicians, surgeons, and multiple care redesign teams.

II.B.1.3. Exit Interviews

Semi-structured exit interviews were conducted with Awardee hospital administrators involved in Model 1 oversight for Awardees that terminated their Awardee Agreement with CMS. As of July 1, 2014, nine hospitals terminated their Awardee Agreement. These interviews were scheduled with BPCI Model 1 administrators within 30 days after notification from CMS that an Awardee had elected to terminate its Awardee Agreement.

The domains included for the exit interviews differ from those of the telephone interviews and site visits. These interviews solicited information about how care redesign, financial savings, engagement, and other confounding initiatives contributed to the hospital's decision to terminate. Interviewees were also probed on whether they expected to continue their care redesign and quality improvement activities that the hospital.

II.B.2. Analyses

All conversations carried with stakeholders during telephone interviews, site visits, and exit interviews were recorded and transcribed. QSR NVivo 10 qualitative data analysis software was used to organize, code, and analyze the information. A codebook was developed based on specific themes, including physician engagement, gainsharing, care redesign interventions progress, financial incentive mechanisms, and reasons for termination. Thematic coding is used to identify passages within text, and references to these evaluation domains were assigned.

Comparative analysis of information collected from the various stakeholders across Awardees was conducted to assess model implementation and organizational responses of BPCI Model 1. This analysis relied on a triangulation process that first determined similarities and differences

[45] The New Jersey Awardees currently only enroll physicians for BPCI Model 1 gainsharing.
[46] Care Redesign teams were typically composed of nurses, patient coordinators, and other non-physician clinical staff; however, some enrolled physicians were included.

within stakeholder types (e.g., comparing administrator responses from telephone interviews and focus groups for a given Awardee hospital). Then cross-stakeholder responses were compared within an Awardee hospital and to findings across Awardees to determine common themes. In looking for commonalities, differences that might reflect current or future model success (or failure) were also assessed.

II.C. Secondary Data Methods and Analyses

Data come from the 100-percent research identifiable Medicare Claims ("claims") and Enrollment Database from the CCW. Medicare claims data were used, including inpatient, outpatient, carrier, durable medical equipment (DME), skilled nursing facility (SNF), home health agency (HHA), and hospice claims. Claims from long-term care hospitals and rehabilitation hospitals were included in other inpatient claims. The Master Beneficiary Summary File AB and Chronic Condition components to extract patient demographics, program eligibility, and chronic condition flags were also used. Other data used include Awardee hospital-level data, such as clinical staff enrolled in the model or hospital characteristics from the Provider of Service file. Future secondary analyses will incorporate hospital cost-savings data.

These claims data are analyzed by quarter and as aggregate time periods of Baseline, Since BPCI Inception, Year 1, and PQ 5. Table 6 details the time period dates for the aggregate periods.[47]

Table 6: Time Period Dates for Aggregate Periods

Time Period Name	Time Period Dates
Baseline	January 1, 2011, through March 31, 2013
Since BPCI Inception	April 1, 2013, through June 30, 2014
Year 1	April 1, 2013, through March 31, 2014
PQ	Exclusive 3-month periods from April 1, 2013, through June 30, 2014

The remainder of this section discusses episode and post-episode identification for measures assessed in this report, comparison hospital selection methodology, and Awardee groupings; it concludes with a discussion of how comparison hospitals and hospital groupings are utilized within presented impact analyses.

II.C.1. Episode Construction

The time periods in Table 6 are composed of patient-level episodes: acute-care inpatient hospital stays that do not have an associated MS-DRG excluded[48] by the awardee. The episode itself captures data on services provided to a patient from admission into the hospital through their discharge. For any given measure (Table 3), episodes were included if:

1. The episode occurred at Awardee hospitals or comparison hospitals.
2. The episode occurred from January 1, 2011, through June 30, 2014.

[47] KSRC's Since BPCI Inception period begins on January 1, 2014.
[48] For example: NJHA Awardees all exclude Psychiatry and Normal Newborns diagnosis related groups.

3. The episode was not excluded from analysis by numerator or denominator defined exclusions listed in Appendix A for that measure.

Some measure statistics rely solely on services that occur within patient episodes. These measures can be identified by their measure name including "in episode" or "over/during episode" in Table 3. As an example, a patient's length of stay or whether they had an ICU stay is identified by examining time spent at the hospital for that stay or whether intensive care resources were utilized during that patient's episode.

Other measures require accounting for outcomes (e.g., mortality), utilization (e.g., post-acute care centers), or Medicare payments (e.g., doctor visits) over specific time frames *after* a patient's episode. These measures, identified as "post-episode" measures in Table 3, focus on outcomes or services utilized within 30-day timeframes after a patient's episode. Note that in some instances, two or more episodes (i.e., rehospitalizations) may occur within a 30-day timeframe. In these instances, outcomes or services utilized are attributed to their most recent preceding episode. In other words, outcomes and services were attributed to the most recent episode to account for post-discharge events/services for all episodes assessed under Model 1.

Note that episodes within the last months of this report's analysis period (ending on June 30, 2014) may utilize claims 30 days beyond June 30, 2014, for post-episode measures (e.g., readmissions). Medicare claims data used for report analyses were pulled from CCW servers on November 14, 2014, to allow for a minimum of 4 months of maturation for claims data pulled through June 2014. Medicare data are expected to be 80 to 96 percent complete for a 3- to 6-month run-out period.

II.C.2. Comparison Hospital Methodology

Quantitative BPCI Model 1 impacts on measures in this report come from comparative analyses of Model 1 Awardees to *similar* non-Awardee hospitals ("comparison hospitals" or "comparison group"). By construction, comparison hospitals allow for inferences of a counterfactual scenario: What would Model 1 Awardee measure performance have been had they not participated in BPCI Model 1?

Construction of the comparison hospital group was a multistep process that included identifying a pool of non-Awardee hospitals (hospitals that did not participate in BPCI Model 1) and then determining non-Awardee hospital similarity to Awardee hospitals. In this construction, similarity was determined by statistical examination of observable characteristics between each Awardee hospital and all non-Awardee hospitals. Observable characteristics were examined over the Baseline period of this report (January 1, 2011, through March 31, 2013) and included the number of hospital beds, patient case mix, and Baseline values of select measures (e.g., average episode length). The resulting comparison hospital group was composed of non-Awardee hospitals deemed statistically similar in their Baseline period to Awardees across these characteristics. Multiple statistically similar non-Awardee hospitals were *matched* to each Awardee to minimize anomalous performance of any singular non-Awardee hospital over time.

Thus, BPCI Model 1 impacts were ascertained by comparing differences in performance measures between Baseline and model implementation periods (April 1, 2013, through June 30, 2014) between two different hospital groups:

- *BPCI Model 1 Awardee hospitals.* This group includes all Awardee hospitals (i.e., those that have executed BPCI Model 1 Awardee Agreements with CMS).
- *BPCI Model 1 comparison hospitals.* This group includes non-Awardee hospitals that did not participate in BPCI Model 1 but are otherwise comparable to the Model 1 Awardee hospitals in terms of observable and measurable characteristics. Outcomes for hospitals in this comparison group may be inferred as counterfactual scenarios for a given performance measure and allow for inferences as to what Model 1 Awardee outcomes would have been had they not participated in Model 1.

The remainder of Section II.C.2 provides an overview of issues that drove comparison hospital selection methods, rationale for criteria used to define the comparability of a non-Awardee hospitals, and comparison hospital selection results. Appendix C provides more technical detail on methods used.

II.C.2.1. Selection of Comparison Hospitals

The selection of comparison hospitals requires a similarity assessment of each Awardee hospital to all non-Awardee acute-care inpatient hospitals. These non-Awardee hospitals comprised a pool from which comparison hospitals could be selected. Non-Awardee hospitals deemed statistically similar to a given Awardee hospital, within a predefined degree of statistical similarity, may be *matched* to that Awardee hospital and included in the comparison hospital group. There are a variety of methods for such assessments of statistical similarity ("matching"); however, a common component of these methods is that they require a set of criteria (e.g., observable characteristics) over which similarity is assessed.

Subsections C.2.2 and C.2.3 detail the set criteria over which similarity between Awardee and potential comparison hospitals was assessed, the statistical metrics for assessment, and results of the comparison hospital selection process.

II.C.2.2. Criteria for Similarity Assessment

Table 7 describes the characteristics utilized in the matching process and their sources.[49] For these characteristics, data before April 1, 2013, were utilized to avoid characteristics whose measurement was potentially affected by BPCI Model 1 Awardees.

First, 2 hospital characteristics were used to winnow down the pool of over 3,000 potential comparison hospital matches to 1,851 hospitals. These "Pre-Match Characteristics" (Table 7) identified whether or not potential comparison hospitals shared two characteristics that all BPCI Model 1 Awardees shared: being located in an urban area and receiving IPPS payments.

[49] Several other matching specifications were considered (e.g., using 1:1, 2:1, and 3:1 NN matching). Other covariates were also considered, such as change in readmission and mortality rates, case mix index, average MS-DRG weight, disproportionate share percentage, and proportion of patients who are dually eligible.

Table 7: Variable Names, Specifications, and Sources

Variable Name	Technical Specification	Source
Pre-Match Characteristics		
Urban	Equal to 1 if the hospital is located in an urban area; 0 otherwise	FY 2014 Final Rule
Provider type	Nine types of acute care hospitals: – IPPS – Rural Referral Center (RRC) – Indian Health Service – Medicare Dependent Hospital (MDH) – MDH/RRC – Sole Community Hospital (SCH) – SCH/RRC – Essential Access Community Hospital (EACH) – EACH/RRC	FY 2014 Final Rule
Hospital Characteristics		
Indicator for general hospital	Equal to 1 for general hospital; 0 otherwise	AHA Annual Survey Database FY 2011
Indicator for teaching hospital	Equal to 1 for teaching hospital; 0 otherwise	AHA Annual Survey Database FY 2011
Indicator for presence of ED	Equal to 1 if the hospital has an ED; 0 otherwise	AHA Annual Survey Database FY 2011
Number of beds	Number of hospital beds	FY 2014 Final Rule
Urban type	Equal to 1 if hospital is located in a large urban area; equal to 2 if hospital is located in another type of urban area	FY 2014 Final Rule
Patient Episode Characteristics		
Indicator for high surgical MS-DRG percentage	Equal to 1 if hospital has more than 90 percent surgical MS-DRGs	2008 Inpatient 5% Base Claims Database
Percent of hospital days paid by Medicare	Medicare days as a proportion of total	FY 2014 Final Rule
Average HCC score	Calculated from inpatient claims and HCC files.	Inpatient claims HCC files
Length of stay	Calculated from inpatient claims file	Inpatient claims
Change in length of stay between: – Q1 2011 and Q2 2011 – Q3 2011 and Q4 2011 – Q1 2012 and Q2 2012 – Q3 2012 and Q4 2012	Calculated from inpatient claims file	Inpatient claims
All-cause mortality	Calculated from inpatient claims and Beneficiary Summary File	Inpatient claims Beneficiary Summary File
All-cause readmissions	Calculated from inpatient claims file	Inpatient claims

Second, with the reduced universe of potential comparison hospitals, a Mahalanobis distance metric with a *nearest neighbor* (NN) matching algorithm was employed. The Mahalanobis distance metric computes a *distance* between characteristics of Awardee hospitals and potential comparison hospitals. The NN algorithm then compares these distances for a given Awardee hospital to each potential comparison hospital. In the simplest NN algorithm, only the closest potential comparison hospital is matched to an Awardee hospital for inclusion into the

comparison hospital group. This requirement was relaxed such that the four closest potential comparison hospitals were matched to each Awardee hospital. The rationale for relaxing this requirement was so that the resulting comparison hospital group would not be sensitive to issues of a singular comparison hospital that were not otherwise captured by the aforementioned matching criteria. Further, a comparison hospital could be matched to multiple Awardee hospitals.[50]

In addition to Pre-Match Characteristics, Table 7 also lists the characteristics used in the Mahalanobis distance metric with NN matching algorithm. Following the advice of matching methodologists (Rubin, 2001,[51] 1980,[52] and 1973;[53] and Stuart, 2010[54]), selection of these criteria was based on scientific understanding of different confounding factors that could drive changes in performance measures and program participation. Namely, the following categories of characteristics were considered: hospital characteristics and patient and episode characteristics—averaged at the hospital level. Hospital characteristics included indicators for whether a hospital was a general hospital or teaching hospital and whether it had an ED. Other characteristics, including bed size and whether a hospital was located in a large urban or other urban area, were also used for matching.

These characteristics capture aspects that may fundamentally affect how hospitals function. For example, general and specialty hospitals are likely to be affected differently by new payment models because specialty hospitals tend to serve more complex patients. Similarly, teaching hospitals operate differently from non-teaching hospitals. In particular, hospitals with a graduate medical education program receive additional indirect medical education payments "to reflect the higher patient care costs of teaching hospitals relative to non-teaching hospitals."[55] Characteristics that proxy a hospital's capacity were also included—presence of an ED and number of hospital beds—as there may be patient differentials in hospitals with differing capacities.

This matching process also differentiated between hospitals located in large urban areas and all other urban areas. This was a relevant characteristic in this Model 1 evaluation as all but one Awardee hospitals are located in New Jersey. Unlike most States, New Jersey is predominantly urban with very few rural areas. However, New Jersey does present multiple suburban areas whose hospitals are likely to behave differently than those in large urban centers. In an approach that merely distinguished between urban and rural settings ("Pre-Match Exclusions"), these

[50] This *matching with replacement* allows for decreased differences between comparison and Awardee hospitals as, in the event where the *next best* match is not statistically similar to an Awardee, a previously matched comparison hospital that is statistically similar can be rematched. Appendix C provides additional information.

[51] Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services & Outcomes Research Methodology,* 2, 169–188.

[52] Rubin, D. (1980). Bias reduction using Mahalanobis-Metric matching. *Biometrics,* 36, 293–298.Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science,* 25, 1–21.

[53] Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics,* 29, 159–84.

[54] Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science,* 25, 1–21.

[55] Source: http://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Indirect-Medical-Education-IME.html.

suburban areas would be categorized together with large urban centers. Urban area size may also be associated with quality outcomes such as readmissions and care experience.

A number of patient and episode characteristics were included in the matching algorithm, such as an indicator for hospitals with a high (greater than 90 percent) proportion of episodes falling within surgical MS-DRGs. This particular indicator facilitated matching any specialty surgical hospitals with a very small proportion of non-surgical inpatient stays. This distinction was important due to differences in services provided to medical patients, relative to surgical patients. Further, medical and surgical patients may have different outcomes and types of complications and may rely on different provider processes and payment streams.

The percent of hospital days paid by Medicare captures information that could affect hospital participation decisions. BPCI Model 1 is based on inpatient stays paid by Medicare. The larger the proportion of hospital days paid by Medicare, the larger the potential incentive under Model 1 (e.g., from the IPPS episode discount). In addition, matching on proportion of Medicare days ensures that treatment and comparison hospitals face similar incentives such as those provided by the IQR program, which reduces the annual payment update for Medicare hospitals that fail to report certain quality measures.

The average CMS-HCC score for patients with an inpatient stay at a hospital summarizes the clinical complexity of the population served by the hospital. Chronic condition burden, which the CMS-HCC score captures, is correlated with patient outcomes[56] and expenditures. Patient acuity could also be linked to ability and/or appetite to take on additional risk by enrolling in the BPCI initiative. For example, hospitals serving patients with higher CMS-HCC scores may have more potential savings opportunities.

Finally, select baseline performance measures in the matching process were included, such as episode length of stay and hospital-level mortality and readmission rates. While CMS-HCC scores were meant to capture patients' expected chronic disease burden, one can infer, to some extent, average resources provided to patients across episodes, based on episode lengths (patient lengths of stay). Furthermore, length of stay is an important channel through which hospitals may reduce internal costs by reducing a patient's length of stay. Indeed, it is touted as an objective for some care redesigns pursued by Awardee hospitals (e.g., by improving patient flow through the hospital system). In addition to matching on the average length of stay at a hospital, matches using changes in length of stay over four intervals were attempted: Quarter 1 of 2011 through Quarter 2 of 2011, Quarter 3 of 2011 through Quarter 4 of 2011, Quarter 1 of 2012 through Quarter 2 of 2012, and Quarter 3 of 2012 through Quarter 4 of 2012. Matching on these *changes* in this characteristic helped ensure that assumptions necessary for DiD analyses (parallel paths assumption, discussed later) were met. Readmission and mortality were also matched on due to their importance as indicators of baseline quality-of-care and coordination-of-

[56] Li, P., Kim, M. M., & Doshi, J. A. (2010). Comparison of the performance of the CMS Hierarchical Condition Category (CMS-HCC) risk adjuster with the Charlson and Elixhauser comorbidity measures in predicting mortality. *BMC Health Services Research,* 10: 245. Accessed November 20, 2014, from http://www.biomedcentral.com/1472-6963/10/245.

care performance. In addition, these variables are strongly linked to patient demographic and community factors not necessarily observable in available data.

II.C.2.3. Statistical Measurement of Similarity and Result

The aforementioned matching methodology resulted in four comparison hospitals for every Awardee hospital, where a matched comparison hospital could be matched to more than one Awardee. Statistical similarity ("quality") of this match was assessed by examining the *bias* across each and over all Table 7 characteristics before and after the matching process.

Statistically, bias is defined using the formula in Rosenbaum and Rubin (1985):

$$\text{Bias} = \frac{\bar{X}_T - \bar{X}_C}{\left(\frac{(\sigma_T^2 + \sigma_C^2)}{2}\right)^{1/2}}, \quad (1)$$

where $\bar{X}_T$ and $\bar{X}_C$ represent the sample means of Awardee and selected comparison hospitals, respectively, for Table 7 characteristics; σ_T^2 and σ_C^2 represent the variances of these characteristics for Awardee and comparison hospitals.[57]

A high-quality match (i.e., low bias statistics that imply similarity over characteristics) would result in small standardized differences between BPCI Model 1 Awardee hospitals and selected comparison hospitals across observable characteristics. The statistical threshold for a high-quality match varies within related literature[58] and was set to 10 percent for this process. Table 8 shows overall bias *before* matching—the bias between Awardee hospitals (in aggregate) and the *potential* comparison hospital group (1,851 hospitals)—and the overall bias after matching, which compares the Awardee hospital group to *actual,* matched, comparison hospitals. The overall bias is the average percentage of biases across individual characteristics matched on. The overall bias before matching of 30.1 percent is reduced to 9.3 percent after matching. DiD regression models will account for a portion of this remaining bias.

Table 8: Overall Matching Quality

Mean Bias Before Matching	Mean Bias After Matching
30.1%	9.3%

Table 9 shows the number of BPCI Model 1 Awardee hospitals, potential comparison hospitals, matched comparison hospitals, and uniquely matched comparison hospitals.

[57] The ratios of these variances are also assessed when determining the quality of the match.
[58] Rosenbaum and Rubin (1985) suggest a 10-percent (in absolute value) threshold for the standardized difference after matching; Stuart (2010) suggests a 5-percent threshold.

Table 9: Number of Awardee and Comparison Hospitals

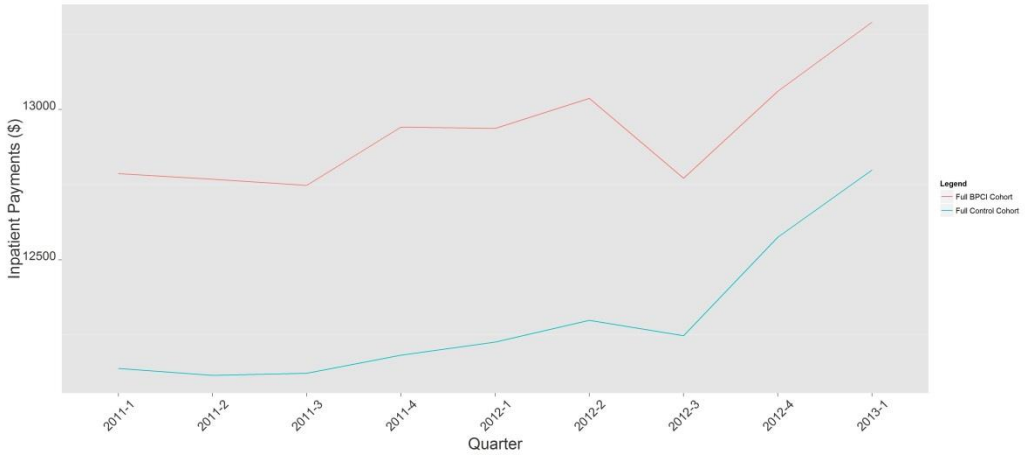
Number of BPCI Model 1 Awardee Hospitals	Number of Potential Comparison Hospitals	Number of Matched Comparison Hospitals	Number of Uniquely Matched Comparison Hospitals
24	1,851	96	82

Appendix C provides additional detail on the similarity among comparison hospitals used in this report and the Mahalanobis and NN matching methodology. However, an additional point of concern in this comparison hospital selection process was meeting an expected assumption for DiD models to produce unbiased impact estimates. This assumption, known as *parallel paths,* requires that trajectories of the dependent/outcome variables in a DiD model must not differ for two cohorts before the introduction of a study *intervention*. The translation of this assumption for DiD impact analyses in this report is that performance measure trends for Awardee and comparison hospital groups be parallel[59] over Baseline (i.e., the time period *preceding* BPCI Model 1 (the "intervention") implementation periods). The validity of this assumption is inspected visually.

Figures 2 through 5 exhibit pre-model trends of four performance measures for BPCI Model 1 Awardees and comparison hospitals; namely, average patient length of stay, 30-day all-cause mortality, 30-day all-cause readmissions, and total Medicare payments over episodes. The horizontal axis indicates the quarters prior to the start of the BPCI program and runs from the first quarter of 2011 through the first quarter of 2013. For average inpatient payments (Figure 2), average length of stay (Figure 3), 30-day all-cause mortality rate (Figure 4), and 30-day all-cause readmission rate (Figure 5), the paths are quite parallel, and for Figure 4 they are almost identical. This provides evidence that the parallel trend assumption is satisfied. These figures are based on unadjusted averages, while DiD models will include additional hospital and patient characteristics to capture some of the residual differences between Awardee and comparison hospital cohorts. Section III presents similar trends for measures assessed in this report.

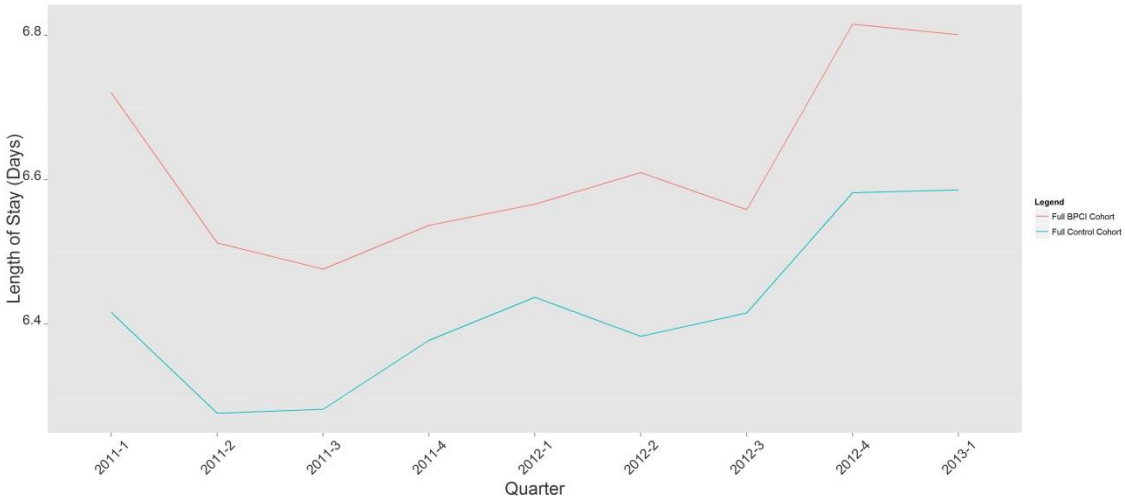
[59] Differences in the *level* of the dependent variable (e.g., historically, BPCI sites tend to have higher expenditures than non-BPCI Awardees) do not pose a problem for DiD. This assumption is discussed in Trivedi, P. K., & Cameron, A.C. (2005). *Microeconometrics: Methods and Applications*. New York: Cambridge University Press, p. 770.

Figure 2: Average Inpatient Payments by Quarter for BPCI Model 1 Hospitals and Comparison Hospitals (in Dollars)



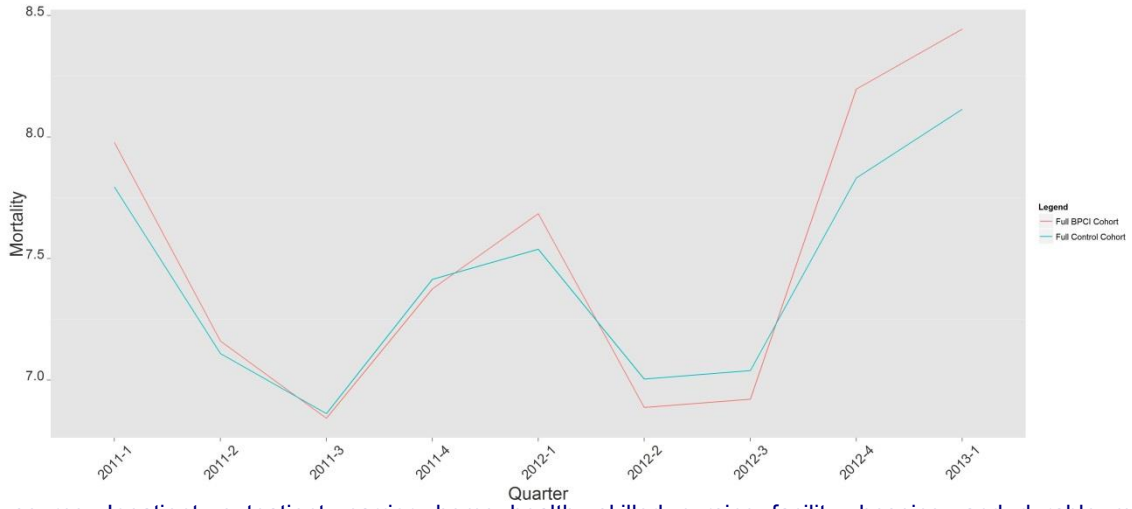
Data source: Inpatient, outpatient, carrier, home health, skilled nursing facility, hospice, and durable medical equipment claims.

Figure 3: Average Length of Stay by Quarter for BPCI Model 1 Hospitals and Comparison Hospitals (in Days)



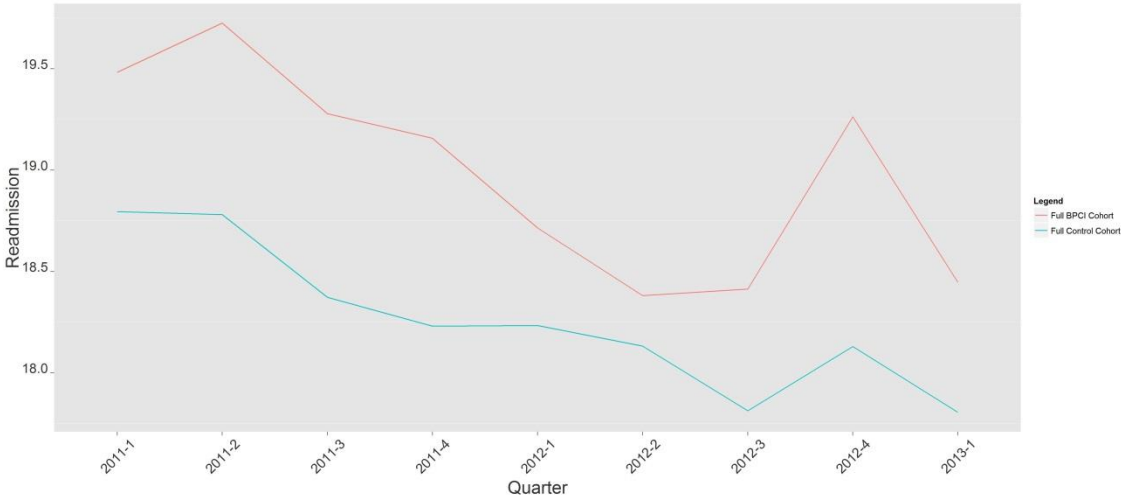
Data source: Inpatient, outpatient, carrier, home health, skilled nursing facility, hospice, and durable medical equipment claims.

Figure 4: 30-Day All-Cause Mortality by Quarter for BPCI Model 1 Hospitals and Comparison Hospitals (%)



Data source: Inpatient, outpatient, carrier, home health, skilled nursing facility, hospice, and durable medical equipment claims.

Figure 5: 30-Day All-Cause Readmissions by Quarter for BPCI Model 1 Hospitals and Comparison Hospitals (%)



Data source: Inpatient, outpatient, carrier, home health, skilled nursing facility, hospice, and durable medical equipment claims.

II.C.3. Hospital Groupings

Awardee hospitals pursue care redesigns that differ in number, type, objective, and focus on patient populations. Additionally, from April 1, 2013, through June 30, 2014, nine hospitals terminated their Awardee Agreement, with six of these nine hospitals having recent gainsharing experience in PHC. To assess differences between Awardees that terminated their agreements and those that remained active and to account, in part, for differences in care redesigns across Awardees, impacts were analyzed across the following subcohorts:

- *Active, Exiting,* and *PHC* Awardees.
- Awardees with *Expansive* or *Targeted* Care Redesigns.

Active and Exiting Awardee cohorts are exclusive sets of Awardee hospitals, as are Expansive and Targeted cohorts. The next two subsections detail subcohort compositions.

II.C.3.1. Active and Exiting Awardees

As previously noted, 23 Awardees began BPCI Model 1 on April 1, 2013, and 1 Awardee began on January 1, 2014. As of July 1, 2014, nine of these Awardees terminated their Awardee Agreement with CMS. In other words, nine Awardee hospitals were not active in BPCI Model 1 for the duration of this report's Medicare claims analyses, including episodes through June 30, 2014.

These 9 Awardees comprised the BPCI Exiting Cohort, while the remaining 15 comprised the BPCI Active Awardee Cohort. It is important to note that these nine exiting Awardee hospitals exited at different points in time: six exited in PQ 5, and three exited in PQ 3. The six Awardee hospitals that participated in the PHC gainsharing program were among these nine Awardee exiting hospitals. These Awardees were classified in the *Exiting* cohort and comprised the entirety of the *PHC* cohort. All exiting Awardees were *active* in BPCI Model 1 for at least 60 percent of the program performance period analyzed in this report.

II.C.3.2. Expansive and Targeted Care Redesign

Care redesigns undertaken by Awardees are an integral design component of this model. To better understand how they might affect model progress, Awardee care redesigns were classified by their potential and expected ability to affect episode or post-episode health care utilization, outcomes, and Medicare payments. Awardees chose to pursue anywhere between 2 and 9 care redesigns under this model, most of which are composed of various care (sub) processes.[60] Further, these care redesigns vary across Awardees, both in measurement of progress and success and even objective. A systematic classification of Awardee care redesigns was developed to group similar care redesigns based on their potential and projected effects and scope, in terms of targeted patient population(s).. This classification process yielded the following two subcohorts:

- *Expansive Care Redesign Awardees.* Awardees in this group included hospitals that, on average, had care redesigns that focused on a larger proportion of a hospital's population or many hospital leverage points (e.g., the entire patient care continuum within a hospital).
- *Targeted Care Redesign Awardees.* Awardees in this group included all other hospitals. These Awardees, on average, had care redesigns that were more focused in terms of patient populations or specific hospital leverage points (e.g., focused on discharge efficiency).

This classification was done prior to Awardee terminations from the model and will be updated for the 2015 Annual Report to account for any changes to hospital care redesigns. Exiting Awardees were included in this classification, since the nine exiting Awardees were active in the

[60] Section III.A provides more detail on the intricacies of care redesigns.

model for the majority of the analysis period and all expressed the intention to maintain their care redesigns.[61]

The data source for this classification were Awardee-submitted IPs that included care redesign names, methods of implementation, targeted population(s), and mechanism(s) for achieving Model 1 goals. This information was submitted to and accepted by CMS prior to an Awardee executing a BPCI Model 1 Awardee Agreement with CMS.

II.C.3.2.1. Potential to Affect BPCI Model 1 Goals

A full overview of this classification process is in Appendix D. Briefly, a modified Delphi process[62] was employed to classify care redesigns pursued at Awardee hospitals by their potential/expected ability to affect Model 1 goals. For this taxonomy, care redesigns were scored across each Awardee in terms of their expected effect and the patient populations over which that effect may occur. Respectively, these two considerations accounted for the potential effect of care redesigns and the scope of that effect. Then, these scores were combined and aggregated across care redesigns *within* an Awardee hospital to determine the *average* care redesign effect and scope at each Awardee hospital for classification into the Expansive or Targeted group.

Scoring the potential/expected effect of care redesigns on Model 1 goals was a multistep process that involved categorizing the type of care redesign, assessing its potential to affect model goals, and then considering the scope of its implementation. The general steps of this process are reviewed below.

Step 1:

Each care redesign was categorized by the *type* of activity and/or leverage point that it was intended to target. Leverage points are the areas, processes, or precise decision points that hospitals/providers target to achieve cost savings and/or improve care coordination and the quality of care they provide. Care redesign type categorizations were aligned with potential organizational responses under this model that include the following:

- Care coordination (e.g., management of transition of care services).
- Material management (e.g., standardization of hardware or other materials).
- Business operations (e.g., internal reporting processes).
- Standardized orders/protocols (e.g., evidence-based checklists).
- Quality improvement initiative (e.g., fall prevention program).
- Education/training (e.g., physician education programs).

[61] One-third were active for at least 60 percent of this report's analysis timeframe; most were active up through the end of the analysis timeframe in PQ 5. See Table 2 and the section above (II.C.3.1) for more information.

[62] Each team member who participated in this exercise independently scored the interventions at each hospital based on predetermined type scores (Appendix D). Once the care redesigns were independently scored, the team members convened to share their scores for each care redesign at each hospital. The "agreement rate" among scorers was quite high, which meant that the team gave the same score for a large majority of the care redesigns. Any discrepancies were discussed at the time of comparison, and a consensus score was determined. Deviating from a "pure" Delphi decision-making process (e.g., the process was not monitored by a facilitator but by consensus) allowed for meeting the needs of the activity while preserving the integrity of the process.

Each of these care redesign types was scored on its potential to reduce cost and improve quality on a scale of 1 to 5, where 1 indicated the least potential effect on BPCI outcomes and 5 indicated the greatest potential effect. When calculating these scores, three factors were considered: (1) published peer-reviewed literature, (2) impacts of previous gainsharing demonstrations, and (3) the remaining potential of commonly used care redesign categories to reduce costs and/or improve quality. A single score for each of the care redesign types was determined through an iterative consensus process inherent to the Delphi method.

To illustrate, researchers have found that education and training activities frequently have a limited impact on the actual quality and cost outcomes experienced by patients, as these activities have a significant degree of separation from the clinical care being provided on a daily basis. As such, care redesigns categorized as education/training should have a care redesign type score of 2. Comparatively, literature indicates that care coordination activities tend to have a broader impact on quality and cost outcomes, since such activities directly affect the way medical professionals provide care. Consequently, care redesigns typed as care coordination efforts received a higher score of 4 than those typed as educational.

Step 2:
The *intensity* of each care redesign type was assessed at each Awardee hospital. A care redesign's intensity was scored in terms of three elements: (1) the care redesign's capacity to affect costs and/or quality, (2) the scale of the care redesign (e.g., one departmental unit versus hospital-wide or implementation of a care redesign versus planning a care redesign), and (3) the compliance and adherence policies associated with the care redesign (e.g., physicians must meet certain requirements or thresholds to receive gainsharing). These *intensity* scores ranged from 1 to 5, with 1 indicating the least projected effect on BPCI outcomes and 5 indicating the greatest projected effect.

Step 3:
A unified *care redesign effect score* for each care redesign was created through a multiplicative process (Appendix D) using care redesign *type* and *intensity* scores. When constructing the multiplicative scoring model, the care redesign type score—a potential score—was considered to be the maximum possible score, with the intensity score—a projected score—acting as a limiting factor. As an example, consider a care redesign that implements standardized order sets that are designed to reduce post-operative infection rates. This care redesign would receive a care redesign *type* score of 4 on its ability to impact model goals (e.g., promote efficiency of care and ultimately reduce costs). However, if the stated goal of the order set was only to maintain the current performance of the hospital's infection rates, then a limited impact of this care redesign was expected and would receive a low care redesign *intensity* score. These *care redesign effect scores* were averaged across Awardee care redesigns to classify the combined impact of the multiple care redesigns each Awardee planned under this model.

Step 4:

Care redesign effect scores were qualified by a care redesign's expected *reach* in terms of patient population(s). For example, if a care redesign's focus was on implementation of standardized order sets but implementation was limited to cardiology patients, the care redesign effect score would be adjusted by the expected number of cardiology patients at that hospital.

First, IP information was utilized to determine whether any care redesigns targeted specific patient populations.[63] Next, these targeted patient populations were broadly classified into seven distinct populations, across care redesigns:

1. Cardiology
2. Orthopedic
3. Neurology
4. Pulmonary
5. Surgery
6. Intravenous immunoglobulin (IVIg)
7. Sepsis patients

These populations were based on the broadest categorization of the patient population deemed reasonable upon consultation with physicians and public health experts. For example, if an Awardee specified the targeted population as heart failure, then the care redesign was considered to be relevant to all cardiology patients. When in doubt of a patient population, the most inclusive population was chosen. This was done in an effort to capture cases in which the care redesigns might have spillover effects into related patient populations and to account for potential variation in coding practices.

Using claims data from the CMS Public Use File (PUF), the fraction of each Awardee's Medicare claims in 2012 for these patient populations was calculated. This calculation relied on the Major Diagnostic Category (MDC) designation included as part of the CMS PUF file or relevant MS-DRG designations. This fraction represents the fraction of patients potentially affected by the given care redesign (i.e., that care redesign's target area/scope). For example, suppose a care redesign targeted orthopedic DRGs, and claims data showed one-third of that Awardee's patient population had the corresponding MDC code (MDC = 8 for orthopedics); in this case, that care redesign would receive a target area score of 0.33. If a care redesign was expected to apply for all patients (e.g., overall improved discharge coordination), then that care redesign's target area score was given a value of 1.0. Finally, each care redesign effect score was multiplied by its target area score and the result was averaged for all care redesigns within a hospital.

As previously noted, the objective was to create a taxonomy over these varied care redesigns by their potential effect on model goals by accounting for the potential effect of an Awardee's care redesigns and the populations over which that effect may occur. Higher scores indicated a broader attempt with care redesigns in terms of their expected effect and scope and the top third

[63] Typically, targeted populations were indicated or implied by care redesign descriptions in an Awardee's IP.

of Awardee hospitals were placed in the Expansive Care Redesign cohort[64]. Section II.C.3.3 list all (sub)cohorts analyzed in this report.

II.C.3.3. Awardee Cohorts for Analyses

The aforementioned hospital groupings (Active, Exiting, Expansive, Targeted, and PHC) all consisted of Awardee hospitals. Comparison hospitals matched to an Awardee hospital are also included in that Awardee hospital's group. In other words, if comparison hospitals B, C, D, and E were matched to Awardee hospital A and Awardee hospital A was in the Active and Expansive cohort, then comparison hospitals B, C, D, and E were included in the Active and Expansive cohorts. In addition to these five hospital groupings, analyses in this report also comparatively analyzed all Awardee and comparison hospitals as a Full cohort. Table 10 details membership of these groupings.

Table 10: Hospital Cohort Membership

Cohort	Number of Awardee Hospitals	Number of Matched Comparison Hospitals
Full	24	96
Active	15	60
Exiting	9	36
PHC	6	24
Expansive	8	32
Targeted	16	64

II.C.4. Analyses

Each measure included in this report may include analyses on unadjusted and adjusted statistics in quarterly trend analyses, aggregate time period comparisons, and DiD impact analyses. These analyses were conducted at the *Awardee level* and the *Awardee cohort level*.

Awardee-level analyses involved analyzing each Awardee relative to their matched comparison hospitals across observed and (risk) adjusted measure statistics over time. Further, these analyses assessed BPCI Model 1 Awardee-level impacts on measures by utilizing quasi-experimental DiD regression models. These regression models leveraged the aforementioned counterfactual scenario directly by comparing Awardee measure differences before and after BPCI Model 1 inception (i.e., Baseline vs. Since BPCI Inception) to like differences for comparison hospitals. Additionally, these models controlled for both episode and patient characteristics to account for residual differences between Awardee and comparison hospitals that were not otherwise captured by the comparison hospital selection process.

Awardee *cohort*-level analyses also included analyses of observed and adjusted measure statistics and DiD measure impacts for cohorts noted in Table 10.

[64] The lowest ranked Awardee in the top third of the distribution was replaced with an Awardee that had the same average score and almost double Medicare admissions of the lowest ranked Awardee in the prior year.

Table 3 and Appendix A detail inclusion and exclusion criteria for each measure presented in this report. Sections II.C.2 and II.C.3 detailed comparison hospital selection and hospital groupings for cohorts analyzed. The next subsections detail unadjusted, adjusted, and DiD methodologies. Analyses were conducted in STATA 13.

II.C.4.1. Unadjusted Statistics

Unadjusted statistics are calculated at the episode level in accordance with specifications provided in Appendix A and averaged at the hospital and time period level (quarter, Baseline, Since BPCI Inception, and Year 1). These statistics were subsequently averaged across Awardee or comparison hospitals within a cohort (e.g., Full cohort). Thus, for a given cohort and time period, Awardee hospital measure statistics within that cohort are averaged and compared to the appropriate comparison hospital average.

For a given measure, these average statistics are presented as weighted (by the number of hospital episodes over a certain time period) or unweighted (i.e., arithmetic mean) at cohort levels. Statistical significance is tested within Awardee or comparison cohort groups across time periods by two-tailed t-tests that allow for unequal variance across samples tested.

II.C.4.2. Adjusted Statistics and DiD Impact Analysis

Adjusted statistics and adjusted impact (DiD) estimates were also calculated. These estimates adjust for patient demographic and clinical risk factor differentials across episodes at Awardee and comparison hospitals to further equitable comparative analyses.

Adjusted measure statistics were estimated over all study period quarters from January 1, 2011, through June 30, 2014, with regression models that account for these differentials and differences across time and (Awardee and comparison) hospitals in the study sample. Further, the entire study population was utilized—all patients from Awardee and comparison hospitals—as a reference population for adjusted measure statistics. In other words, after accounting for the factors above, equivalent patient populations were imposed for each hospital in the study sample.

Impact analyses come from DiD regression models. DiD is a quasi-experimental policy analysis tool that enables longitudinal comparisons of measure outcomes for Model 1 Awardee hospitals with those of comparison hospitals. An assumption for DiD models, known as *parallel paths,* requires that trajectories of the dependent variables (i.e., measure outcomes) in a DiD model must be parallel for two cohorts before the introduction of a study *intervention*. The translation of this assumption for the DiD impact analyses is that performance measure trends for Awardee and comparison hospital groups be parallel[65] over the Baseline period; i.e., the time period *preceding* BPCI Model 1 (the "intervention") implementation periods.

Comparison hospital measure outcomes over time allow inference of a counterfactual scenario: What would Model 1 Awardee outcomes have been had they not participate in Model 1? In

[65] Differences in the *level* of the dependent variable (e.g., historically, BPCI sites tend to have higher expenditures than non-BPCI participants) do not pose a problem for DiD. This assumption is discussed in Trivedi P.K. and Cameron A.C., *Microeconometrics: Methods and Applications*. New York: Cambridge University Press, 2005, p. 770.

comparing measure outcomes during the BPCI Model 1 implementation period between Awardee and comparison hospitals, differences in measure performance can be attributed to BPCI Model 1 participation. DiD regression models follow the specification below:

$$Y_{iht} = \alpha + \sum_{t=1}^{k} \beta_t D_{ht} + D_h + \theta POST_t + \gamma X_{iht} + \tau_t + \lambda_h + \epsilon_{iht}$$

The dependent variable Y_{iht} is the measure of interest for episode i[66] receiving services at hospital h at time t. Further,

D_h: Cohort indicator equal to 1 if an episode occurred at a BPCI Model 1 Awardee hospital and 0 otherwise.

Post: Time indicator equal to 1 if an episode occurred on or after BPCI Model 1 performance period and 0 otherwise.

D_{ht}: Policy indicator equal to 1 if an episode occurred at a BPCI Model 1 Awardee hospital, after they executed a BPCI Model 1 Awardee Agreement with CMS and 0 otherwise. β_t is the estimated impact of BPCI Model 1.

τ: Quarter (time) fixed effects.

λ: Hospital fixed effects[67].

X_{iht}: Additional hospital and episode characteristics listed in Table 11.

Table 11: DiD Model Specification Patient and Episode Characteristics

Measure
Patient Age
Patient Gender
Patient Dual Medicare/Medicaid Eligibility Status
CMS- HCC Score
Episode MS-DRG

A second model is also estimated. This additional model includes an interaction of quarter period indicators ([68]) with the policy variable (D_{ht}) and allows inferences on whether particular program quarters are influencing the overall impact estimate of BPCI Model 1 for a given measure.

In this model, hospital fixed effects account for time-invariant heterogeneity among Awardee and comparison hospitals. Further, this DiD specification also accounts for time effects (by quarter) and episode characteristics. This regression specification is utilized for BPCI Model 1 impact estimates in Section III. Table 12 lists the DiD model types for each measure assessed in

[66] These reference patient-level outcomes; e.g., whether a patient had a mortality event 30 days after inpatient admission.
[67] All Awardee and comparison hospitals were included in every quarter of the study.
[68] *T:* Set of quarter indicator variables (not shown in model above).

this report. Note that unadjusted and adjusted statistics for mortality, readmission, and ICU utilization are in terms of percentages, while DiD impact estimates for these measures assess binary outcomes at the episode level such as whether or not an episode lead to a readmission event. As such, these impact estimates for these measures are assessed as likelihoods, presented as odds ratios[69], with logistic regression assumptions. Regression model standard errors were clustered at the hospital-quarter level.

Table 12: DiD Model Type by Measure

Domain	Measure	Model Type
Health Care Outcomes and Resource Utilization	**30-day all-cause mortality**	Logistic Regression
	30-day condition-specific mortality:	
	Acute myocardial infarction	Logistic Regression
	Pneumonia	Logistic Regression
	Heart failure	Logistic Regression
	30-day all-cause readmissions post-episode	Logistic Regression
	30-day condition-specific readmissions post-episode:	
	Acute myocardial infarction	Logistic Regression
	Pneumonia	Logistic Regression
	Heart failure	Logistic Regression
	ICU use during episode	Logistic Regression
	Episode length of stay	OLS Regression
Medicare Payments	**30-day post-episode Medicare payments**	OLS Regression
	Total episode payments	
	Non-hospital payments during episode	OLS Regression
	Hospital-only payments during episode	OLS Regression

[69] Future work will also assess marginal impacts for these estimations.

Section III. Results and Findings

III.A. Awardee Implementation and Organizational Responses to Model 1

This section provides insights into the progress of the BPCI Model 1 implementation. Data from telephone interviews and in-person focus groups ("site visits") captured Awardee progress, successes, challenges, and outcomes from multiple perspectives across Awardee hospitals. As previously noted, Model 1 Awardees engage and enroll physicians and other medical practitioners to implement redesigned care processes and share internal hospital cost savings engendered under this model with enrolled medical practitioners ("gainsharing"). Gainsharing is expected to promote successful implementation of model activities, adherence to model requirements, and efficiency gains from redesigned care processes. Further, these efficiency gains may translate to the Model 1 goal of reducing Medicare expenditures without sacrificing quality of care or patient access to care.

Collection and analysis methods were discussed in Section II.B, and Tables 4 and 5 provided telephone interview and site visit characteristics, participation, and periodicity. Data collection activities and findings from the data are summarized below:

Data Collection Mechanisms:

- *Semi-structured telephone interviews,* which consisted of 45-minute discussions with the designated administrators (e.g., BPCI Model 1 PC or CMO) of Awardee hospitals. These interviews were conducted in two waves (April 1, 2013, to December 23, 2013, and January 3, 2014, to May 31, 2014).
- *Site visits,* which consisted of 60- to 90-minute sessions with three categories of focus group participants: (1) Hospital Administrators, (2) BPCI Model 1 Enrolled Practitioners, and (3) Care Redesign Team Leaders. Focus groups ranged from five to eight participants. In situations where three or fewer participants attend a focus group session, the moderator switched to a face-to-face interview format.

Common Themes:

- Strong leadership and clinical staff education are critical in deterring skepticism about the model and promote engagement. While hospital administrators are generally satisfied with the levels of engagement, some reported dissatisfaction with the engagement of physician staff and claimed that a lack of or low physician engagement was detrimental to achieving quality improvement and hospital internal cost reduction goals.
- The sole FC provided valuable assistance to New Jersey Awardee hospitals at the onset of the model in terms of administrator and physician education, understanding of model design components, and interactions with CMS. This guidance and support eased the Model 1 implementation process. For example, the FC prepared a Model 1 program physician handbook that provided further guidance on program requirements and implementation) for New Jersey Awardees.

- All care redesigns were fully implemented, but more work is required to realize efficiency gains (e.g., reduction in duplications of health care services) and, potentially, quality improvements.
- To date, Awardees have distributed at least one round of gainsharing payments. Overall, Awardee hospital administration felt that gainsharing payments did incentivize behavior change. However, most Awardees with employed physicians did not see gainsharing as necessary for employed physicians relative to voluntary physicians, as employed physicians were already expected to follow hospital protocols, including care redesigns. Several administrators expressed dissatisfaction with the timeliness of internal hospital cost savings and gainsharing payment data as delays inhibited (near) real time information to Awardee staff (including physicians) that may otherwise allow Awardees to assess and adjust their care redesign implementation if necessary.
- Telephone interviews with administrators at nine exiting Awardees revealed challenges these Awardees were facing in relation to their BPCI participation. Their decision to terminate the Awardee Agreement was driven by many factors, including their inability to associate perceived or realized cost savings from care redesigns, the timing of cost-saving realizations from care redesigns, lack of physician enrollment and engagement, and/or the hospital-physician employment structure (i.e., level of employed versus voluntary physicians).

The following subsections discuss these findings in further detail and address their implications.

III.A.1. BPCI Model 1 Participation

The longevity of BPCI Model 1 relies on continued Awardee participation and the number and engagement of enrolled physicians. This subsection provides an overview of Awardee participation and physician enrollment and engagement since model inception in April 2013. Various factors impacting these participation aspects are also presented.

III.A.1.1. Awardee Participation

At the inception of Model 1, Awardees included 23 hospitals located in the State of New Jersey that voluntarily participated in the program. KSRC was the last hospital to sign an Awardee Agreement with CMS for the model, in January 2014. As of July 1, 2014, 9 Awardee hospitals terminated their Awardee Agreement with CMS, bringing the number down to 15 Awardees.[70] Telephone exit interviews were conducted with one to three Awardee hospital administrators[71] 30 days prior to the effective termination date[72] to identify motivations for Awardee Agreement termination. Results from these interviews indicate that Awardees terminated for the following reasons:

[70] As of March 2015, 12 of the 24 Awardees terminated their Awardee Agreements with CMS.
[71] These administrators typically included the model PC, CMO, and Chief Financial Officer; however, interview participation did vary across exiting Awardees.
[72] Awardees may terminate their Awardee Agreements by providing CMS with at least 60 calendar days' prior notice of the effective date of such termination.

1. Hospital-physician employment structure already required/motivated physician adherence to protocols and participation in care redesigns through existing employment/contractual relationships.
2. Inability to associate perceived or realized internal hospital cost savings from care redesigns. Relatively long timescales involved in realizing potential internal hospital cost savings from care redesigns.
3. Lack of physician enrollment and engagement resulting in insufficient critical mass to generate the change in practice required to deliver cost savings from care redesigns.

Section III.A.4 further details these reasons. Awardee hospitals remaining in BPCI Model 1 cited some of these reasons as areas of concern in their continued participation but, overall, believed it worthwhile to stay in Model 1 to continue furthering clinician and hospital alignment for efficiency gains, despite Model 1 requirements and its IPPS discount.

III.A.1.2. Physician Enrollment and Engagement

Administrators employed a number of methods to recruit and engage physicians into the model and maintain their engagement once involved. The recruitment methods, which generally varied based on an Awardee's organizational staffing structure, included the following:

- *Targeted approach.* This approach focused on enrolling high-volume or high-impact physicians.[73] It involved one-on-one meetings between model administrators and physicians identified as having high patient caseloads (relative to other physicians) or focusing on high-cost patients where cost reductions may occur from specific care redesigns that focused on those patients.
- *Scaled approach.* This approach focused on one or two service lines within a hospital, with plans to extend enrollment efforts at a later date.
- *Cascade approach.* This approach focused on initiating the enrollment process at the Board or Medical Executive Committee level, then "cascading" enrollment efforts down through department heads to other physician staff.
- *Mixed-methods approach.* This approach incorporated multiple methods, utilizing a range of techniques, including presentations, email newsletters, opportunistic meetings in the physician's lounge (if present), and other methods to promote enrollment.
- *Opt-out approach.* This approach automatically enrolled all eligible physicians into the program. If physicians chose not to enroll, they were required to opt out. There was no penalty to physicians for choosing to opt out.

All hospital administrators were initially satisfied with the levels of physician enrollment and engagement in the BPCI Model 1 care redesigns. However, as the program progressed, some Awardees struggled to engage physicians, while others succeeded. Saint Michael's Medical Center used a mixed-methods approach to provide physicians with information about BPCI Model 1, an effort that only enrolled 12 physicians. As expected, this Awardee was concerned

[73] High impact can be characterized as physicians who take on patients with high cost and cost variation and/or a large number of Medicare patients.

with low levels of physician enrollment for its care redesign and ultimately the program to succeed. The Awardee found itself in a tenuous situation with an impending change in ownership, so administrators felt physicians questioned the logic of enrolling in BPCI Model 1 at a "failing" hospital and resisted enrolling.

In contrast, RWJU employed a targeted approach to recruit physicians, which focused on enrolling physicians with high Medicare patient volume. Administration at this hospital used its position as a teaching hospital to its advantage, involving residents in care redesign processes to promote efficient practices. The hospital faced challenges initially due to an incomplete understanding of the program by physicians. The hospital overcame this challenge, in part, by mandating the use of the Crimson tool,[74] and encouraging enrolled physicians to review their performance data quarterly.

Based on the findings from the interviews, it is difficult to determine which recruitment strategy was more effective in increasing enrollment levels. This difficulty is further compounded by the lag in physician enrollment data and the use of multiple, possibly concurrent, recruiting approaches across Awardee hospitals.

Next, common factors that have impacted physician engagement in this model are discussed.

III.A.1.3. Impacting Factors

The involvement of clinical staff and Awardee hospital progress in BPCI Model 1 is influenced by various factors, including:

- Awardee characteristics.
- Administrative leadership.
- Physician skepticism.
- Physician education.
- Longevity/tenure of clinical staff.
- Facilitator Convener.
- Physician incentives.

III.A.1.3.1. Awardee Characteristics

Awardee hospitals have distinct features, including size, staffing structure, organizational structure, and culture. These features appear to have a substantial impact on the overall physician engagement and progress towards model and Awardee goals. Administrators identified the *size* of hospitals, particularly small (less than 100 staffed beds) to medium (101 to 499 staffed beds), as a potential facilitator of success. One Awardee considered *staffing structure* (such as having a "closed" or employed/internal staffing structure) as another enabling factor. Because *staff members were more invested*, another felt it was their *organizational structure* (as a freestanding community facility that was not affiliated with a larger system) that helped it succeed.

[74] The Crimson tool is a physician performance data management platform marketed by The Advisory Board Company.

Administrators, physicians, and non-physician clinical staff across Awardees generally believed that a hospital's *culture* was a key attribute for successful implementation.[75]

For example, administrators at RWJU viewed its teaching hospital status as a mechanism by which support for BPCI Model 1 is enhanced; the academic atmosphere and innovation-friendly environment were reported as catalysts that enhance support for BPCI Model 1 among hospital staff. Saint Peter's unique organizational staffing structure leaves many members of the medical staff ineligible to enroll in BPCI Model 1. A significant proportion of this Awardee's staff are pediatricians and obstetricians, specialties that are not included in the program. The inability to involve a large proportion of the medical staff was thought to be a challenge for engagement and decreases the ability to diffuse change across the institution.

III.A.1.3.2. Administrative Leadership

Administrators and clinical staff cited *strong* administrative leadership as a primary contributing factor for successful BPCI Model 1 practitioner engagement. Committed and responsive leadership reportedly helped the program succeed (e.g., by increasing engagement, diffusing information about program activities), despite competing or concurrent organizational priorities such as Medicare's Sustainable Growth Rate or other sequestration effects. Leadership was also reported as instrumental in overcoming resistance to implementing model care redesigns across multiple levels of staff.[76] For example: Saint Peter's employed a new Six Sigma-certified[77] PC to "relaunch" BPCI Model 1 among clinical staff and decrease physician skepticism of the model (discussed below).

III.A.1.3.3. Physician Skepticism

Physicians' skepticism concerning Model 1 design and implementation emerged as a major barrier to engagement and overall initiative implementation success. Administrators noted some difficulty in enrolled physician engagement due to physician skepticism of the model—in regard to gainsharing performance data calculations—or general lack of interest. Site visit interviews[78] and focus groups conducted with enrolled physicians were valuable in obtaining physician perspective on skepticism, which varied across Awardees.

Physicians' skepticism appears to stem from multiple factors:

- *A lack of understanding about the program and gainsharing arrangement.* For example, physicians at JFK, St. Joseph's, and RWJ Rahway expressed concerns that BPCI Model 1 was inherently unfair, as it rewarded "bad" doctors. BPCI Model 1 was thought to reward physicians who waste hospital resources and/or provide a lower quality of care, because these physicians have more room to improve and get higher gainsharing payments from their performance outcomes than physicians who consistently provide high-quality care.

[75] Respondent types (i.e., administrative, physician, nonclinical physician) did not necessarily agree on whether an *effective* culture was in place at an Awardee hospital.
[76] See Sections A.1.3.4, Physician Education, and A.1.3.5, Longevity/Tenure of Clinical Staff, for more information.
[77] Six Sigma is a management philosophy developed to improve existing business processes by utilizing statistical analysis and a DMAIC (define, measure, analyze, improve, control) approach to problem solving.
[78] Some site visits, especially those occurring during the first program year, had low physician turnout. In some instances, if there were only one or two attendees, interviews were conducted in place of expected focus groups.

- *Capabilities to achieve BPCI Model 1 goals*. Physicians at RWJ Rahway believed they could no longer reduce utilization without compromising care.
- *Changes in the Nation's health care policy* (e.g., health care delivery models, reimbursement strategies). Physicians at CentraState, JFK, RWJ Hamilton, and RWJ Rahway had reservations about enrolling in BPCI Model 1 due to potential policy changes in the future that may negatively affect their compensation.

In an effort to deter physician skepticism about the model, hospitals and the FC (for New Jersey Awardees) employed various approaches. Information dissemination regarding the model through educational material, learning sessions, and enhanced communication (e.g., newsletters) aimed to increase awareness of the model[79] and dispel misconceptions among staff. The impact of these approaches is not clear.

III.A.1.3.4. Physician Education

To encourage a higher degree of engagement, administrative staff reported that it is essential to provide clinical staff access to a broad range of information and education regarding the components and goals of BPCI Model 1.

Administrators and physicians at RWJ Rahway and Saint Peter's believed physician education is essential to overcoming physician skepticism associated with BPCI Model 1. Education, typically through informational sessions, proved to increase physicians' faith in the data presented through the program and adoption of care redesigns. However, some hospitals did not have the time or resources to devote to education. JFK and RWJU, for example, reported that the amount of time necessary to educate physicians and bring them up to speed with BPCI Model 1 activities was burdensome relative to the financial gain. This was especially relevant at RWJU, whose leadership explained that different methods of outreach and education were required for different types of physicians (e.g., hospitalist, community physician, surgeon); this presented a challenge for this Awardee's hectic teaching hospital environment.

III.A.1.3.5. Longevity/Tenure of Clinical Staff

The longevity of clinical staff in the hospital emerged as an additional factor influencing physicians' enrollment and engagement. Interviews with administrators and focus groups with physicians revealed mixed responses regarding the nature of the impact of clinical staff longevity. JFK, RWJU, and RWJ Rahway characterized longevity as an advantage, because staff members with relatively longer tenure were likely more willing to undertake quality initiative-type efforts with a positive outlook. Low turnover and positive morale play a critical role in fostering significant leadership and staff support for the BPCI Model 1 initiative. Longevity helps establish long-term relationships and contributes to increased institutional knowledge about how to effectively work together on such initiatives as BPCI Model 1. Conversely, JFK raised the issue that longevity can have negative aspects as well. Physicians may demonstrate unwillingness to accept changes in set practices.

[79] See Section III.A.1.2, Physician Enrollment and Engagement.

III.A.1.3.6. Facilitator Convener (FC)

New Jersey Awardees generally believed that their FC positively influenced the success of the program. Awardees described the organization as "very supportive," "always available," "able to answer any questions," and "quick to respond." Awardees expressed satisfaction with the range of support provided by the FC, including support with the application process, physician recruitment and enrollment, tutorials and online (reporting) portals, explanations of cost-savings data and performance reports, and facilitation of cross-hospital learning. This support was noted as extremely helpful during initial implementation, but now that the program is up and running, the demand for convener assistance diminished.

There were few concerns with the FC's assistance. An administrative respondent from one Awardee expressed frustration with the FC's expectations regarding information requests and the timeliness of internal hospital cost-savings data. Another Awardee reported a negative experience regarding the implementation and use of the Provider Enrollment, Chain, and Ownership System, which affected physician enrollment.

III.A.1.3.7. Physician Incentives

The effects of physician gainsharing/incentive payments on physician enrollment and engagement vary across Awardees. Some Awardees did not report changes in physician enrollment and engagement due to incentive payments. Lack of changes in enrollment could be due to insufficient peer communication about physician incentives[80] or beliefs that incentive payments are unnecessary for practice change. With respect to physician engagement, physician focus groups identified that enrolled physicians felt that the time and resources to meet program requirements were too burdensome when compared to the amount of incentive payment received.

Other Awardees reported that incentives improved enrollment and engagement. Capital Health, Inspira Woodbury, and RWJU reported increased enrollment after the first distribution of physician incentive payments. Specifically, Awardees with lower initial enrollment stated that incentive payments had an impact on enrollment. These payments likely initiated a cultural change and increased physician understanding of the rationale for practice change. For example, one Awardee cited that physicians now research alternate supplies in an effort to be more cost-conscious. Finally, it is believed that incentive payments induced early adoption of new practice protocols. RWJ Rahway and Inspira Vineland are looking to expand Awardee-specific criteria required to receive incentive payments because payments had a positive effect on the adoption of care redesigns, and increased compliance for some interventions.

III.A.2. BPCI Model 1 Design Characteristics

BPCI Model 1 aims include maintaining or improving the quality of care provided to patients and lowering health care costs. Under this model, Awardees engage in redesigns to care processes to meet these goals and align physician behavior to support care redesigns through incentive payments (i.e., gainsharing). Participation within this model was presented in the

[80] Administrators noted that enrolled physician communication about BPCI Model 1– in addition to administrator efforts - with non-enrolled physicians may increase enrollment of non-enrolled physicians.

preceding section. This section presents Awardee perspectives on model activities that relate to two key model design components: (1) care redesigns and (2) incentive mechanisms.

III.A.2.1. Care Redesigns: Hospital-Driven Changes

For BPCI Model 1, each Awardee developed and implemented its organizational infrastructure early on. These include developing care redesigns plans, establishing goals and expectations for these redesigns, and establishing groups and committees to guide model activities. The Steering Committee at an Awardee hospital determined internal hospital rules and care redesigns for that Awardee under BPCI Model 1. The BPCI Model 1 PC at a hospital oversees all model activities.

Awardee Steering Committees defined the BPCI Model 1 care redesigns, defined gainsharing criteria, and ensured that the content and pace of implementation occurred in equilibrium with BPCI Model 1 goals. Thus, members of an Awardee's Steering Committee ultimately decided which care redesigns to pursue, their steps for implementation, and markers for success. The composition of the Steering Committee varied by Awardee. According to the Awardee Steering Committee minutes, a majority of participants were physicians. Care redesigns are hospital-driven changes; however, physicians enrolled in BPCI Model 1 at an Awardee hospital primarily support their success. Respondents—physicians, non-physician clinical staff, and administrators—further pressed the idea of physician *champions* who would strongly advocate care redesign changes, especially among other physicians, and adherence to quality reporting requirements. Not all Awardees had designated physician champions for BPCI Model 1.

Some care redesigns were underway prior to BPCI Model 1 in many Awardee hospitals—either through a gainsharing pilot, other initiatives, or an Awardee's own initiative to improve quality of care. Thus, most Awardees did not "reinvent the wheel" with respect to their BPCI Model 1 care redesigns. Some Awardees did, however, start care redesigns "from scratch" and noted a steeper learning curve. As expected, these Awardees were relatively less advanced in care redesign implementation than other Awardees. In some instances, Awardees used BPCI Model 1 as a tool to expand existing care redesign activities across their respective facilities; this expansion was accomplished primarily through varied targeting approaches noted earlier.

Overall, BPCI Model 1 care redesigns and corresponding goals varied by Awardee. Each Awardee hospital implemented between two and nine care redesigns. Across Awardees, these care redesigns have also varied in type (e.g., patient flow improvement versus fall prevention programs), ease of implementation, timing of implementation, definition of "fully implemented," and desired outcome(s) from implementation. The care redesigns that Awardees pursued also varied by scope (as measured by patient populations affected) and intensity (measured by clinical staff involved). For example, some Awardees planned to focus on patients with chronic diseases. This care redesign, generally labeled "Chronic Disease Management," would establish a chronic disease management team and/or improve the process of an existing team's predetermined "best practices."[81] Within this broadly defined care redesign, subprocesses (such as improving outpatient management through patient follow-up) were expected to affect readmission rates for

[81] Example: New Jersey participants generally determined best practices through clinical guidelines and statewide practices.

these patients positively. Another care redesign, "Patient Flow Improvement," was also common across Awardees; however, its implementation and scope varied across these Awardees.

Some Awardee hospitals focused on improving (or increasing the use of) patient-tracking software (or processes) to anticipate patient discharges and increase the efficiency of bed turnover for incoming patients. This software could decrease patient wait time, such as wait time in a hospital's ED. Other Awardees implemented patient flow improvement care redesigns that focused on patient discharge planning. This improvement was expected to discharge patients earlier and potentially reduce a patient's length of stay. Other care redesign processes were more nuanced in their aim to meet BPCI Model 1 goals. For example, care redesigns focused on specific outcomes (e.g., length of stay, or readmissions) aimed to educate clinical staff on how to interpret outcome data and monitor the data on a periodic basis. Crimson, a tool that allows physicians to monitor their performance on outcome measures associated with the model with a 2- to 3-month lag, is an example of an Awardee's effort to make clinical staff more data driven. The assumption with this type of care redesign is that physicians would monitor their performance and see where improvements could be made (when compared with best practices).

It is important to note the role of non-physician clinical staff in relation to the implementation of care redesigns. Previous interviews and focus group sessions indicated that non-physician clinical staff members (e.g., nurses) were relevant to care redesign implementation. These staff members helped align physicians to care redesign guidelines by assisting with care redesign compliance, especially when the guidelines differed from what physicians were accustomed to. For example, nurse care redesign leaders mentioned how they initially had to stress that physicians follow a newly developed checklist for their falls prevention care redesign to reduce the number of falls. Or a care redesign implemented at Saint Peter's to decrease 30-day readmission rates primarily involved nursing staff following up with patients by phone. In this instance, this Awardee indicated that there was good engagement from non-physician staff.

In early interviews and focus group sessions (e.g., the start of program year 1), administrative and clinical respondents emphasized that these care redesigns aimed to affect quality first and cost of providing care second, as efficient and efficacious care was expected to reduce cost. Recently, administrators appeared more cost-conscious by discussing whether the care redesigns have resulted in sufficient cost savings to validate a business case for continuing in the program, where they face an across-the-board IPPS discount.

The next section discusses mechanisms used to incentivize physician behavior to align with Awardee care redesigns and goals.

III.A.2.2. Incentive Mechanisms

III.A.2.2.1. Gainsharing

Gainsharing is the incentive mechanism that aims to align physician practice behavior with Awardee care redesigns and goals under BPCI Model 1 through incentive payments. While model restrictions for gainsharing/incentive payments do not vary across Awardees, the physician eligibility criteria for incentive payment do differ across hospitals. The aim of establishing criteria for incentive payment is to increase physician awareness of the Awardee's endeavors in the model directly (e.g., outcome-based payment criteria such as a decrease in

readmission rates) or indirectly (e.g., physicians review their own performance data through Crimson).

AMS provided preliminary gainsharing data during the third PQ[82]. Most Awardees distributed incentive payments by the fourth and fifth PQs. As such, interviews and focus groups in the first three PQs obtained *perceptions* of the incentive payment/gainsharing mechanism of BPCI Model 1. Administrators expected gainsharing to be a positive mechanism for their success in the model, while physicians were generally skeptical of its effects. Specifically, administrators expected gainsharing to increase physician interest and engagement in the model by encouraging competition and resolving physician skepticism. Medical and surgical physicians noted skepticism of or disinterest in the gainsharing component of this model during the first round of interviews. Generally, physician skepticism stemmed from a misunderstanding of the gainsharing computation methodology. Some physicians believed that the methodology rewarded bad doctors, as these doctors had the most room for improvement on measures analyzed (e.g., average patient length of stay relative to statewide best practices). Other physicians expressed disinterest in gainsharing and the model because they believed the monetary incentive was not necessary to provide effective, quality care, which is the very nature of their vocation.

Generally, most Awardees that distributed incentive payments indicated that they perceived the financial incentive as being a key driver of performance enhancement. Administrators mentioned increases in physician enrolment and engagement after the first round of gainsharing/incentive payments were distributed. Inspira Vineland described unenrolled physicians as "intrigued by their colleagues who are getting checks." At St. Joseph's, incentive payments helped generate interest from and engage staff members who were previously uninterested. This new interest triggered further physician enrollment in the model, with five new physicians recruited in the month prior to the interview. Several other Awardees also reported increases in physician enrollment after the first round of incentive payments were distributed. Saint Peter's reported more than doubling its physician enrollment in the program, from 43 to 98 physicians.

Administrators also offered anecdotal evidence of physicians asking PCs how to improve their performance and potentially earn more in gainsharing payments. This was corroborated by an Awardee that reported that one of its most effective recruitment techniques was to demonstrate potential incentive payments physicians could be receiving under BPCI Model 1.

Other Awardees perceived physician attitudes as ambivalent about or even unaffected by the incentive payments. Administrators at these Awardees noted that doctors were independently motivated to participate in care redesign activities and comply with best practice care because of the improvements to patient safety and experience, not because of the financial incentive (RWJ Hamilton and Overlook). Administrators, physicians, and non-physician clinical staff believed that cost reduction and quality improvement goals could be achieved without incentive payments. (Overlook terminated its Awardee Agreement with the BPCI Model 1 program at the time of this writing.)

[82] For New Jersey Awardees.

Several respondents expressed dissatisfaction with the timeliness of internal hospital cost-savings data.[83] The performance and incentive data covering the first PQ was finalized by the fourth PQ. In addition to timeliness, physicians at RWJ Hamilton were frustrated with the penalties associated with physician incentives. Physicians do not feel adequately informed about how they are being graded on the performance indicators. Enrolled physicians also feel that consultants negatively affect their performance and incentive payments. As a result, physicians are less likely to reach out to consultants for guidance.

III.A.2.2.2. Performance Review

Gainsharing payments were conditioned by hospital-determined physician performance metrics. Awardees conditioned gainsharing payments to these metrics and committed to performance reviews (e.g., one-on-one meetings with physicians) to increase physician awareness and potentially fuel physician-peer competition.

Performance reviews provided physicians an opportunity to review recent individual performance data and discuss potential improvement with hospital administrators and Steering Committee members. Leadership at different Awardee hospitals received mixed responses from physicians on their performance reviews. Some physicians appreciated the opportunity to receive and review their performance data and suggestions on areas needing improvement, while others were not showed less interest in granular data, preferring big picture information. Additionally, some Awardees (e.g., RWJ Rahway) experienced challenges engaging or scheduling meetings with physicians. RWJ Rahway found that limiting the meetings to a maximum of 15 minutes and allowing enrolled physicians to schedule a time provided an effective means of encouraging the reluctant enrolled physicians to review data. In addition, presenting the data in a format that is easy to access and user friendly was an important consideration for Awardees.

The approaches to performance monitoring varied by Awardees and its care redesigns. Almost all of the Awardees reported that it was too early in the initiative to provide any definitive answers on progress. Despite these challenges, several Awardees had a number of key performance indicators in place, which included both internal and external sources of information. Key sources of performance monitoring data and information include:

1. Reviews of data.
2. Clinical audits.
3. Patient satisfaction and experience information.

Overall, Awardees used data maintained by AMSto monitor physician and BPCI initiative-wide performance. Several Awardees used other data sources, such as QuadraMed, Press Ganey (for patient satisfaction), HCAHPS, National Patient Safety Goals, Medical Data Exchange or MDX (for case management), and Surgical Compass (for tracking surgical data and time starts). These data may be presented in a variety of ways such as dashboards, internal report cards, and performance metrics. Other Awardees, including KSRC, indicated their intention to use patient satisfaction and long-term outcomes as a measure of care redesign success in the future.

[83] The first round of these data was finalized (approximately) in the fourth PQ, and data in these reports were only for the first PQ of BPCI Model 1.

Saint Peter's noted that the program is moving slower than expected. Leadership believed that the BPCI Model 1 program is highlighting some of the barriers to efficient care, but the data discussions and incentive payment distributions happen infrequently, and changes in behavior are occurring slowly.

III.A.3. Motivations for Termination of BPCI Model 1 Awardee Agreement

III.A.3.1. Overall Assessment of Termination

Nine exit interviews were conducted with administrators at Awardee hospitals that terminated their Awardee Agreement. CentraState, Cooper, and Jersey Shore exited during the first wave of data collection[84]; JFK, Morristown, Deborah, Hunterdon, Overlook, and Saint Michael's exited during the second wave of data collection; and Saint Michael's exited after the second wave of telephone interviews.

Overall, administrators saw "value" in BPCI Model 1. Generally, they believed that CMS initiatives aimed at influencing physician behavior patterns for improved patient health care at reduced costs could be advantageous. However, for various reasons, exiting Awardees could ultimately not support the business case for continuing in a voluntary program, where their profits and revenues were at risk from the IPPS discount. Key reasons for termination included the following:

1. Hospital-physician employment structure already required/motivated physician adherence to protocols and participation in care redesigns through existing employment/contractual relationships.
2. Inability to associate perceived or realized cost savings from care redesigns. The relatively long timescales involved in realizing cost savings from care redesigns.
3. Lack of physician enrollment and engagement resulting in insufficient critical mass to generate the change in practice required to deliver cost savings from care redesigns.

Generally, BPCI Model 1's gainsharing component was seen as a useful mechanism for Awardees with a specific type of internal staffing structure, composed of mostly voluntary physicians whose loyalty to the hospital may need to be incentivized. Indeed, exiting Awardees with employed physician staff reported not seeing gainsharing as a necessary incentive mechanism, as employed (physician) staff members were already invested in hospital care redesigns. Additionally, multiple exiting Awardees explained that it takes more than 6 months to realize implemented care redesigns and potentially more time to realize possible cost savings under the model. Further, if cost-savings and measure data could be attributed to care redesigns and obtained faster,[85] hospital representatives believed that they might be able to justify the IPPS discount and continue participation in BPCI Model 1.

All six Awardees that participated in PHC terminated their Awardee Agreement from BPCI Model 1. Their general reason for termination centered on their actual—or perceived—ability to

[84] The first wave of data collection occurred from April 1, 2013 to December 23, 2013.
[85] As previously noted, cost-savings and measure data from AMS for PQ 1 were finalized by hospitals sometime after PQ 3.

achieve cost savings from care redesigns under BPCI Model 1 or additional cost savings from care redesigns continued from PHC. Effectively, they believed that they had achieved relatively easy efficiency gains in care delivery ("low-hanging fruit") under PHC and would not see further cost savings in time to offset the IPPS discount[86] or would not see cost savings at all.

Looking forward, Awardees explained that enrollment in BPCI Model 1 allowed them to identify cost-reduction opportunities and inefficiencies in their organizations, and all exiting Awardees asserted that they would continue care redesigns proposed under this model.

III.A.3.2. Specific Rationale for Termination

III.A.3.2.1. Organizational and Internal Structural Changes

The clinical staffing structure across Awardee hospitals varied. Under this model, differentials in physician affiliations with Awardees influenced Awardee termination decisions.

Physicians practicing at Awardee hospitals were either employed by the Awardee or practiced at that hospital on a voluntary basis[87]. Few Awardees maintain an all employed or all voluntary physician staff model, most have a mix of employed and voluntary physicians.. As mentioned previously, employed physicians are reportedly already invested in hospital care redesigns and are less incentivized by gainsharing.

For example, Cooper recently increased its employed physician staff and decreased its voluntary physician staff. Consequently, this Awardee decreased its expected reliance on gainsharing to align its physicians to its goals.

Morristown noted that its physicians were not complying with some of the care redesigns proposed under the program. As a result, the hospital had difficulty in progressing in its care redesigns. One leader specifically noted challenges getting medical staff to discharge patients earlier.

Two other Awardees, Saint Michael's and Overlook, also reported issues with physician engagement and enrollment that hindered implementation of the model care redesigns. For example, administrators at Saint Michael's reported that only 13 of the 129 physicians eligible to enroll in the model did so. Admittedly, administrators had not attempted to recruit any new physicians during the third and fourth quarters. This Awardee had been undergoing a change in ownership over the past year, which reportedly discouraged physician engagement and stalled recruitment efforts. These administrators expected increased engagement and enrollment once the Awardee solidified its new ownership. This Awardee terminated its Awardee Agreement with CMS after PQ 5.

Overlook administrators reported physician willingness to enroll in the model. However, enrolled surgeons at this Awardee hospital were skeptical of the value of BPCI Model 1 and noted that the feedback offered during meetings with leadership failed to result in action. Overlook participated in a prior CMS gainsharing demonstration.

[86] PHC did not have an IPPS discount and relied on achieving budget neutrality instead.
[87] This would include physicians that had, or were part of private practice groups.

III.A.3.2.2. Finances, Timing, and the IPPS Discount

All exiting Awardees reported the IPPS discount as a major factor in their decision to terminate their Awardee Agreement.

Awardees exiting within the first year of the BPCI Model 1 initiative were primarily concerned with their *potential* to offset lost revenue from the IPPS discount with cost savings from their care redesigns. These care redesigns varied in time and scope for implementation, which was reported to take 3 to 6 months. As reported by the initial exiting Awardees, this led to an expectation that the IPPS discount would be out of sync with their expected timeline for cost savings. That is, as the IPPS discount increased from the grace period (0 percent) to 0.5 percent and 1 percent, Awardees would incur revenue losses without reaping potential cost savings from their care redesigns. Further, in some cases, Awardees had not yet fully implemented all of their care redesigns.

Awardees exiting in PQs 4 and 5 cited concrete concerns, as they realized their inability to offset the IPPS discount. These hospitals were active for approximately 1 year in the BPCI Model 1 initiative. Their realized cost savings had not offset the overhead costs for model implementation and the IPPS discount (now frozen at 1 percent). Here are some examples of this unmet offset for the recent exiting Awardees:

CentraState cited that the overhead costs for continuing in BPCI Model 1 were too burdensome for the hospital. Specifically, it noted that it had achieved the "low-hanging fruit" efficiency gains in care delivery and would need additional financial resources and time to improve further. This Awardee attributed approximately "70 percent of its decision" to the IPPS discount and associated Medicare rules. Cooper administrators also cited that continuation in the program would require higher levels of effort and resources to realize the benefits expected from care redesigns under the model. As with other Awardee hospitals, Cooper's leadership initially believed the costs associated with the program could be offset, but the hospital was unable to recover IPPS reductions through cost savings or other increases in revenue and had low cost-saving projections.

Cooper and CentraState—along with Jersey Shore, Valley, and JFK—reported that the direct costs for implementing BPCI Model 1 would equal roughly more than a million dollars, and IPPS discounts would be close to an additional million for each Awardee. As a result, these financial risks were considered a detriment to the Awardee's ability to recoup savings. Jersey Shore was also facing budget cuts, staff reductions, and flat admission rates. This hospital, an Awardee that had previously participated in the PHC, found that it was unable to offset the revenue lost from the IPPS discount through cost savings. Under PHC, Jersey Shore reportedly saved approximately $10 million, with an initial investment of approximately $3 million.

Hunterdon estimated the IPPS discount to be between $250,000 and $280,000, which was not deemed viable, as this Awardee stated it was having a difficult year financially.

Deborah also did not experience savings that could be attributed to BPCI Model 1. This Awardee learned about the IPPS freeze at 1 percent at the time of its decision to terminate and reported

being at the cusp of breaking even at 1 percent. The IPPS freeze did not change its termination decision.

III.A.3.2.3. Other Initiatives

A majority of Awardee hospitals were involved in other CMS initiatives. Some of these other programs included CMS' Hospital Value-Based Purchasing, Horizon's Reward Leap Frog Program for Quality Improvement, and the Partnership for Patients (New Jersey Hospital Engagement Network). Many Awardees reported difficulties attributing success in cost savings or quality improvement solely to BPCI Model 1 due to the concurrent initiatives. Attribution of outcomes and cost savings was especially difficult with care redesigns (or their objectives) that overlapped with other initiatives. For example, Awardees could not parse out the effect of care redesigns under BPCI Model 1 affecting readmissions from the effect from CMS' Hospital Readmission Reduction Program.

III.A.3.2.4. Requirements of the Model 1 Program

CentraState terminated its Awardee Agreement prior to CMS removing the requirement for the B-CARE Tool for BPCI Model 1. Leadership reported the duplicative nature of the data collection activities associated with the tool as an issue. This Awardee was frustrated because the majority of information was already collected through its patient records, yet the tool would be unable to pull that information directly from those records. Therefore, the Awardee foresaw the need to hire an additional full-time employee solely for the purpose of data entry.

Hunterdon reported having to pay a data fee of $125,000, a prime factor in decision to terminate its participation in Model 1.

Morristown believed that the additional administrative duties to implement and manage BPCI Model 1 were too great a burden for its current staff. In addition to routine data collection, this Awardee cited that loading data into physician portals and reporting requirements to CMS were onerous tasks and the hospital did not hire additional staff for these additional responsibilities.

III.A.4. Analysis of Implementation and Organizational Responses Domain Findings

At the inception of Model 1, Awardee hospitals included 23 hospitals located in New Jersey that voluntarily participated in the program with the FC. One additional hospital in Kansas initiated an Awardee Agreement with CMS for the model in January 2014. Over the first 5 PQs, 9 Awardee hospitals terminated their Awardee Agreement with CMS, bringing the number of Awardee hospitals down to 15. Results from these exit interviews indicated that hospitals terminated for the following reasons:

1. Hospital physician employment structure already required/motivated physician adherence to protocols and participation in care redesigns through existing employment/contractual relationships.
2. Inability to associate perceived or realized internal hospital cost savings from care redesigns. Relatively long timescales involved in realizing potential internal hospital cost savings from care redesigns.

3. Lack of physician enrollment and engagement resulting in insufficient critical mass to generate the change in practice required to deliver cost savings from care redesigns.

Awardee hospitals enrolled and engaged physicians to assist in their Model 1 activities (e.g., care redesigns), designed to achieve model goals. Awardee hospitals employed various recruiting techniques to engage physicians but have been hindered, in part due to physician skepticism of gainsharing, misunderstanding of model components, or added effort required for model (reporting) requirements. Increased physician enrollment and engagement was noted after the initial gainsharing distribution (end of PQ 3 through PQ 5).

Each Awardee hospital implemented between two and nine care redesigns. Across Awardees, these care redesigns varied in type (e.g., patient flow improvement versus reduction in 30-day readmissions), ease of implementation, timing of implementation, definition of "fully implemented," and desired outcome(s) from implementation. Reports on care redesign implementation stated implementation timeframes spanning from 3 to 6 (or even 9) months. Early in the series of interviews and focus group sessions (e.g., the start of program year 1), administrative and clinical respondents emphasized that these care redesigns aimed to affect quality first and cost of providing care second, as efficient and efficacious care should reduce cost. Recently in PQs 4 and 5, administrators appeared more cost conscious by discussing whether the care redesigns induced sufficient cost savings to validate a business case for continuing in the program, where they face an across-the-board IPPS discount.

III.B. BPCI Model 1 Episode Characteristics and Impacts

Section III.A discussed common findings from hospital administration and clinician perspectives collected on BPCI Model 1. These findings primarily centered on internal activities and progress under Model 1. Section III.B provides insight into how these activities translated to Model 1 goals such as reducing health care expenditures and utilization while maintaining or improving patient quality of care.

First, patient and episode characteristics such as age and an episode's clinical severity were examined between BPCI Model 1 Awardees and comparison hospitals. Then, model impacts on Medicare payment, resource utilization, and health care outcome measures were assessed.

III.B.1. Episode Case Mix and Patient Characteristics

BPCI Model 1 is designed to potentially improve quality of care and contain health care costs. Many health care services are rendered during a patient's hospital stay. Coupled with MS-DRG-based payments hospitals receive for these services and Model 1's aim of reducing hospital (and Medicare) expenditures for care could produce unintended consequences that may negatively impact care quality. These unintended consequences may stem from or manifest as propitious patient selection (e.g., hospital cherry picking or servicing less sick patients) or stinting of care.[88]

[88] Care stinting through decreased care utilization (e.g., to reduce costs associated with care) is examined in the Health Care Outcomes and Resource Utilization domain.

This section assesses patient populations of Awardee and comparison hospital and provides initial insight on whether cherry picking, the action of selecting more potentially profitable patients, has occurred as Awardees attempt to meet model goals. Further, the selective overview of Awardee patient characteristics will contextualize subsequent impact estimates.

Various measures are presented for multiple cohorts: all BPCI hospital Awardees (Full cohort), active Awardees (Active cohort), exiting Awardees (Exiting cohort), Awardees designated as having expansive care redesigns (Expansive cohort), and Awardees designated as having targeted care redesigns (Targeted cohort). As comparison hospitals were statistically matched to individual model Awardees, these hospitals are included/compared within a given cohort to their matched Awardee hospitals.

Below, a cross-measure overview of results and implications for all measures in this domain is presented. Further, in-text examination of the following measures is presented in the remainder of this section:

- Patient Age.
- Patient CMS-HCC Score.[89]
- Episode MS-DRG Weight.

Appendix B presents all data for these measures. Additionally, Awardee-level data may be presented to examine if particular Awardee hospitals within a cohort are driving results.

Measures in this domain are presented as continuous variables that capture select patient characteristics. Weighted and unweighted unadjusted measure rates (Tables 13, 14, and 15) assess actual cohort-level data for the Full, Active, and Exiting cohorts (other cohorts are left to Appendix B), and Figure 6 provides cohort time trends for PHC, Expansive, and Targeted cohorts.

[89] CMS-HCC scores for 2014 were unavailable when claims data for this report were pulled from the CCW. For patient claims analyzed over PQs 4 and 5, the last known CMS-HCC score for that patient is taken. In 2014, 5.3 percent of inpatient stays analyzed had no CMS-HCC scores since 2010; individuals for these records were given missing CMS-HCC scores.

Table 13: Patient and Episode Characteristics – Full Cohort

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	Patient Age								
BPCI	**Mean**	76.63	76.08	76.42***	76.09	76.42***	76.11	76.42***	76.00
	95% CI	[76.58, 76.67]	[75.31, 76.85]	[76.36, 76.49]	[75.33, 76.85]	[76.35, 76.49]	[75.35, 76.87]	[76.28, 76.56]	[75.24, 76.76]
	N	282,896		149,233		118,799		30,434	
Comparison	**Mean**	75.18	74.70	74.87***	74.61***	74.85***	74.63***	74.95***	74.53***
	95% CI	[75.15, 75.20]	[73.81, 75.59]	[74.83, 74.91]	[73.77, 75.45]	[74.81, 74.89]	[73.81, 75.45]	[74.87, 75.04]	[73.63, 75.42]
	N	892,495		464,815		372,482		92,333	
	CMS-HCC Score								
BPCI	**Mean**	1.72	1.69	1.64***	1.64***	1.66***	1.67***	1.55***	1.53***
	95% CI	[1.72, 1.73]	[1.61, 1.77]	[1.63, 1.64]	[1.56, 1.72]	[1.65, 1.67]	[1.59, 1.75]	[1.54, 1.56]	[1.46, 1.61]
	N	282,896		149,233		118,799		30,434	
Comparison	**Mean**	1.68	1.62	1.61***	1.60***	1.63***	1.62	1.54***	1.51***
	95% CI	[1.67, 1.68]	[1.53, 1.70]	[1.61, 1.62]	[1.51, 1.68]	[1.63, 1.63]	[1.53, 1.70]	[1.53, 1.55]	[1.43, 1.60]
	N	892,497		464,815		372,482		92,333	
	Episode MS-DRG Weight								
BPCI	**Mean**	1.58	1.58	1.62***	1.59***	1.62***	1.58	1.64***	1.62***
	95% CI	[1.57, 1.59]	[1.50, 1.65]	[1.62, 1.63]	[1.50, 1.67]	[1.61, 1.63]	[1.50, 1.66]	[1.63, 1.66]	[1.54, 1.69]
	N	282,896		149,233		118,799		30,434	
Comparison	**Mean**	1.62	1.62	1.66***	1.65***	1.66***	1.65***	1.66***	1.67***
	95% CI	[1.61, 1.62]	[1.53, 1.71]	[1.65, 1.66]	[1.57, 1.74]	[1.65, 1.66]	[1.56, 1.74]	[1.65, 1.67]	[1.58, 1.76]
	N	892,497		464,815		372,482		92,333	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline within adjusted or unadjusted statistics. Table 10 details cohort membership.

Table 14: Patient and Episode Characteristics – Active Cohort

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	Patient Age								
BPCI	**Mean**	76.23	75.84	76.00***	75.88	76.01***	75.92	75.97**	75.75
	95% CI	[76.16, 76.29]	[74.99, 76.70]	[75.91, 76.09]	[75.04, 76.73]	[75.90, 76.11]	[75.08, 76.76]	[75.77, 76.18]	[74.90, 76.59]
	N	144,068		74,042		59,129		14,913	
Comparison	**Mean**	74.80	74.35	74.53***	74.38	74.53***	74.44***	74.56***	74.16***
	95% CI	[74.76, 74.83]	[73.42, 75.29]	[74.49, 74.58]	[73.54, 75.23]	[74.47, 74.58]	[73.63, 75.26]	[74.45, 74.66]	[73.22, 75.10]
	N	571,049		298,017		238,789		59,228	
	CMS-HCC Score								
BPCI	**Mean**	1.76	1.68	1.67***	1.66***	1.70***	1.69	1.58***	1.53***
	95% CI	[1.75, 1.76]	[1.60, 1.77]	[1.66, 1.68]	[1.57, 1.74]	[1.69, 1.71]	[1.60, 1.78]	[1.56, 1.60]	[1.45, 1.61]
	N	144,068		74,042		59,129		14,913	
Comparison	**Mean**	1.70	1.62	1.64***	1.62	1.65***	1.64***	1.57***	1.53***
	95% CI	[1.70, 1.71]	[1.53, 1.70]	[1.63, 1.64]	[1.53, 1.70]	[1.65, 1.66]	[1.56, 1.72]	[1.56, 1.58]	[1.44, 1.62]
	N	571,050		298,017		238,789		59,228	
	Episode MS-DRG Weight								
BPCI	**Mean**	1.49	1.51	1.51***	1.49***	1.50**	1.48***	1.51**	1.52*
	95% CI	[1.48, 1.50]	[1.43, 1.58]	[1.50, 1.52]	[1.41, 1.57]	[1.49, 1.51]	[1.40, 1.56]	[1.49, 1.54]	[1.44, 1.60]
	N	144,068		74,042		59,129		14,913	
Comparison	**Mean**	1.61	1.61	1.64***	1.63***	1.64***	1.62***	1.64***	1.64***
	95% CI	[1.60, 1.61]	[1.52, 1.70]	[1.64, 1.65]	[1.54, 1.71]	[1.64, 1.65]	[1.54, 1.71]	[1.63, 1.65]	[1.55, 1.73]
	N	571,050		298,017		238,789		59,228	

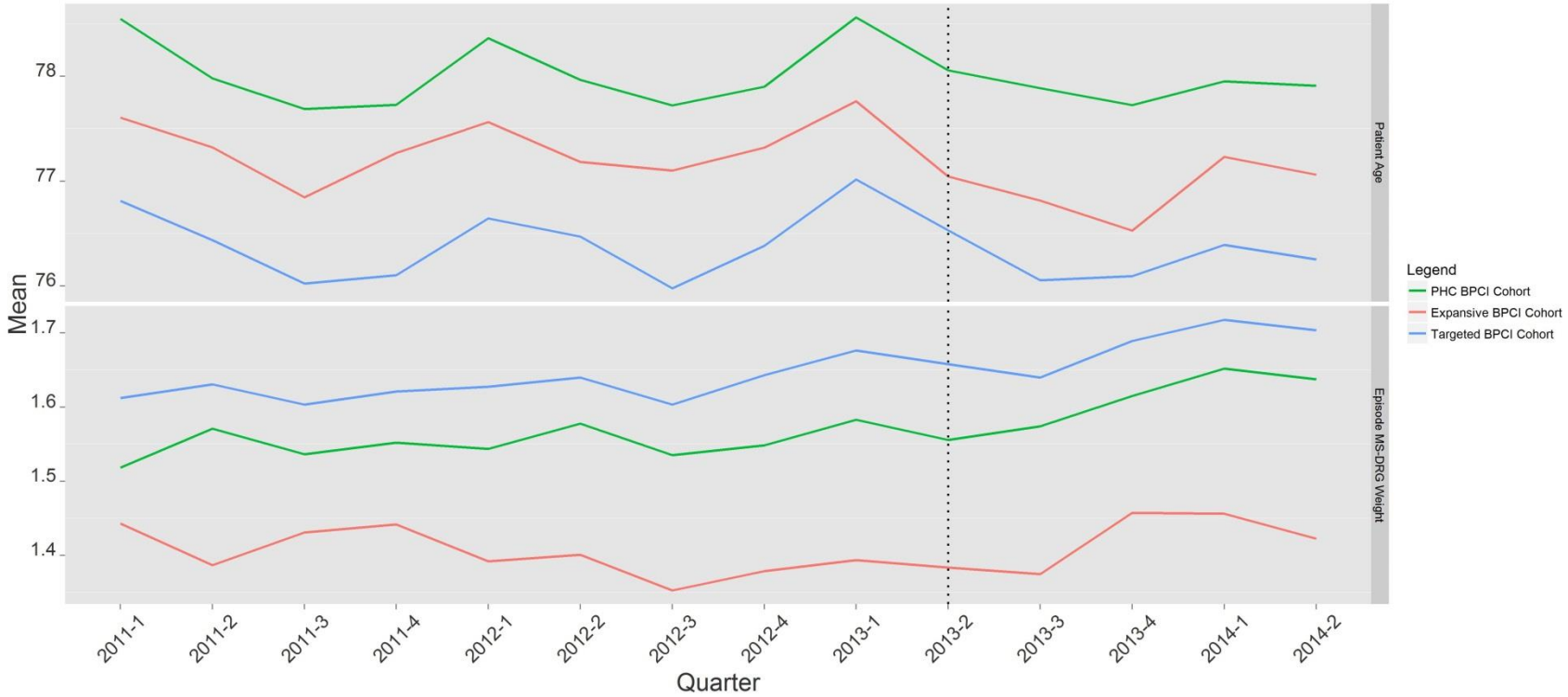
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline within adjusted or unadjusted statistics. Table 10 details cohort membership.

Table 15: Patient and Episode Characteristics – Exiting Cohort

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted	Weighted	Unweighted
	Patient Age								
BPCI	**Mean**	77.04	76.48	76.84***	76.42	76.84***	76.41	76.85*	76.42
	95% CI	[76.98, 77.11]	[75.86, 77.11]	[76.75, 76.93]	[75.78, 77.05]	[76.74, 76.93]	[75.78, 77.05]	[76.67, 77.04]	[75.81, 77.04]
	N	138,828		75,191		59,670		15,521	
Comparison	**Mean**	75.11	74.91	74.77***	74.58***	74.73***	74.54***	74.92***	74.76**
	95% CI	[75.07, 75.15]	[74.14, 75.69]	[74.71, 74.82]	[73.79, 75.38]	[74.67, 74.79]	[73.74, 75.35]	[74.80, 75.05]	[73.98, 75.54]
	N	403,965		209,467		167,988		41,479	
	CMS-HCC Score								
BPCI	**Mean**	1.68	1.70	1.60***	1.62***	1.62***	1.64***	1.52***	1.53***
	95% CI	[1.68, 1.69]	[1.63, 1.77]	[1.59, 1.61]	[1.55, 1.69]	[1.61, 1.63]	[1.57, 1.71]	[1.50, 1.54]	[1.47, 1.60]
	N	138,828		75,191		59,670		15,521	
Comparison	**Mean**	1.63	1.62	1.57***	1.56***	1.59***	1.58***	1.50***	1.48***
	95% CI	[1.63, 1.63]	[1.54, 1.70]	[1.56, 1.58]	[1.48, 1.64]	[1.58, 1.59]	[1.50, 1.66]	[1.49, 1.51]	[1.41, 1.56]
	N	403,966		209,467		167,988		41,479	
	Episode MS-DRG Weight								
BPCI	**Mean**	1.67	1.70	1.74***	1.74***	1.73***	1.74***	1.77***	1.77***
	95% CI	[1.67, 1.68]	[1.61, 1.78]	[1.73, 1.75]	[1.66, 1.82]	[1.72, 1.75]	[1.65, 1.82]	[1.74, 1.80]	[1.69, 1.85]
	N	138,828		75,191		59,670		15,521	
Comparison	**Mean**	1.66	1.66	1.70***	1.71***	1.70***	1.71***	1.70***	1.73***
	95% CI	[1.65, 1.66]	[1.57, 1.76]	[1.69, 1.71]	[1.62, 1.81]	[1.69, 1.71]	[1.62, 1.80]	[1.69, 1.72]	[1.63, 1.82]
	N	403,966		209,467		167,988		41,479	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline within adjusted or unadjusted statistics. Table 10 details cohort membership.

Figure 6: Unadjusted Quarterly Trends for Patient Age, Episode MS-DRG Weight, and CMS-HCC Score for PHC, Expansive, and Targeted Cohorts



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

III.B.1.1. Overview of Results

Table 13 indicates that there were statistically significant changes in case mix over time for all three measures (patient age, CMS-HCC score, and MS-DRG weight) among the Full cohort Awardee and comparison hospitals. The weighted results (weighted by the number of episodes at each hospital) show that among Awardee hospitals, average patient age decreased by approximately 2.5 months between Baseline and Since BPCI Inception periods. Average patient age across comparison hospitals decreased by 3.7 months over this time period. The unweighted age measure, which gives each hospital equal weight regardless of the number of episodes, did not display any statistically significant changes for Awardee hospitals but did display statistically significant changes for comparison hospitals. In addition to decreases in age, Full cohort Awardee and comparison hospitals displayed comparable decreases in average patient CMS-HCC score and 0.04-point increases in episode MS-DRG weight between Baseline and Since BPCI Inception periods.

Table 14 provides the same statistics as Table 13 but for the Active cohort. Full cohort decreases in patient age and CMS-HCC score, and increases in MS-DRG weight are quantitatively similar for the Active cohort. Among the active Awardees and their comparison cohort, beneficiaries tend to be younger and have lower MS-DRG weights than the Full BPCI cohort. The Active cohort has higher CMS-HCC scores than the Full cohort over the Since BPCI Inception period.

Table 15 presents these characteristics for the Exiting cohort (Awardee and comparison) and generally exhibits similar changes between Baseline and Since BPCI Inception. The Exiting cohort Awardees and comparisons tend to have higher MS-DRG weights and lower CMS-HCC scores than Full cohort Awardees and comparisons. Exiting cohort Awardee beneficiaries tend to be older than Full cohort Awardee beneficiaries.

Figure 6 provides similar information for the PHC, Expansive, and Targeted subcohorts. Longitudinal patterns across these subgroups are similar, though the levels of the measures differ among cohorts. Generally, Awardee hospitals in the Expansive cohort serve an older population but have lower episode MS-DRG weights for those patients over time, relative to Awardees in the Targeted cohort.

III.B.1.2. Analysis of Case Mix and Patient Characteristics Domain Findings

Full cohort statistics in Table 13 indicate that BPCI Awardee hospitals serviced a slightly younger Medicare population since Model 1 inception. As expected, this was accompanied by a decrease in average CMS-HCC score. At the same time, however, it was found that MS-DRG weight increased despite decreases in both age and CMS-HCC score. This indicates that, despite their younger age and lower chronic disease burden, patients received more resource-intensive care as implied by higher payment MS-DRGs. Holding patient acuity constant, there *may* be an incentive to "up code" inpatients so that they are more likely to fall into a higher MS-DRG category and yield higher payments. However, these increases in episode MS-DRG weights may not be indicative of up-coding practices, as comparison hospitals also exhibit increases in this measure. Tables 14 and 15 provide similar findings for the Active and Exiting cohorts.

III.B.2. Health Care Outcomes and Resource Utilization

BPCI Model 1 is designed to incentivize hospitals and physicians to lower health care expenditures while maintaining and/or improving quality of care. Therefore, physicians and hospitals are expected to increase care coordination to provide more efficient and potentially improved care However, unintended consequences, such as care stinting, may manifest in attempts to engender efficiency gains and negatively impact care quality.

The model's effect on quality and utilization will be tested through various measures. In this report, the following health care outcome and resource utilization measures are presented over two time periods: during and up to 30 days after an episode:

- 30-day all-cause mortality[90].
- 30-day all-cause readmissions.
- Episode ICU utilization.
- Episode length of stay.

The length of stay measure is presented in units of days, while mortality, readmission, and ICU utilization are presented as means (as percentages). Additionally, since DiD impact analyses are conducted across Awardee and comparison hospitals at the episode level, mortality, readmission, and ICU use impact estimates are presented in terms of odds ratio (ORs) that indicate the *likelihood* of a mortality, readmission, or ICU event for BPCI Model 1 patients relative to Baseline and comparison hospital patients. Appendix A presents detailed specifications for these measures. Appendix B presents data for these and related (e.g., condition-specific mortality) measures.

The remainder of Section III.B.2 first provides a high-level overview of measures in the Health Care Outcomes and Resource Utilization domain (Section III.B.2.1) and then presents individual measure result overviews, analysis, and data (Sections III.B.2.1.1–III.B.2.1.4). Findings focus on changes from Baseline to Since BPCI Inception periods for unadjusted and adjusted statistics and DiD impact estimates. Additionally, Awardee-level data may be presented to examine whether particular hospitals within a cohort are driving results.

[90] This measure may span the during and up to 30 days after an episode time periods as the mortality measure begins from admission and may include in-hospital or post-discharge mortality.

III.B.2.1. Domain-Wide Overview of Results

Generally, there were no statistically significant changes in mortality rates from Baseline to the Since BPCI Inception period for Awardee and comparison hospital cohorts. DiD impact analyses exhibited no program-wide[91] statistically significant changes in the likelihood of a patient experiencing a mortality event within 30 days over the model performance period, relative to Baseline and patients discharged from comparison hospitals. Subcohort analyses, however, exhibited an increased impact estimate among Active cohort Awardees, stemming from an elevated PQ 4 likelihood estimate (1.09 OR, $p < 0.05$). Examination of Active Awardee-level impact estimates[92] indicate that two-thirds of the Active cohort Awardees had elevated mortality impact estimates (i.e., OR above 1.0), but only two Active Awardee estimates were statistically significant at the 10-percent level. Additional analyses indicate that these elevated mortality impact estimates were driven by post-episode mortality, not in-hospital (i.e., episode) mortality.

Changes in ICU use from Baseline to the Since BPCI Inception period varied in direction and magnitude across Awardee and comparison hospital cohorts. Program-wide DiD impact analyses exhibited a statistically significant increase in the likelihood of a patient having an ICU stay at an Awardee hospital relative to Baseline and comparison hospitals (1.10 OR, $p < 0.01$). This result appears to be driven by Active cohort Awardee hospitals, with an OR impact estimate of 1.20 ($p < 0.01$). This cohort was previously noted to have a higher average (CMS-HCC) risk profile relative to Exiting cohort Awardees. Hospital-level DiD impact analyses indicate that all but three of the Active Awardees exhibited increased likelihoods of their patients having an ICU stay during their episode (OR above 1.0). However, the attribution of these elevated ICU estimates to Model 1 must be qualified. Placebo DiD tests[93] for the Full and Active cohorts indicate that these elevated ICU likelihood estimates were found *before* BPCI Model 1 implementation and may consequently not be attributable to Model 1.

Changes in episode length of stay from Baseline to the Since BPCI Inception period varied in direction and magnitude across Awardee and comparison hospital cohorts but were generally minimal with differences less than 0.5 days. There were no program-wide DiD impact estimates in this measure. However, hospital-level impact analyses indicate that 11 of the 24 hospital Awardees did achieve statistically significant decreases in episode length of stay ranging from decreases of 0.22 to 0.62 days (all statistically significant, $p < 0.01$). Of these 11 Awardees, 8 are in the Active cohort, and 5 of these 8 comprise the majority of the Expansive cohort.

[91] Full cohort analysis.
[92] An individual Awardee hospital performance compared to its four, matched comparison hospitals.
[93] Placebo DiD Tests test the reliability of associating DiD impact estimates with an intervention (e.g., BPCI Model 1) by assuming placebo/pseudo start dates *before* the intervention's *actual* start date. Statistically significant estimates in the placebo periods can indicate an inappropriateness of impact estimate attribution to BPCI Model 1.

III.B.2.1.1. 30-Day All-Cause Mortality

III.B.2.1.1.1. Overview of Results

Table 16 presents 30-day all-cause mortality as unadjusted and adjusted rates for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5 periods. This table also includes these statistics for Active and Exiting cohorts. Among the Full cohort, the unadjusted 30-day all-cause mortality rates increased by 0.08 and 0.02 percentage points for Awardee and comparison hospitals, respectively (Since BPCI Inception vs. Baseline). After adjusting for patient and episode characteristics, the 30-day all-cause mortality rates increased by 0.18 and 0.1 percentage points for Awardee and comparison hospitals, respectively. The unadjusted rate increased by 0.19 percentage points for Active cohort Awardee hospitals and a decrease of 0.02 percentage points was noted among the comparison hospitals. The adjusted rates increased by 0.34 percentage points for Active cohort Awardee hospitals while the adjusted rates increased by 0.06 percentage points for the comparison cohort. Unadjusted rates decreased by 0.03 percentage points among Exiting cohort Awardee hospitals while the rates increased by 0.04 percentage points among the comparison cohort. Adjusted, the rates increased by 0.04 and 0.13 for Awardee and comparison hospitals, respectively.

Table 17 presents unadjusted, unweighted 30-day all-cause mortality statistics for Full, Active, and Exiting cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted rates in the Full and Active BPCI cohorts, as these estimates differ from unadjusted, weighted statistics presented in Table 16. Hospital specific unadjusted mortality rates (not shown) indicates that the lower rates in Table 17 are driven down by certain specialty hospitals (e.g., KSRC and its comparison hospitals), which are expected to have low mortality (and readmission) events due to the nature of their specializations[94] and small patient population.

Table 18 presents Full, Active, and Exiting cohort DiD estimates. These impact estimates assess the likelihood of patients experiencing a mortality event within 30 days during the Since BPCI Inception period, relative to the Baseline period and similar timeframes for comparison hospitals. The Full cohort estimate is a non-statistically significant OR of 1.01. Subcohort analysis shows that the estimate is driven by Active cohort Awardees (1.04, $p < 0.05$), while Exiting cohort Awardees had no statistically significant impact. Figure 7 breaks these estimates down by PQ and show that within the Since BPCI Inception period, Active Awardee mortality estimates were significantly elevated in PQ 4, relative to comparisons.

Table 19 presents 30-day all-cause mortality as unadjusted and adjusted rates for PHC, Expansive, and Targeted cohorts over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. Generally, Awardee and comparison hospitals across these cohorts show increases in average 30-day all-cause mortality in the Since BPCI Inception period relative to Baseline for Expansive and Targeted cohorts but not the PHC cohort. Table 20 presents unadjusted, unweighted 30-day all-cause mortality as unadjusted and adjusted rates for Expansive and Targeted Care Redesign and PHC cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted rates in the Targeted cohorts, as these estimates differ from unadjusted, weighted statistics presented in Table 19. Table 21 presents DiD impact estimates for these

[94] For example, KSRC focuses on orthopedic procedures and other outpatient surgical care with inpatient beds for extended stays as needed.

cohorts. The Expansive cohort's odds ratio was 1.05 ($p < 0.1$), while the Targeted and PHC cohorts show no statistically significant impact on a patient's likelihood for a mortality event. This result is not surprising, as the majority of the Expansive cohort is composed of Active Awardee hospitals (and their comparisons).

III.B.2.1.1.2. Analysis of Measure Findings – Mortality

BPCI Model 1 may provide complex incentives directly related to mortality. All else being equal, model incentives to improve care coordination would likely reduce mortality rates. However, Awardee incentive to reduce internal costs could result in care stinting and early discharge that may lead to adverse outcomes such as mortality. Further, as an outcome that extends beyond the episode period, mortality events occurring after a patient is discharged from the acute-care setting may not be directly under Awardee control.

The Full cohort DiD estimate exhibits no statistically significant changes in 30-day all-cause mortality rates. In other words, when compared to comparison hospitals, BPCI Model 1 did not have a statistically significant impact on this outcome. This result suggests that, on average, BPCI hospitals are maintaining quality as measured by the mortality outcome, relative to comparison hospitals. Although there is no statistically significant impact for the entire cohort of Awardee hospitals, differences appear when analyzing subsets of Active and Exiting Awardees. Active Awardee hospitals exhibited a statistically significant increase in the likelihood of a patient experiencing a mortality event within 30 days (relative to comparisons), while Exiting cohort Awardees did not. Examination of the first five PQs shows elevated mortality estimates in PQs 1 and 5 (not statistically significant) and a statistically significant likelihood in PQ 4 (1.09 OR, $p < 0.05$). Table 16 showed that Active cohort Awardee changes between Baseline and Since BPCI Inception outpaced changes in Active cohort comparison hospitals.[95] When hospital specific impact estimates (i.e., individual Awardees relative to their 4 matched comparison hospitals) are examined, 10 of the Active Awardees have increased likelihood estimates for patient mortality; that is, they have OR impact estimates above 1.0. However, only 2 of these 10 hospitals (Awardees Inspira Vineland and St. Joseph's) have statistically significant, increased likelihood impacts ($p < 0.1$). Of these 10 Active Awardees, 6 are in the Expansive cohort, contributing to the similar statistically significant impact noted in Table 21. The reasoning for these increased likelihoods is unclear, and it can be further noted that two of the nine Exiting cohort Awardee hospitals also had elevated patient mortality likelihoods.

Additional analyses (not shown) across all Model 1 Awardees indicate that statistically significant impacts in mortality likelihood occur during the post-episode period (i.e., are not attributable as in-hospital mortality events). However, it must be emphasized that this finding is not (statistically) consistent across PQs and does exhibit a decrease in PQ 5 (1.05, no statistically significance). Further, as a post-episode event, factors outside of Awardee control may have affected these outcomes.

[95] Active cohort Awardee hospitals and comparison hospitals had changes in unadjusted means of 0.19 and –0.02 percentage points, respectively and with respective adjusted differences of 0.34 and 0.06 when comparing Since BPCI Inception to Baseline.

Table 16: Episode Weighted, Unadjusted and Adjusted 30-Day All-Cause Mortality for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean (%)**	7.54	8.13	7.62	8.31	7.67	8.35	7.45	8.18
	95% CI	[7.45, 7.64]	[7.82, 8.43]	[7.49, 7.76]	[7.99, 8.63]	[7.52, 7.82]	[8.03, 8.66]	[7.16, 7.75]	[7.86, 8.50]
	N	284,383		150,096		119,578		30,518	
Comparison	**Mean (%)**	7.44	7.23	7.46	7.33	7.49	7.36	7.30	7.21
	95% CI	[7.39, 7.50]	[6.77, 7.69]	[7.38, 7.53]	[6.86, 7.80]	[7.41, 7.58]	[6.90, 7.83]	[7.14, 7.48]	[6.73, 7.69]
	N	890,054		463,915		372,086		91,829	
	Active Cohort								
BPCI	**Mean (%)**	7.42	8.06	7.61	8.40	7.67*	8.44	7.41	8.22
	95% CI	[7.29, 7.56]	[7.61, 8.50]	[7.42, 7.81]	[7.93, 8.87]	[7.45, 7.88]	[7.97, 8.91]	[6.99, 7.84]	[7.76, 8.69]
	N	144,411		74,410		59,560		14,850	
Comparison	**Mean (%)**	7.45	7.26	7.43	7.32	7.48	7.36	7.22**	7.16
	95% CI	[7.38, 7.52]	[6.80, 7.73]	[7.33, 7.52]	[6.85, 7.79]	[7.37, 7.59]	[6.89, 7.83]	[7.01, 7.43]	[6.67, 7.66]
	N	569,740		297,237		238,463		58,774	
	Exiting Cohort								
BPCI	**Mean (%)**	7.66	9.37	7.63	9.41	7.67	9.43	7.49	9.32
	95% CI	[7.52, 7.80]	[8.96, 9.78]	[7.44, 7.82]	[8.99, 9.83]	[7.45, 7.88]	[9.02, 9.85]	[7.08, 7.91]	[8.89, 9.76]
	N	139,972		75,686		60,018		15,668	
Comparison	**Mean (%)**	7.42	6.92	7.46	7.05	7.48	7.06	7.38	6.98
	95% CI	[7.34, 7.50]	[6.50, 7.33]	[7.34, 7.57]	[6.63, 7.47]	[7.35, 7.60]	[6.64, 7.48]	[7.13, 7.64]	[6.55, 7.41]
	N	401,702		208,678		167,376		41,302	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 17: Unweighted, Unadjusted 30-Day All-Cause Mortality for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Full Cohort				
BPCI	**Mean (%)**	7.37	7.50	7.57**	7.25
	95% CI	[5.95, 9.04]	[6.01, 9.26]	[6.07, 9.33]	[5.80, 8.97]
	N	284,383	150,096	119,578	30,518
Comparison	**Mean (%)**	7.37	7.41	7.44	7.25
	95% CI	[5.90, 9.13]	[5.89, 9.24]	[5.92, 9.28]	[5.75, 9.07]
	N	890,054	463,915	372,086	91,829
	Active Cohort				
BPCI	**Mean (%)**	7.33	7.56**	7.61**	7.36
	95% CI	[5.81, 9.12]	[5.95, 9.46]	[6.00, 9.52]	[5.77, 9.25]
	N	144,411	74,410	59,560	14,850
Comparison	**Mean (%)**	7.47	7.46	7.51	7.27*
	95% CI	[6.00, 9.19]	[5.95, 9.24]	[6.00, 9.29]	[5.77, 9.03]
	N	569,740	297,237	238,463	58,774
	Exiting Cohort				
BPCI	**Mean (%)**	7.43	7.41	7.50	7.08*
	95% CI	[6.16, 8.91]	[6.11, 8.94]	[6.17, 9.04]	[5.84, 8.52]
	N	139,972	75,686	60,018	15,668
Comparison	**Mean (%)**	7.21	7.29	7.31	7.19
	95% CI	[5.81, 8.97]	[5.81, 9.14]	[5.83, 9.17]	[5.74, 9.01]
	N	401,702	208,678	167,376	41,302

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

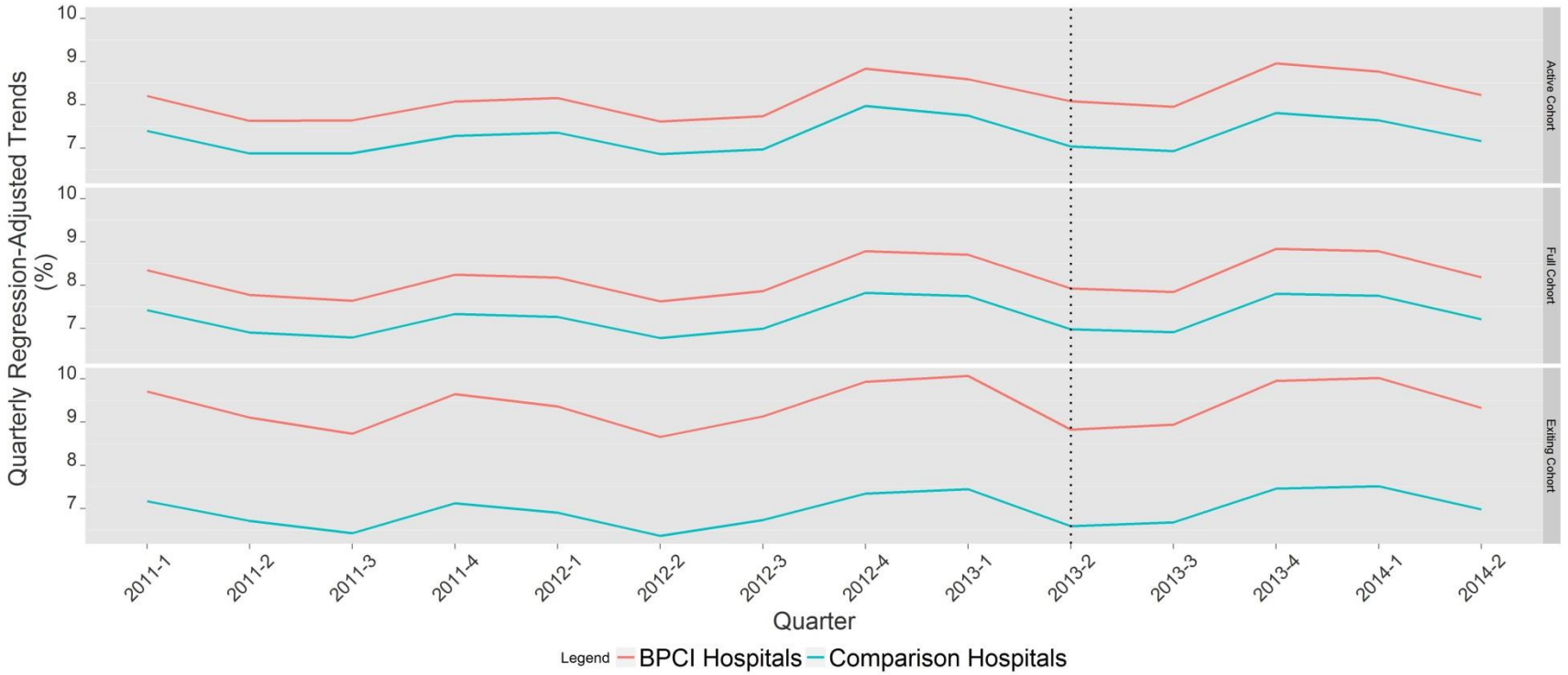
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 18: Since BPCI Inception DiD Estimates for 30-Day All-Cause Mortality for Full, Active, and Exiting Cohort[+]

	Full Cohort	Active Cohort	Exiting Cohort
DiD (Odds Ratio)	1.01	1.04**	0.98
95% CI	(0.98, 1.04)	(1.00, 1.08)	(0.94, 1.03)
Sample Size	1,788,446	1,085,797	826,037

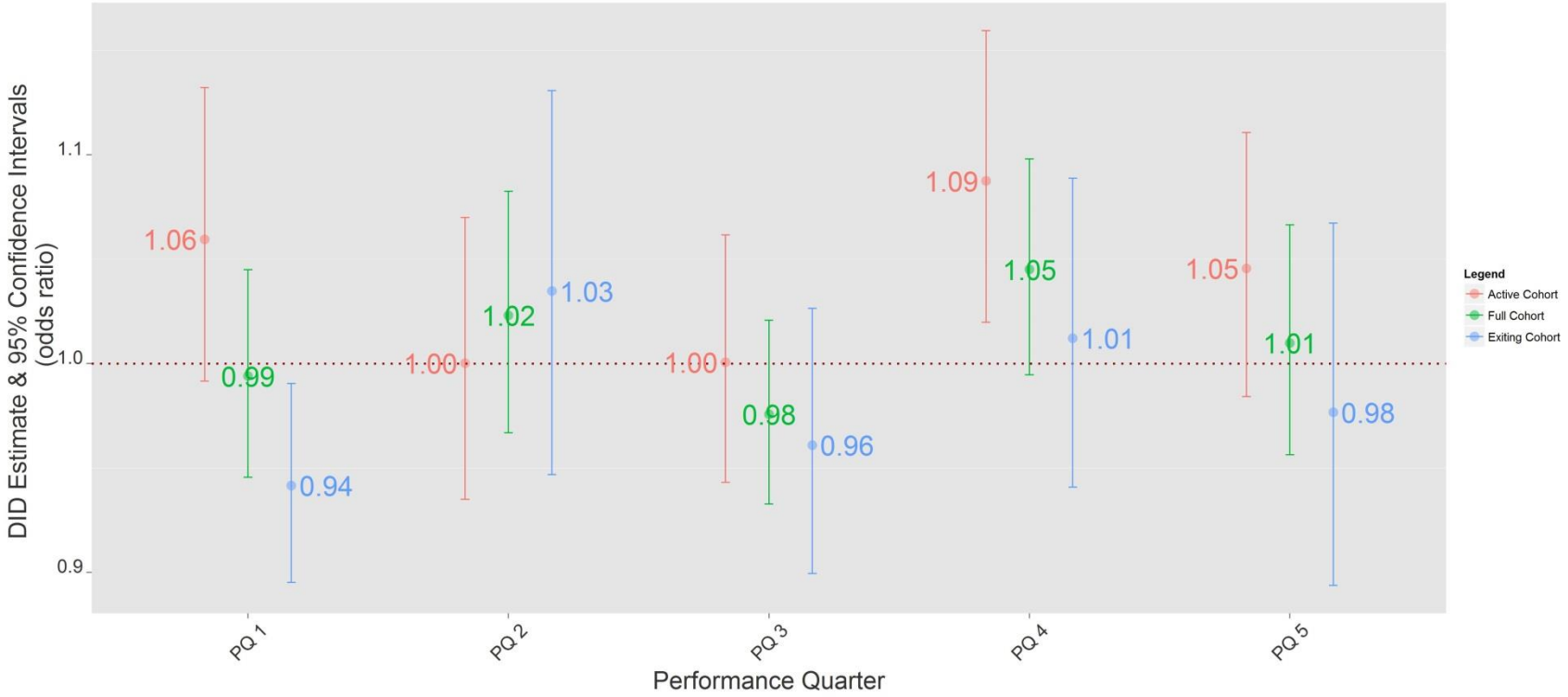
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 7: Adjusted Quarterly Trends for 30-Day All-Cause Mortality for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 8: Performance Quarter DiD Estimates for 30-Day All-Cause Mortality for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 19: Episode Weighted, Unadjusted and Adjusted 30-Day All-Cause Mortality for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Expansive Cohort								
BPCI	**Mean (%)**	7.42	8.28	7.69	8.77	7.72	8.86*	7.57	8.43
	95% CI	[7.22, 7.63]	[7.87, 8.69]	[7.40, 7.99]	[8.32, 9.23]	[7.39, 8.05]	[8.40, 9.32]	[6.93, 8.26]	[8.00, 8.86]
	N	62,147		31,655		25,370		6,285	
Comparison	**Mean (%)**	7.31	7.08	7.37	7.21	7.47*	7.28	6.97**	6.93
	95% CI	[7.21, 7.41]	[6.61, 7.56]	[7.23, 7.51]	[6.73, 7.70]	[7.31, 7.63]	[6.81, 7.76]	[6.66, 7.28]	[6.41, 7.44]
	N	249,646		130,290		104,369		25,921	
	Targeted Cohort								
BPCI	**Mean (%)**	7.57	9.53	7.60	9.66	7.65	9.68	7.42	9.58
	95% CI	[7.47, 7.69]	[9.16, 9.91]	[7.45, 7.76]	[9.27, 10.05]	[7.48, 7.82]	[9.30, 10.06]	[7.09, 7.75]	[9.19, 9.98]
	N	222,236		118,441		94,208		24,233	
Comparison	**Mean (%)**	7.47	6.93	7.49	7.04	7.51	7.06	7.44	6.99
	95% CI	[7.40, 7.53]	[6.55, 7.32]	[7.40, 7.58]	[6.65, 7.43]	[7.41, 7.60]	[6.67, 7.45]	[7.24, 7.64]	[6.59, 7.38]
	N	669,128		348,118		279,356		68,762	
	PHC Cohort								
BPCI	**Mean (%)**	8.22	9.92	8.17	9.91	8.20	9.95	8.05	9.77
	95% CI	[8.05, 8.39]	[9.45, 10.40]	[7.93, 8.41]	[9.43, 10.40]	[7.93, 8.47]	[9.47, 10.43]	[7.53, 8.59]	[9.24, 10.29]
	N	97,506		50,529		40,191		10,338	
Comparison	**Mean (%)**	7.78	7.26	7.76	7.36	7.80	7.39	7.60	7.25
	95% CI	[7.68, 7.88]	[6.80, 7.72]	[7.62, 7.90]	[6.90, 7.82]	[7.64, 7.95]	[6.93, 7.85]	[7.29, 7.92]	[6.77, 7.73]
	N	270,528		140,014		112,220		27,794	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 20: Unweighted, Unadjusted 30-Day All-Cause Mortality for Expansive, Targeted, and PHC Cohorts+

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Expansive Cohort				
BPCI	**Mean (%)**	7.30	7.65*	7.67*	7.58
	95% CI	[5.72, 9.18]	[5.95, 9.68]	[5.97, 9.69]	[5.88, 9.62]
	N	62,147	31,655	25,370	6,285
Comparison	**Mean (%)**	7.33	7.34	7.43	6.97**
	95% CI	[5.79, 9.14]	[5.76, 9.20]	[5.84, 9.30]	[5.43, 8.81]
	N	249,646	130,290	104,369	25,921
	Targeted Cohort				
BPCI	**Mean (%)**	7.40	7.44	7.52	7.10*
	95% CI	[6.05, 8.97]	[6.04, 9.07]	[6.11, 9.17]	[5.76, 8.68]
	N	222,236	118,441	94,208	24,233
Comparison	**Mean (%)**	7.36	7.44	7.45*	7.40
	95% CI	[5.93, 9.09]	[5.95, 9.26]	[5.96, 9.28]	[5.91, 9.21]
	N	669,128	348,118	279,356	68,762
	PHC Cohort				
BPCI	**Mean (%)**	8.25	8.06	8.13	7.81
	95% CI	[6.96, 9.71]	[6.76, 9.55]	[6.81, 9.62]	[6.55, 9.24]
	N	97,506	50,529	40,191	10,338
Comparison	**Mean (%)**	7.90	7.88	7.91	7.78
	95% CI	[6.44, 9.58]	[6.38, 9.62]	[6.41, 9.64]	[6.29, 9.51]
	N	270,528	140,014	112,220	27,794

+ Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort-level statistics consider all hospitals as having equal weight.

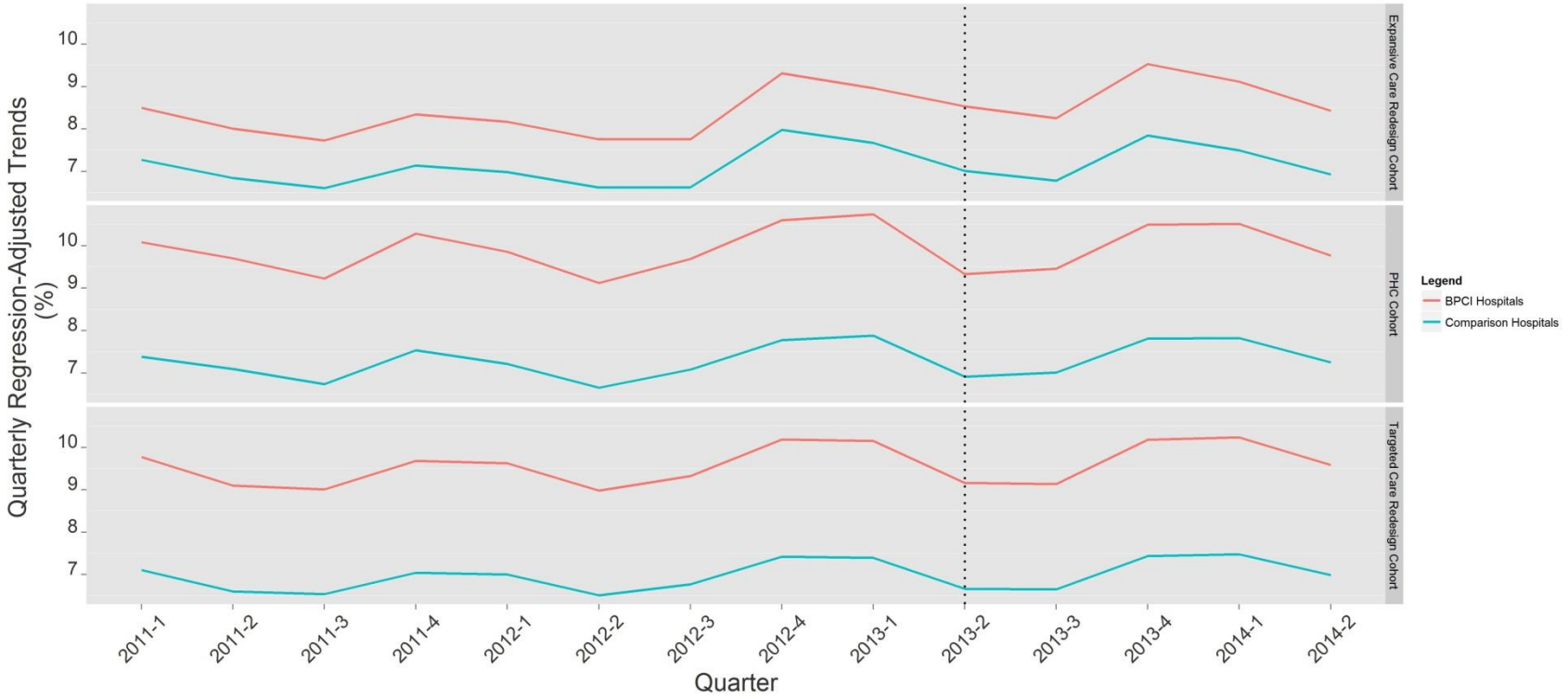
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 21: Since BPCI Inception DiD Estimates for 30-Day All-Cause Mortality for Expansive, Targeted, and PHC Cohorts[+]

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD (Odds Ratio)	0.98	1.05*	1.00
95% CI	(0.93, 1.04)	(1.00, 1.10)	(0.97, 1.03)
Sample Size	558,576	473,738	1,357,921

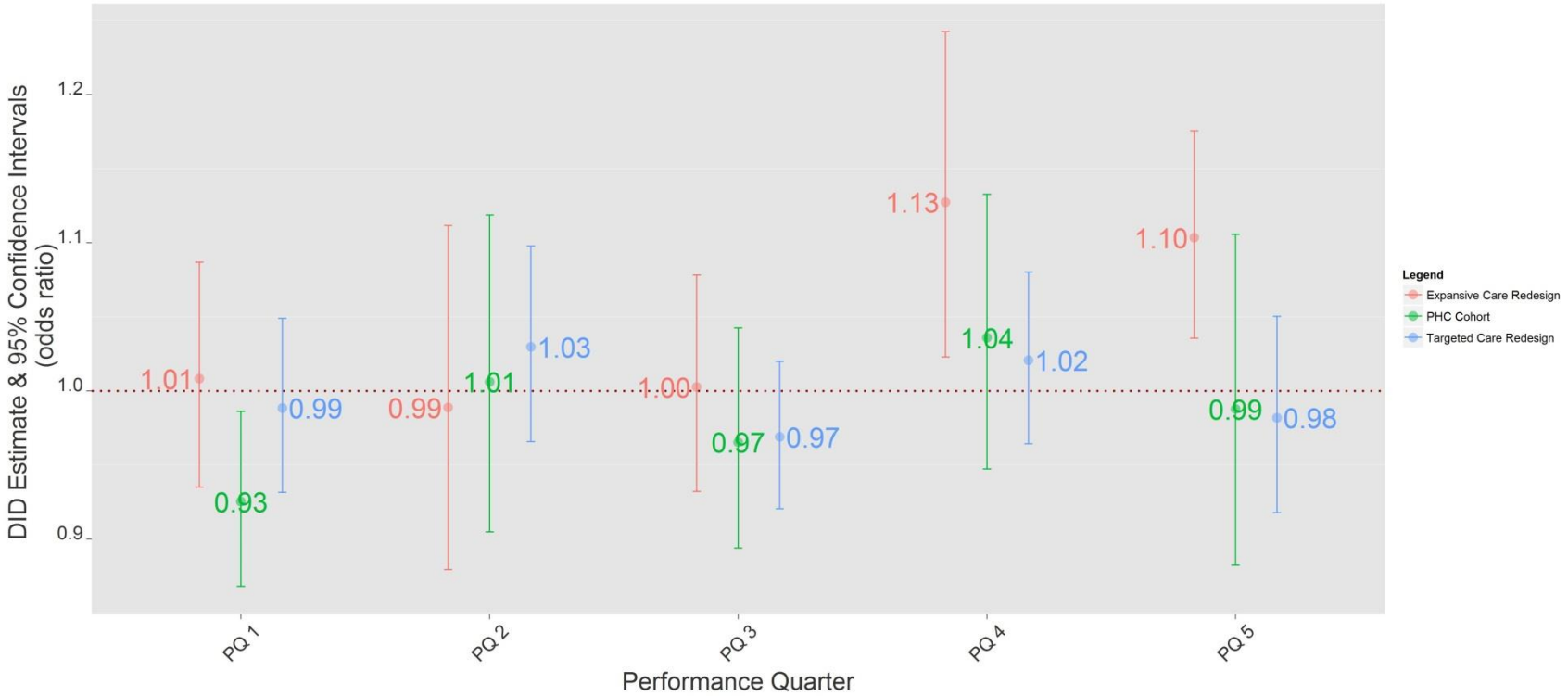
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 9: Adjusted Quarterly Trends for 30-Day All-Cause Mortality for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 10: Performance Quarter DiD Estimates for 30-Day All-Cause Mortality for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

III.B.2.1.2. 30-Day All-Cause Readmissions, Post-Episode

III.B.2.1.2.1. Overview of Results

Table 22 presents 30-day all-cause readmissions as unadjusted and adjusted rates for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. This table also includes these statistics for Active and Exiting cohorts. The Full cohort unadjusted 30-day all-cause readmission rates decreased by0.79 and 0.72 percentage points for Awardee and comparison hospitals, respectively (Since BPCI Inception vs. Baseline; $p < 0.01$). After adjusting for patient and episode characteristics, the rates decreased by 0.62 and 0.65 percentage points for Awardee and comparison hospitals, respectively. The unadjusted rates decreased by 0.46 and 0.76 percentage points among Active cohort Awardee and comparison hospitals, respectively, and the adjusted rates decreased by 0.34 and 0.68 percentage points when comparing the Since BPCI Inception period to Baseline. Exiting cohort Awardee and comparison hospitals had unadjusted rate decreases from Baseline to Since BPCI Inception of 1.08 and 0.48 percentage points ($p < 0.05$). Adjusted, these differences were –0.9 and –0.41 for Awardee and comparison hospitals, respectively.

Table 23 presents 30-day all-cause readmissions as unadjusted, unweighted rates for Full, Active, and Exiting cohorts. The data indicate that some (Awardee and comparison) hospitals may be driving the unadjusted statistics for Full and Active BPCI cohorts downward (relative to weighted, unadjusted data—Table 22). Unadjusted hospital specific trends (not shown) indicate that KSRC and its comparison hospitals are indeed driving these lower readmission rates for Awardee and comparison cohorts, similar to what was seen with mortality rates.

Table 24 presents Full, Active, and Exiting cohort DiD estimates. These estimates assess whether patients discharged from Awardee hospitals had a greater (or lesser) odds of a readmission (i.e., rehospitalization) event within 30 days after episode-discharge during the Since BPCI Inception period, relative to Baseline and the same time period difference for comparison hospitals. The Full cohort estimate is not statistically significant (1.01 OR), indicating no increased likelihood of readmission events for Awardee hospital patients relative to Baseline and comparison hospital patients. Subcohort analysis also shows Active Awardees with non-statistically significant impact estimates. Figure 12, presenting the quarterly breakdown of the Since BPCI Inception period, shows elevated likelihood of readmissions in PQ 4 and 5 for the Active cohort. Further analysis of unadjusted and adjusted readmission rates for Awardee and comparison hospitals in the Active cohort show that readmissions decreased in later PQs from; however, the reduction for Active comparison hospitals outpaced that of Active Awardee hospitals (Table 22). Within the Exiting cohort, reductions in the readmission likelihood of Awardees outpaced that of their comparison hospitals and yielded a statistically significant decrease in the odds of their patients having a readmission event (0.97 OR; $p < 0.10$) relative to Baseline and comparisons.

Tables 25 through 27 and Figures 13 and 14 detail a similar narrative for PHC, Expansive, and Targeted cohorts.

III.B.2.1.2.2. Analysis of Measure Findings – Readmissions

Similar to mortality, BPCI Model 1 provides a complex set of incentives related to readmissions. The incentive to coordinate care without compromising the quality of care may potentially reduce rehospitalizations attributable to inadequate care coordination. Further, hospitals face a nationwide focus on reducing readmissions (e.g., through CMS' Hospital Readmission

Reduction Program, which financially penalizes hospitals for excess readmissions). However, Awardee incentives to reduce internal costs could result in increased readmissions due to care stinting (e.g., reducing a patient's length of stay or decreasing using of imaging services that might otherwise be needed).

Full cohort DiD analysis indicates that there were no statistically significant impacts in readmissions for the combined first five PQs. In other words, when compared to comparison hospitals with similar readmission trends before the model began, BPCI Model 1 did not produce statistically significant changes from Baseline when considering all Awardees. Further, despite reductions in *actual* readmissions (Table 22), these reductions did not translate to an impact attributable to BPCI Model 1. This suggests that the care coordination initiatives that the BPCI hospitals undertook have not translated into post-episode reductions in readmissions beyond that of comparison hospitals.

While the lack of statistically significant findings is also present among the Expansive and Targeted redesign Awardee hospitals, results differ for Exiting and Active Awardee hospitals. Exiting Awardee hospitals produced a statistically significant reduction in the likelihood of patient readmission (0.97, $p < 0.10$), while Active Awardees did not. The rationale for this differential impact between Active and Exiting Awardees may be attributable to patients from Active hospitals having a higher risk profile (as evidenced from higher CMS-HCC scores) than Exiting Awardees (Tables 14 and 15).

Table 22: Episode Weighted, Unadjusted and Adjusted 30-Day All-Cause Readmissions for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean (%)**	18.97	19.94	18.18***	19.32	18.24***	19.34	17.96***	19.24
	95% CI	[18.82, 19.12]	[19.32, 20.56]	[17.98, 18.38]	[18.68, 19.97]	[18.01, 18.46]	[18.70, 19.98]	[17.51, 18.40]	[18.58, 19.89]
	N	267,923		141,494		112,594		28,900	
Comparison	**Mean (%)**	18.24	17.91	17.52***	17.26	17.61***	17.28	17.16***	17.19
	95% CI	[18.16, 18.33]	[17.15, 18.68]	[17.41, 17.63]	[16.51, 18.02]	[17.48, 17.73]	[16.53, 18.04]	[16.91, 17.41]	[16.44, 17.93]
	N	846,223		440,463		352,987		87,476	
	Active Cohort								
BPCI	**Mean (%)**	19.41	20.15	18.95**	19.81	19.02**	19.84	18.68**	19.67
	95% CI	[19.20, 19.63]	[19.33, 20.96]	[18.66, 19.24]	[18.88, 20.73]	[18.69, 19.34]	[18.93, 20.75]	[18.04, 19.34]	[18.71, 20.64]
	N	136,072		69,855		55,772		14,083	
Comparison	**Mean (%)**	18.73	18.52	17.97***	17.84	18.07***	17.87	17.58***	17.71
	95% CI	[18.63, 18.84]	[17.70, 19.35]	[17.83, 18.11]	[17.02, 18.66]	[17.91, 18.23]	[17.04, 18.69]	[17.26, 17.89]	[16.90, 18.52]
	N	541,355		282,591		226,186		56,405	
	Exiting Cohort								
BPCI	**Mean (%)**	18.51	19.98	17.43***	19.08*	17.47***	19.07*	17.26***	19.09*
	95% CI	[18.30, 18.72]	[19.33, 20.62]	[17.15, 17.71]	[18.39, 19.76]	[17.16, 17.78]	[18.40, 19.75]	[16.66, 17.88]	[18.38, 19.79]
	N	131,851		71,639		56,822		14,817	
Comparison	**Mean (%)**	17.65	17.15	17.17***	16.74	17.24***	16.74	16.91***	16.75
	95% CI	[17.53, 17.77]	[16.30, 17.99]	[17.01, 17.34]	[15.93, 17.55]	[17.05, 17.42]	[15.93, 17.55]	[16.54, 17.29]	[15.93, 17.57]
	N	382,746		198,195		159,179		39,016	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 23: Unweighted, Unadjusted 30-Day All-Cause Readmissions for Full, Active, and Exiting Cohorts [+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Full Cohort				
BPCI	**Mean (%)**	18.34	18.10*	18.25	17.51***
	95% CI	[16.02, 20.97]	[15.74, 20.70]	[15.88, 20.85]	[15.18, 20.11]
	N	267,923	141,494	112,594	28,900
Comparison	**Mean (%)**	17.27	17.11**	17.31	16.32***
	95% CI	[14.89, 20.38]	[14.73, 19.90]	[14.93, 20.04]	[13.95, 19.34]
	N	846,223	440,463	352,987	87,476
	Active Cohort				
BPCI	**Mean (%)**	17.98	18.30*	18.57***	17.28**
	95% CI	[15.52, 20.84]	[15.77, 21.12]	[16.02, 21.38]	[14.78, 20.13]
	N	136,072	69,855	55,772	14,083
Comparison	**Mean (%)**	17.21	17.34	17.61***	16.34***
	95% CI	[14.86, 20.51]	[15.01, 20.12]	[15.28, 20.29]	[13.99, 19.49]
	N	541,355	282,591	226,186	56,405
	Exiting Cohort				
BPCI	**Mean (%)**	18.95	17.77***	17.75***	17.88***
	95% CI	[16.86, 21.18]	[15.70, 20.02]	[15.66, 20.01]	[15.83, 20.08]
	N	131,851	71,639	56,822	14,817
Comparison	**Mean (%)**	17.61	17.04***	17.13***	16.70***
	95% CI	[15.27, 20.22]	[14.65, 19.72]	[14.73, 19.82]	[14.34, 19.34]
	N	382,746	198,195	159,179	39,016

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

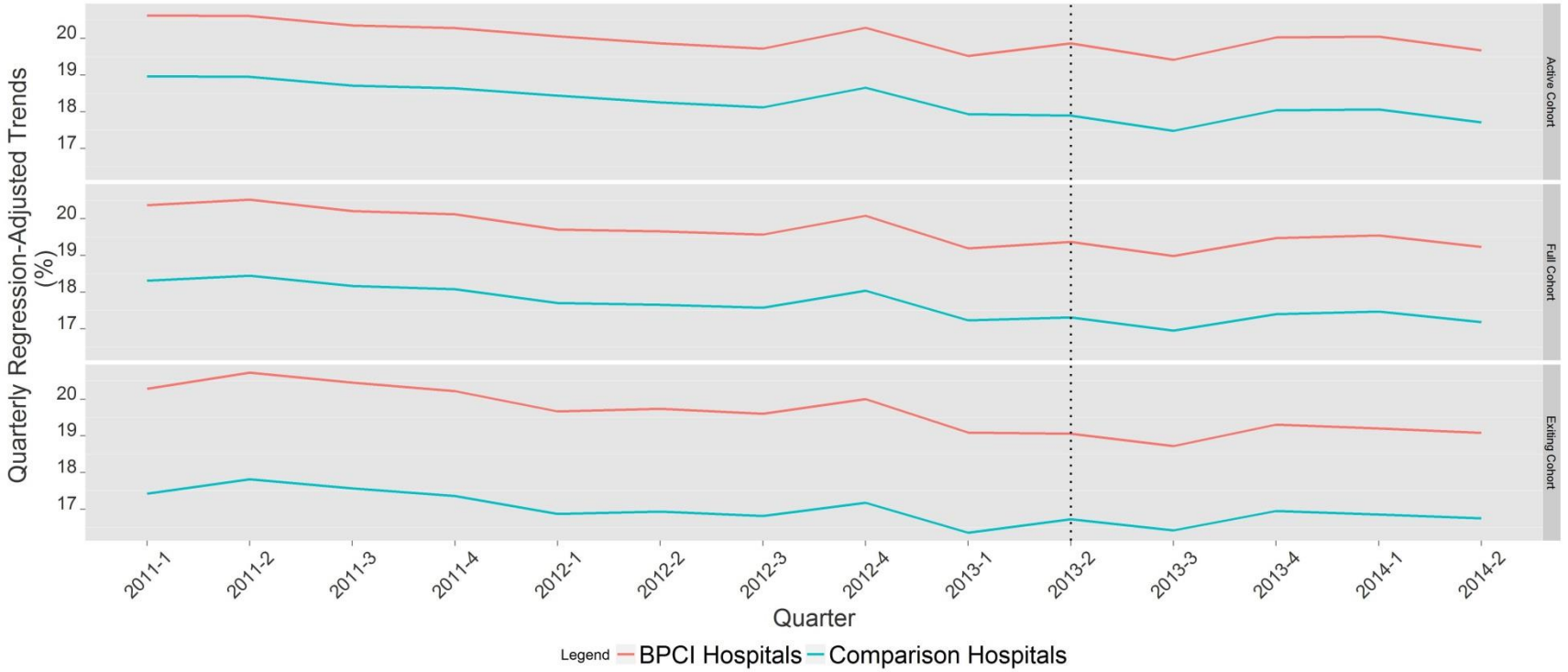
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 24: Since BPCI Inception DiD Estimates for 30-Day All-Cause Readmissions for Full, Active, and Exiting Cohort[+]

	Full Cohort	Active Cohort	Exiting Cohort
DiD (Odds Ratio)	1.01	1.03	0.97*
95% CI	(0.98, 1.03)	(0.99, 1.06)	(0.94, 1.00)
Sample Size	1,696,103	1,029,873	784,431

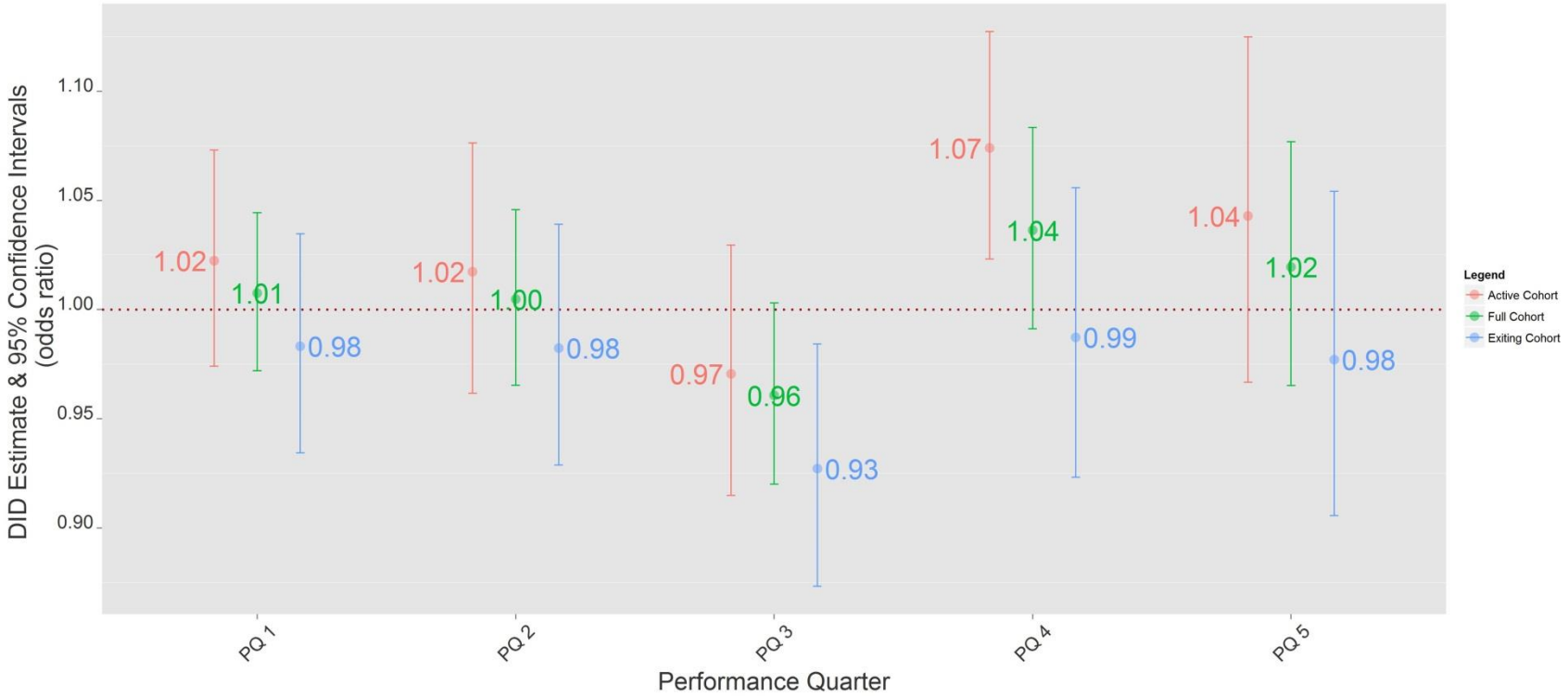
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 11: Adjusted Quarterly Trends for 30-Day All-Cause Readmissions for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 12: Performance Quarter DiD Estimates for 30-Day All-Cause Readmissions for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 25: Episode Weighted, Unadjusted and Adjusted 30-Day All-Cause Readmissions for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Expansive Cohort									
BPCI	**Mean (%)**	18.09	19.11	17.38***	18.43	17.48**	18.46	16.99**	18.29
	95% CI	[17.78, 18.40]	[18.28, 19.94]	[16.95, 17.81]	[17.57, 19.29]	[17.00, 17.96]	[17.59, 19.33]	[16.06, 17.96]	[17.47, 19.10]
	N	59,689		29,979		23,900		6,079	
Comparison	**Mean (%)**	17.91	17.67	17.01***	16.77	17.13***	16.80	16.51***	16.64*
	95% CI	[17.76, 18.06]	[16.80, 18.53]	[16.80, 17.22]	[15.90, 17.64]	[16.90, 17.37]	[15.92, 17.68]	[16.06, 16.98]	[15.83, 17.45]
	N	239,593		123,887		98,818		25,069	
Targeted Cohort									
BPCI	**Mean (%)**	19.22	20.69	18.39***	20.08	18.44***	20.11	18.21***	19.97
	95% CI	[19.05, 19.39]	[20.03, 21.35]	[18.17, 18.62]	[19.38, 20.77]	[18.19, 18.70]	[19.42, 20.79]	[17.71, 18.72]	[19.25, 20.69]
	N	208,234		111,515		88,694		22,821	
Comparison	**Mean (%)**	18.35	17.85	17.67***	17.28	17.75***	17.30	17.32***	17.19
	95% CI	[18.25, 18.44]	[17.00, 18.70]	[17.54, 17.80]	[16.46, 18.10]	[17.61, 17.90]	[16.49, 18.12]	[17.03, 17.61]	[16.36, 18.02]
	N	633,184		330,091		265,010		65,081	
PHC Cohort									
BPCI	**Mean (%)**	18.45	19.70	16.98***	18.42**	17.13***	18.43**	16.39***	18.37***
	95% CI	[18.20, 18.71]	[19.02, 20.37]	[16.64, 17.32]	[17.69, 19.14]	[16.75, 17.52]	[17.71, 19.15]	[15.66, 17.14]	[17.63, 19.10]
	N	91,372		47,525		37,836		9,689	
Comparison	**Mean (%)**	17.31	16.86	16.69***	16.33	16.73***	16.34	16.52***	16.28
	95% CI	[17.16, 17.45]	[15.99, 17.73]	[16.49, 16.89]	[15.49, 17.16]	[16.51, 16.96]	[15.51, 17.17]	[16.07, 16.98]	[15.43, 17.13]
	N	257,641		132,512		106,450		26,062	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 26: Unweighted, Unadjusted 30-Day All-Cause Readmissions for Expansive, Targeted, and PHC Cohorts [+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Expansive Cohort				
BPCI	**Mean (%)**	15.86	16.78***	17.08***	15.69
	95% CI	[13.40, 18.89]	[14.19, 19.72]	[14.49, 19.99]	[13.12, 18.72]
	N	59,689	29,979	23,900	6,079
Comparison	**Mean (%)**	15.52	16.12***	16.48***	14.83***
	95% CI	[13.04, 19.53]	[13.71, 19.22]	[14.07, 19.40]	[12.39, 18.58]
	N	239,593	123,887	98,818	25,069
	Targeted Cohort				
BPCI	**Mean (%)**	19.63	18.71***	18.79***	18.41***
	95% CI	[17.38, 22.05]	[16.46, 21.15]	[16.52, 21.24]	[16.20, 20.80]
	N	208,234	111,515	88,694	22,821
Comparison	**Mean (%)**	18.30	17.58***	17.71***	17.08***
	95% CI	[15.97, 20.86]	[15.22, 20.20]	[15.33, 20.34]	[14.75, 19.67]
	N	633,184	330,091	265,010	65,081
	PHC Cohort				
BPCI	**Mean (%)**	18.33	16.93***	17.06***	16.41***
	95% CI	[16.39, 20.40]	[15.02, 18.99]	[15.13, 19.13]	[14.56, 18.41]
	N	91,372	47,525	37,836	9,689
Comparison	**Mean (%)**	17.38	16.75***	16.80***	16.58***
	95% CI	[15.24, 19.70]	[14.57, 19.13]	[14.62, 19.17]	[14.39, 18.98]
	N	257,641	132,512	106,450	26,062

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

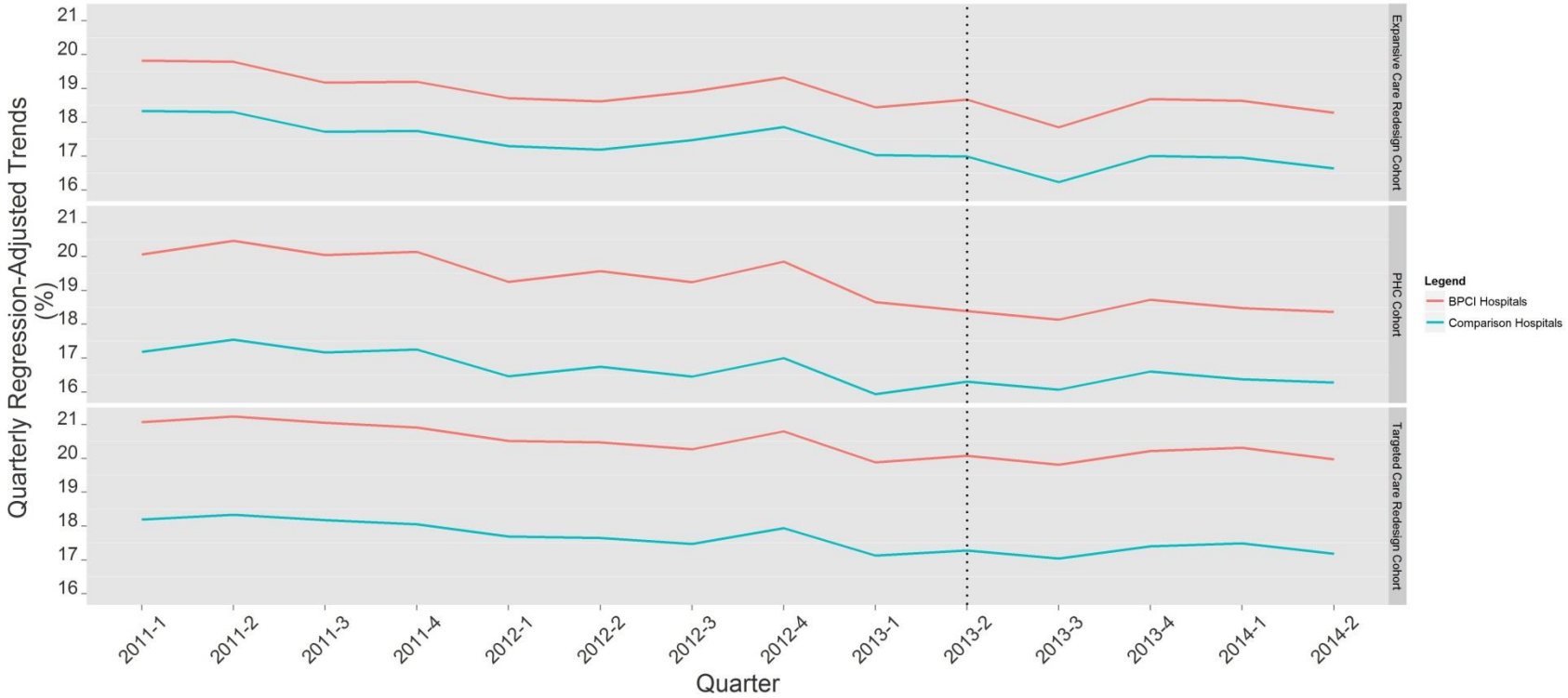
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 27: Since BPCI Inception DiD Estimates for 30-Day All-Cause Readmissions for Expansive, Targeted, and PHC Cohorts[+]

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD (Odds Ratio)	0.96**	1.02	1.00
95% CI	(0.92, 1.00)	(0.97, 1.07)	(0.98, 1.03)
Sample Size	529,050	453,148	1,283,024

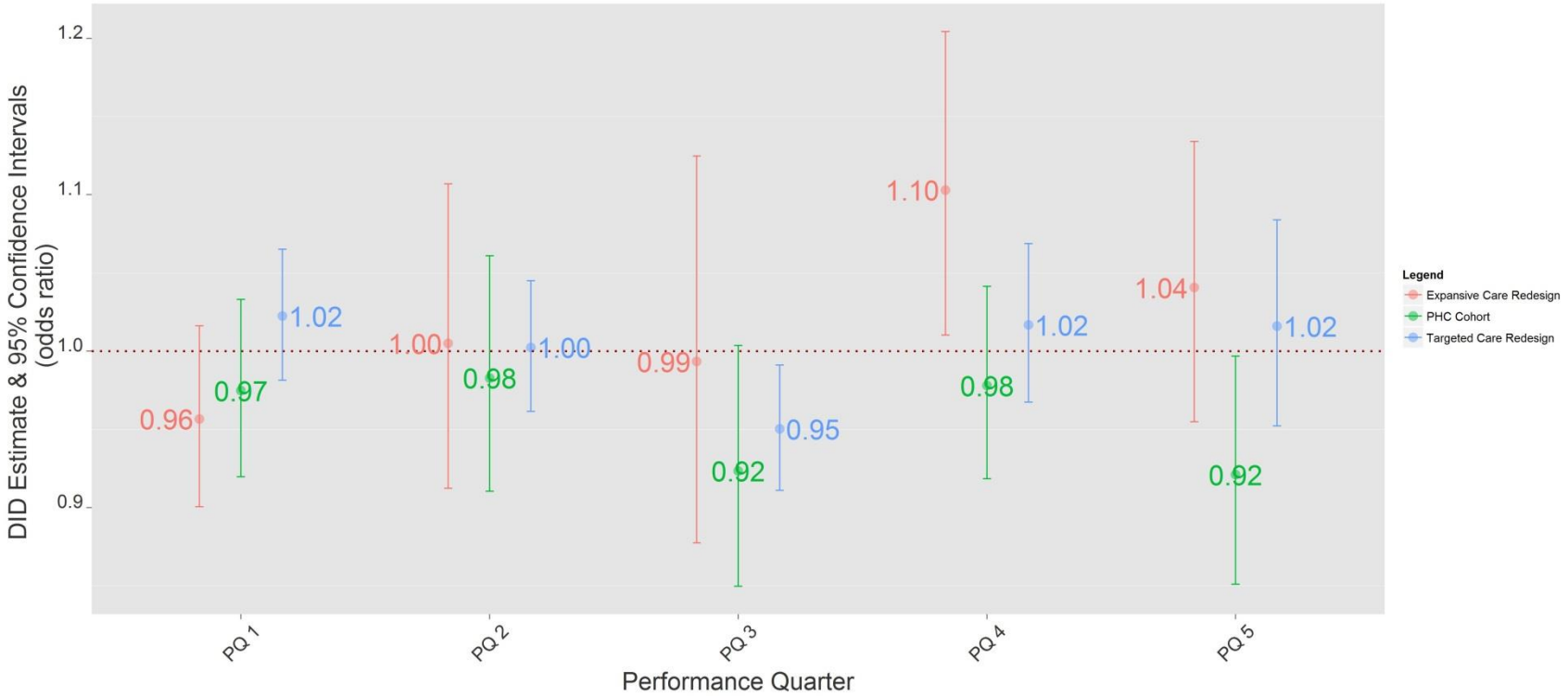
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 13: Adjusted Quarterly Trends for 30-Day All-Cause Readmissions for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 14: Performance Quarter DiD Estimates for 30-Day All-Cause Readmissions for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

III.B.2.1.3. ICU Use During Episode

III.B.2.1.3.1. Overview of Results

Table 28 presents ICU use during an episode as unadjusted and adjusted rates for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. This table also includes these rates for Active and Exiting cohorts. Full cohort unadjusted ICU use rates increased by 0.76 and 0.1 percentage points for Awardee and comparison hospitals, respectively (Since BPCI Inception vs. Baseline). After adjusting for patient and episode characteristics, the ICU use rates increased by 0.44 percentage points for the Awardee hospital, while the rates decreased by 0.35 percentage points for the comparison hospitals. In the Active cohort, the unadjusted ICU use rates increased by 1.23 percentage points for the Awardee hospitals, while the rates decreased by 0.20 percentage points for the comparison hospitals. The adjusted rates increased by 0.55 percentage points for the Awardee hospitals, while the rates decreased by 0.74 percentage points for the comparison hospitals. Unadjusted ICU use rates increased by 0.29 and 0.40 percentage points for the Exiting cohort Awardee hospitals and comparison hospitals, respectively. Adjusted, the rates increased by 0.02 and 0.03 percentage points for Awardee and comparison hospitals, respectively. Table 29 presents unadjusted, unweighted ICU episode use for these same cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted rates in the Full and Active BPCI cohorts to differ from unadjusted, weighted statistics presented in Table 28. However, examination of hospital specific unadjusted trends (Figure 15) indicates that Awardees Capital Health, Capital Health Hopewell, and Deborah are influencing Table 29 statistics.

Table 30 presents Full, Active, and Exiting cohort DiD estimates. These estimates are the impact estimates that assess the likelihood (in ORs) of having ICU services during an episode for BPCI Model 1 Awardees during the model performance period, relative to Baseline and comparison hospitals. The Full cohort estimate indicates a statistically significant increased likelihood (1.10 OR, $p < 0.01$) of patients at Awardee hospitals having an ICU stay relative to Baseline and comparison hospitals. Subcohort analysis shows that this estimate is driven by Active cohort Awardees (1.20 OR, $p < 0.01$), while Exiting cohort Awardee hospitals had no statistically significant impacts on this measure. Figure 17 breaks these estimates down by PQ and shows that BPCI Model 1 impacts were elevated for Active Awardees in all quarters but most heavily in PQ 3 and PQ 4.

Table 31 presents episode ICU use as unadjusted and adjusted rates for PHC, Expansive, and Targeted BPCI and comparison cohorts over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. Generally, Awardee and comparison hospital changes across these cohorts are ambiguous in average ICU use in the Since BPCI Inception period relative to Baseline. Table 32 presents unadjusted, unweighted ICU use as means in percentages for Expansive and Targeted Care Redesign and PHC cohorts. Figure 15 presents unadjusted, hospital-level quarter trends which indicate that Awardees Capital Health, Capital Health Hopewell, and Deborah and their comparison hospitals are influencing the difference between these unweighted statistics and their weighted counterparts in Table 31. Table 34 presents DiD impact estimates for these cohorts and shows a statistically significant impact estimates for Targeted and Expansive Awardees of 1.08 ($p < 0.05$) and 1.15 ($p < 0.01$), respectively. Figure 19 shows the volatility of the Expansive and, to a lesser extent, the Targeted cohort over the first five PQs.

III.B.2.1.3.2. Analysis of Measure Findings – ICU Use

ICU stays represent a costly segment of inpatient care and are associated with more complications and greater chronic disease burden.[96] BPCI Model 1 can provide mixed incentives for ICU use during an inpatient episode. On one hand, the model may reduce ICU use, as Awardee hospitals have an incentive to reduce internal costs Further, higher quality of care—perhaps from better coordinated care—may obviate the need for the higher intensity of care inherent in an ICU environment. Conversely, a push to improve the quality of care—perhaps in reducing post-episode outcomes such as readmission or mortality—may incentivize ICU use for patients that would marginally benefit from intensive care during their hospital stay.

The Full cohort DiD results indicate that the likelihood of a patient having an ICU stay at an Awardee hospital increased (relative to Baseline and comparison hospitals, 1.10 OR, $p < 0.01$). This result appears to be driven by Active cohort Awardee hospitals, with an OR impact estimate of 1.20 ($p < 0.01$). This cohort was previously noted to have a higher average (CMS-HCC) risk profile relative to Exiting Awardees. Hospital specific DiD impact analysis (i.e., one Awardee versus its four comparison hospitals) indicates that all but three of the Active Awardees showed increased likelihoods of patients having an ICU stay during their episode (OR above 1.0), with Awardee Capital Health Hopewell having the largest impact estimate of 2.67 OR ($p < 0.01$). Figure 15 noted Capital Health Hopewell's increasing use of ICU for its patients over time, even before Model 1; however, the reasons for these increases affecting the impact estimate are unclear at this time. Additionally, the attribution of these elevated ICU estimates to Model 1 must be qualified. Placebo DiD tests[97] for the Full and Active cohort indicate that these elevated ICU likelihood estimates were found *before* BPCI Model 1 implementation and may consequently not be attributable to Model 1. Unadjusted ICU statistics do exhibit a 0.76 percentage point increase across Awardee hospitals and a 0.10 percentage point increase across comparison hospitals. This yields an unadjusted DiD result of 0.66 percentage points—indicative of increased ICU utilization—for Awardee hospitals relative to Baseline and comparison hospitals.

[96] Barret et al. (2014). "Utilization of Intensive Care Services, 2011." Statistical Brief No. 185, H-CUP. Retrieved from: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb185-Hospital-Intensive-Care-Units-2011.jsp.

[97] Placebo DiD Tests test the reliability of associating DiD impact estimates with an intervention (e.g., BPCI Model 1) by assuming placebo/pseudo start dates *prior* to the intervention's *actual* start date. Statistically significant estimates in the placebo periods can indicate an inappropriateness of impact estimate attribution to BPCI Model 1.

Table 28: Episode Weighted, Unadjusted and Adjusted ICU Use for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean (%)**	12.24	9.66	13.00***	10.10	13.03***	10.15*	12.90***	9.90
	95% CI	[12.12, 12.37]	[9.29, 10.03]	[12.83, 13.17]	[9.65, 10.56]	[12.83, 13.22]	[9.70, 10.60]	[12.53, 13.29]	[9.42, 10.39]
	N	281,247		148,907		118,635		30,272	
Comparison	**Mean (%)**	13.53	15.12	13.63	14.77	13.71***	14.84	13.30**	14.49
	95% CI	[13.46, 13.60]	[14.53, 15.72]	[13.53, 13.73]	[14.11, 15.42]	[13.60, 13.82]	[14.19, 15.48]	[13.08, 13.52]	[13.78, 15.21]
	N	887,693		463,977		372,062		91,915	
	Active Cohort								
BPCI	**Mean (%)**	12.32	5.95	13.55***	6.50	13.62***	6.54	13.27***	6.33
	95% CI	[12.15, 12.49]	[5.45, 6.45]	[13.31, 13.80]	[5.90, 7.10]	[13.35, 13.90]	[5.95, 7.14]	[12.73, 13.83]	[5.71, 6.94]
	N	142,419		73,716		58,965		14,751	
Comparison	**Mean (%)**	13.99	18.25	13.79**	17.51	13.91	17.61	13.31***	17.09**
	95% CI	[13.90, 14.08]	[17.58, 18.93]	[13.67, 13.92]	[16.71, 18.30]	[13.77, 14.05]	[16.84, 18.37]	[13.04, 13.59]	[16.18, 18.01]
	N	566,246		297,179		238,369		58,810	
	Exiting Cohort								
BPCI	**Mean (%)**	12.17	20.47	12.46**	20.49	12.44*	20.50	12.55	20.46
	95% CI	[12.00, 12.34]	[19.64, 21.31]	[12.22, 12.70]	[19.48, 21.50]	[12.17, 12.70]	[19.48, 21.51]	[12.03, 13.08]	[19.44, 21.48]
	N	138,828		75,191		59,670		15,521	
Comparison	**Mean (%)**	12.35	11.13	12.75***	11.16	12.76***	11.16	12.75**	11.14
	95% CI	[12.25, 12.45]	[9.83, 12.43]	[12.61, 12.90]	[9.80, 12.51]	[12.60, 12.92]	[9.81, 12.51]	[12.43, 13.07]	[9.76, 12.52]
	N	403,966		209,467		167,988		41,479	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 29: Unweighted, Unadjusted ICU Use for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Full Cohort				
BPCI	**Mean (%)**	13.74	15.15***	15.18***	15.06***
	95% CI	[11.82, 15.87]	[13.07, 17.44]	[13.09, 17.46]	[13.01, 17.32]
	N	281,247	148,907	118,635	30,272
Comparison	**Mean (%)**	13.86	13.80	13.91	13.35***
	95% CI	[12.00, 15.98]	[11.89, 15.98]	[12.00, 16.10]	[11.47, 15.50]
	N	887,693	463,977	372,062	91,915
	Active Cohort				
BPCI	**Mean (%)**	13.45	15.21***	15.31***	14.81***
	95% CI	[11.45, 15.69]	[13.01, 17.65]	[13.11, 17.75]	[12.63, 17.22]
	N	142,419	73,716	58,965	14,751
Comparison	**Mean (%)**	14.45	14.07***	14.20***	13.54***
	95% CI	[12.52, 16.57]	[12.13, 16.22]	[12.26, 16.35]	[11.63, 15.67]
	N	566,246	297,179	238,369	58,810
	Exiting Cohort				
BPCI	**Mean (%)**	14.20	15.06***	14.96***	15.46***
	95% CI	[12.40, 16.16]	[13.16, 17.11]	[13.06, 17.01]	[13.59, 17.48]
	N	138,828	75,191	59,670	15,521
Comparison	**Mean (%)**	12.43	12.76***	12.81***	12.56
	95% CI	[10.77, 14.42]	[11.02, 14.85]	[11.06, 14.91]	[10.84, 14.61]
	N	403,966	209,467	167,988	41,479

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

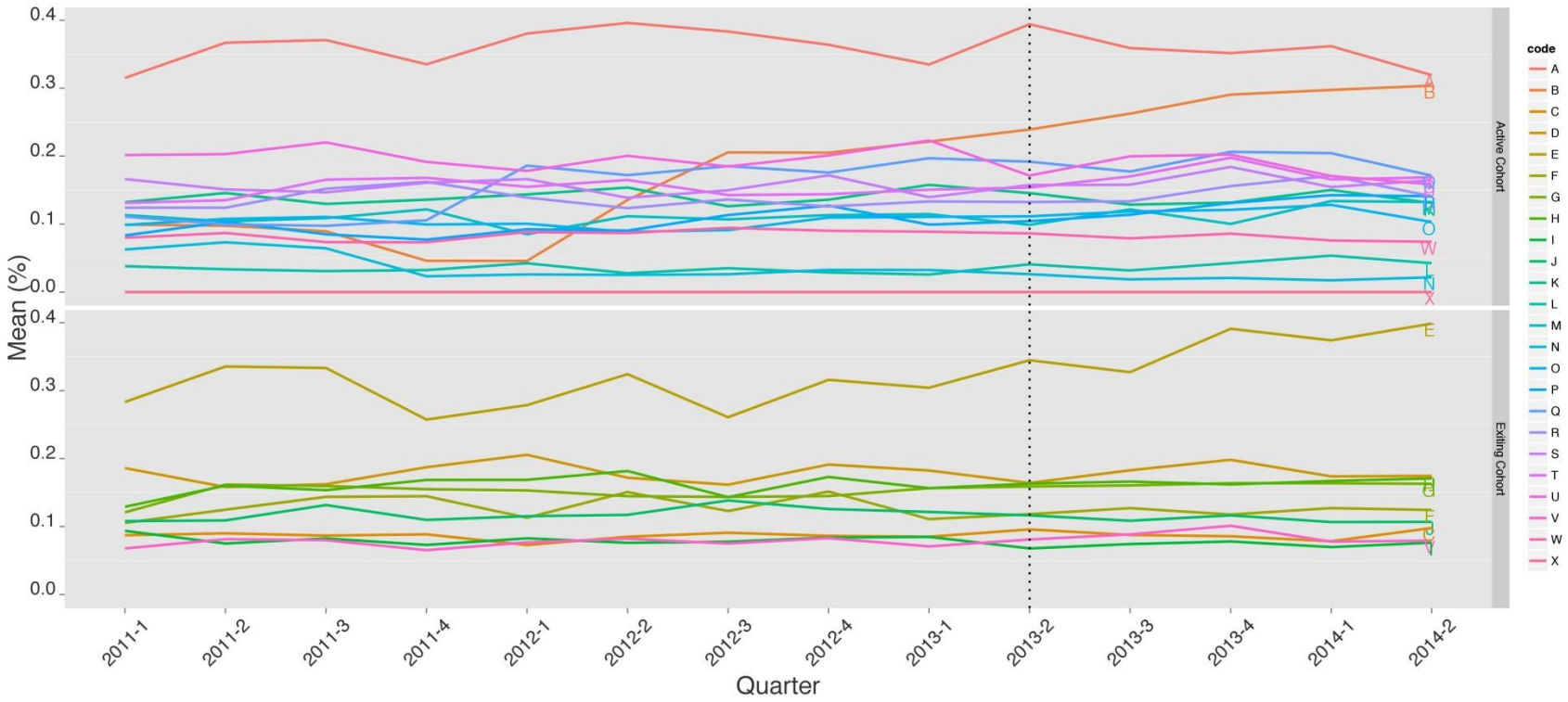
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 30: Since BPCI Inception DiD Estimates for ICU Use for Full, Active, and Exiting Cohort[+]

	Full Cohort	Active Cohort	Exiting Cohort
DiD (Odds Ratio)	1.10***	1.20***	1.00
95% CI	(1.05, 1.16)	(1.12, 1.30)	(0.94, 1.06)
Sample Size	1,764,189	1,079,559	809,818

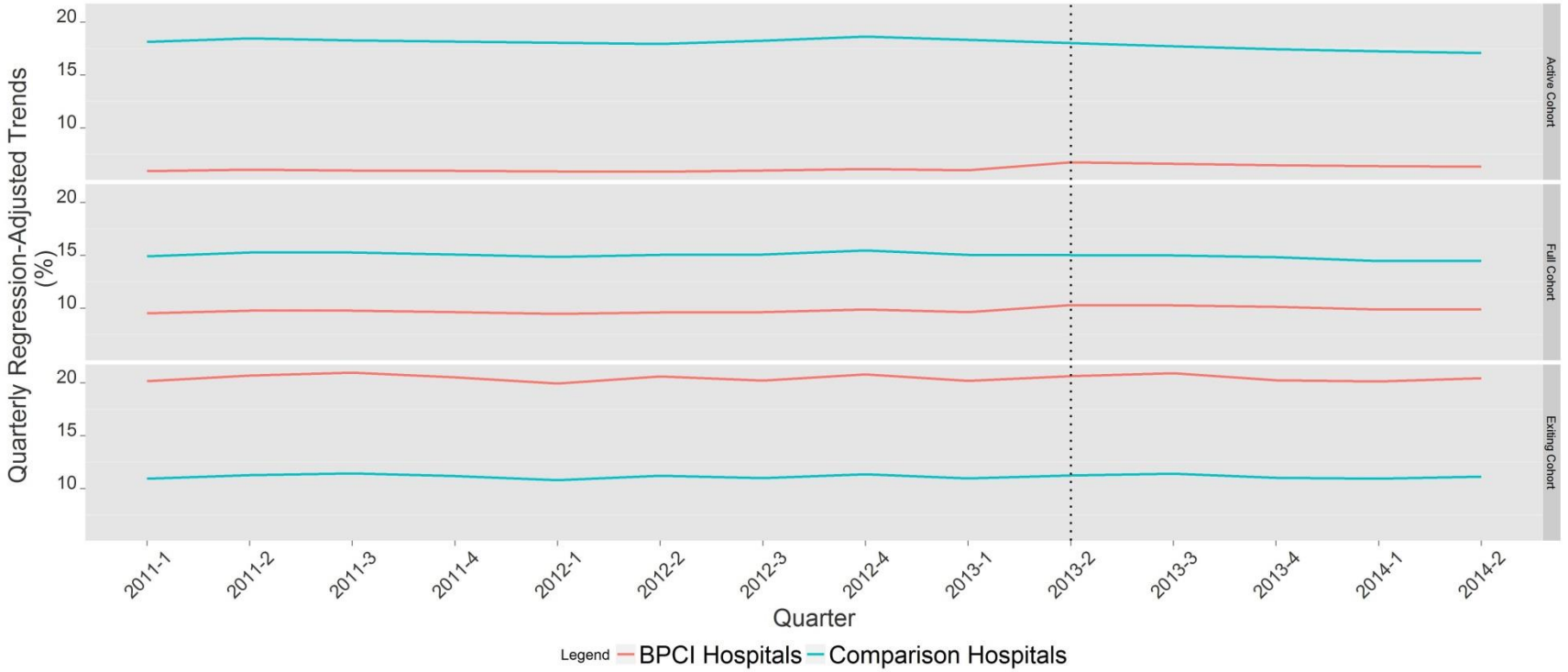
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 15: Unadjusted Quarterly Trends for ICU Use By Awardee Hospital for Active and Exiting Cohorts*



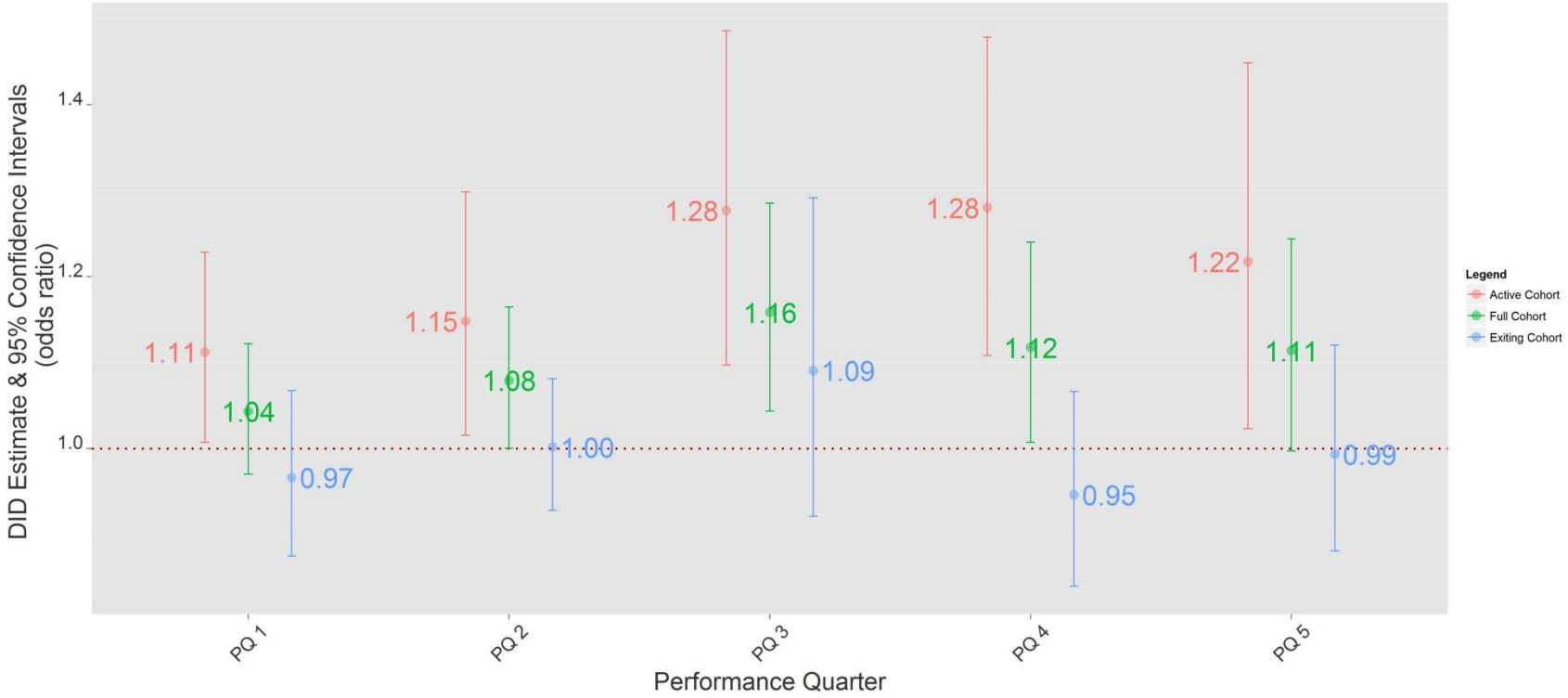
* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014. See Appendix E for Awardee identification.

Figure 16: Adjusted Quarterly Trends for Episode ICU Use for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 17: Performance Quarter DiD Estimates for Episode ICU Use for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 31: Episode Weighted, Unadjusted and Adjusted ICU Use for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Expansive Cohort								
BPCI	**Mean (%)**	13.95	16.23	14.82***	17.23	14.94***	17.44	14.37	16.39
	95% CI	[13.67, 14.22]	[15.34, 17.12]	[14.43, 15.22]	[15.99, 18.47]	[14.50, 15.38]	[16.26, 18.63]	[13.50, 15.26]	[14.94, 17.84]
	N	61,136		31,197		24,981		6,216	
Comparison	**Mean (%)**	14.44	14.09	14.38	13.55	14.60	13.73	13.54***	12.87*
	95% CI	[14.30, 14.58]	[13.33, 14.85]	[14.19, 14.58]	[12.61, 14.50]	[14.38, 14.81]	[12.86, 14.60]	[13.12, 13.96]	[11.63, 14.10]
	N	247,156		129,594		103,821		25,773	
	Targeted Cohort								
BPCI	**Mean (%)**	11.77	21.46	12.52***	22.18	12.52***	22.22	12.52***	22.05
	95% CI	[11.64, 11.91]	[20.72, 22.21]	[12.33, 12.71]	[21.26, 23.10]	[12.30, 12.73]	[21.31, 23.12]	[12.11, 12.95]	[21.09, 23.02]
	N	220,111		117,710		93,654		24,056	
Comparison	**Mean (%)**	13.25	11.75	13.44**	11.57	13.47***	11.58	13.30	11.49
	95% CI	[13.17, 13.33]	[10.38, 13.13]	[13.32, 13.55]	[10.18, 12.95]	[13.34, 13.60]	[10.21, 12.96]	[13.04, 13.55]	[10.09, 12.89]
	N	668,504		348,605		279,659		68,946	
	PHC Cohort								
BPCI	**Mean (%)**	11.63	21.14	12.20***	21.37	12.19***	21.47	12.23*	20.99
	95% CI	[11.43, 11.84]	[20.12, 22.15]	[11.91, 12.49]	[20.28, 22.47]	[11.87, 12.52]	[20.41, 22.52]	[11.60, 12.88]	[19.73, 22.24]
	N	96,243		49,876		39,712		10,164	
Comparison	**Mean (%)**	12.63	10.94	13.42***	11.11	13.47***	11.16	13.22***	10.88
	95% CI	[12.51, 12.76]	[9.62, 12.26]	[13.24, 13.60]	[9.73, 12.48]	[13.27, 13.67]	[9.81, 12.52]	[12.82, 13.62]	[9.43, 12.32]
	N	271,554		140,080		112,235		27,845	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 32: Unweighted, Unadjusted ICU Use for Expansive, Targeted, and PHC Cohorts [+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Expansive Cohort				
BPCI	**Mean (%)**	14.04	15.00***	15.14***	14.44
	95% CI	[11.84, 16.49]	[12.64, 17.64]	[12.77, 17.79]	[12.11, 17.04]
	N	61,136	31,197	24,981	6,216
Comparison	**Mean (%)**	14.75	14.33***	14.58	13.34***
	95% CI	[12.69, 17.03]	[12.28, 16.63]	[12.51, 16.89]	[11.34, 15.58]
	N	247,156	129,594	103,821	25,773
	Targeted Cohort				
BPCI	**Mean (%)**	13.61	15.22***	15.19***	15.34***
	95% CI	[11.81, 15.60]	[13.26, 17.35]	[13.23, 17.32]	[13.40, 17.44]
	N	220,111	117,710	93,654	24,056
Comparison	**Mean (%)**	13.51	13.67**	13.71***	13.51
	95% CI	[11.73, 15.57]	[11.82, 15.81]	[11.85, 15.85]	[11.66, 15.63]
	N	668,504	348,605	279,659	68,946
	PHC Cohort				
BPCI	**Mean (%)**	11.99	12.27	12.25	12.37
	95% CI	[10.42, 13.71]	[10.66, 14.04]	[10.64, 14.02]	[10.78, 14.11]
	N	96,243	49,876	39,712	10,164
Comparison	**Mean (%)**	13.44	13.89***	14.02***	13.37
	95% CI	[11.66, 15.40]	[12.03, 15.94]	[12.15, 16.08]	[11.57, 15.38]
	N	271,554	140,080	112,235	27,845

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

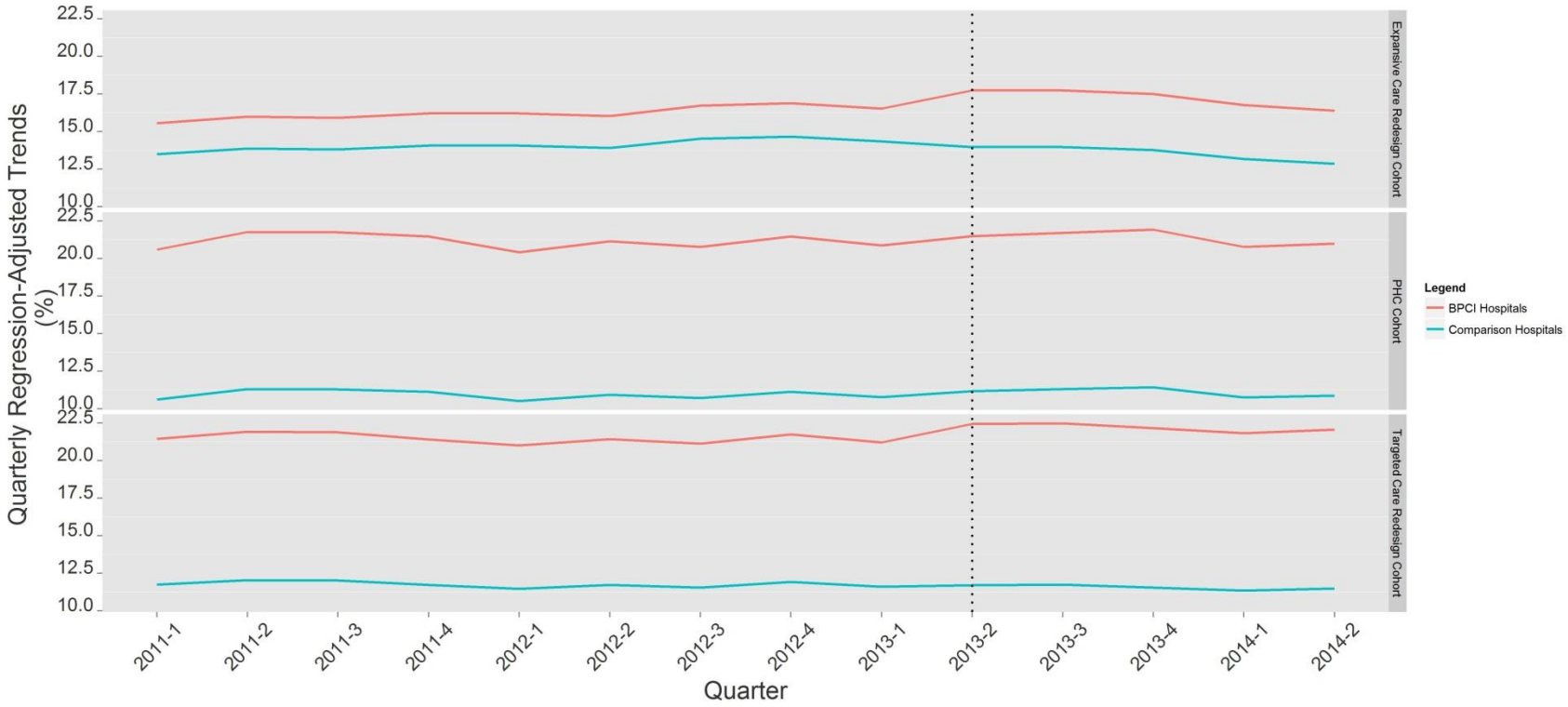
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 33: Since BPCI Inception DiD Estimates for ICU Use for Expansive, Targeted, and PHC Cohorts[+]

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD (Odds Ratio)	1.00	1.15***	1.08**
95% CI	(0.93, 1.07)	(1.05, 1.26)	(1.02, 1.14)
Sample Size	557,752	469,083	1,337,295

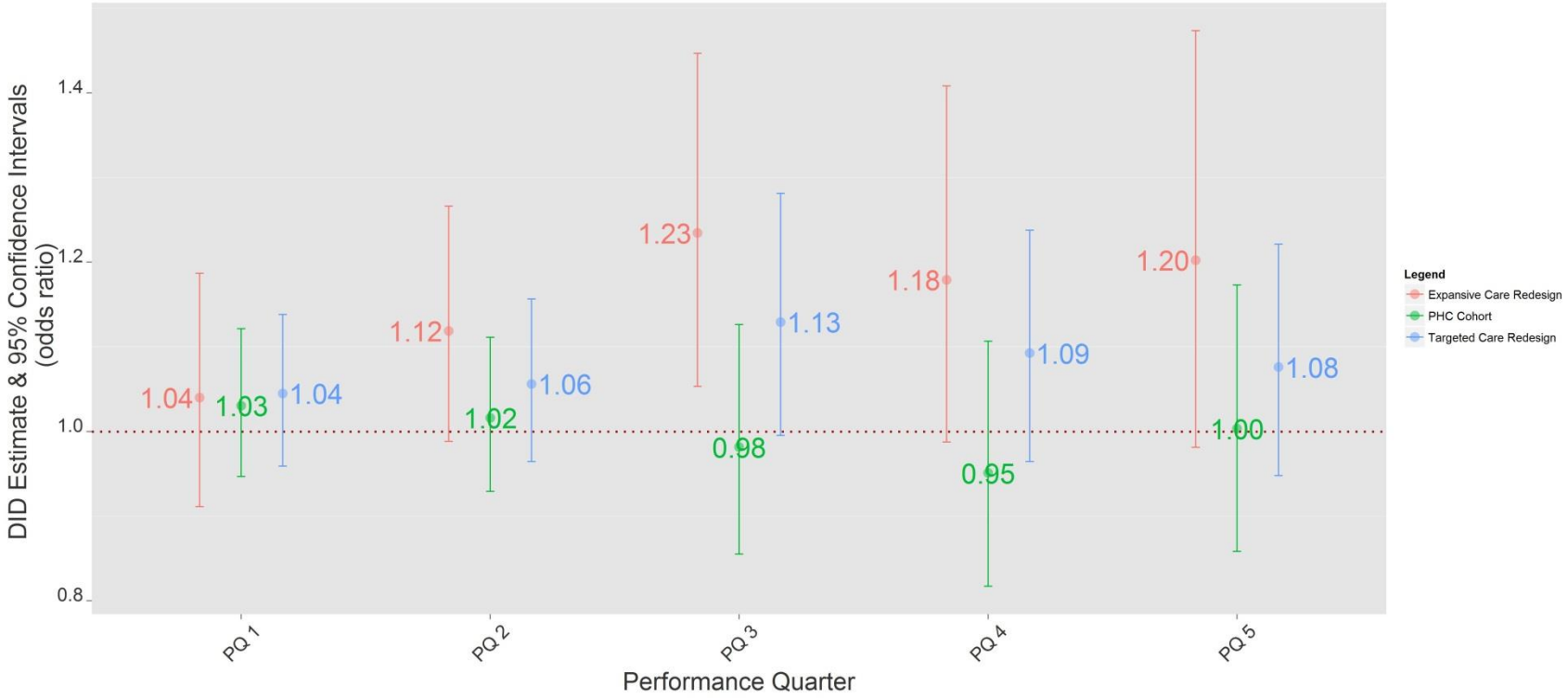
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 18: Adjusted Quarterly Trends for Episode ICU Use for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 19: Performance Quarter DiD Estimates for Episode ICU Use for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

III.B.2.1.4. Episode Length of Stay

III.B.2.1.4.1. Overview of Results

Table 34 presents the length of an episode stay as unadjusted and adjusted average days for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. This table also includes these statistics for Active and Exiting cohorts. Full cohort unadjusted average length of stay did not increase for Awardees, however minimally increased by 0.05 days for the comparison hospitals (Since BPCI Inception vs. Baseline). After adjusting for patient and episode characteristics, the average length of stay decreased by 0.05 and 0.01 days. Active cohort Awardee and comparison hospitals had changes in unadjusted means of –0.04 and 0.05, respectively, and adjusted differences of –0.06 and –0.01 when comparing Since BPCI Inception to Baseline. Exiting cohort Awardee and comparison hospitals had unadjusted average increases from Baseline to Since BPCI Inception of 0.04 and 0.02 days. Adjusted, these differences became –0.04 and –0.03 days for Awardee and comparison hospitals respectively.

Table 35 presents average length of stay as unadjusted, unweighted statistics in days for Full, Active, and Exiting cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted statistics in the Full and Active BPCI cohorts downward (relative to unadjusted, weighted statistics—Table 34). Unadjusted hospital-level trends (not shown) indicate that KSRC is a consistent outlier for the Active and Full BPCI cohorts with its average length of stay centered around 4 days. KSRC's matched comparison hospitals engender a similar effect on comparison cohorts.

Table 36 presents Full, Active, and Exiting cohort DiD estimates. These estimates are the impact estimates of BPCI Model 1 when Awardee hospital episode durations are compared to those of comparison hospitals over Baseline and Since BPCI Inception periods. The Full cohort estimate is an average reduction in days spent of –0.05 (not significant). Subcohort analysis shows that this estimate is driven by Active cohort Awardees (–0.06, not significant), while Exiting cohort Awardees had no statistically significant impacts on this measure. Figure 21 breaks these estimates down by PQ and show BPCI Model 1 impacts of lesser (albeit not statistically significant) magnitudes.

Table 37 presents length of stay as unadjusted and adjusted days for PHC, Expansive, and Targeted BPCI and comparison cohorts over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. Generally, Awardee and comparison hospitals across these cohorts show ambiguous changes in average episode length of stay from Baseline to Since BPCI Inception. Table 38 presents unadjusted, unweighted length of stay in mean days for Expansive and Targeted Care Redesign and PHC cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted statistics for the Expansive BPCI cohort downward (relative to weighted, unadjusted data—Table 37). Unadjusted hospital-level trends (not shown) indicate that KSRC is likely responsible, with its previously noted average length of stay around 4 days across Baseline and PQs.

Table 39 presents impact estimates for these cohorts. This table exhibits the Targeted Care Redesign and PHC cohort have no statistically significant impact, but the Expansive cohort deceases by –0.16 ($p < 0.01$).

III.B.2.1.4.2. Analysis of Measure Findings – Episode Length of Stay

Reducing episode length (of stay) is an important component of internal hospital cost reductions. Therefore, all else being equal, BPCI Model 1 Awardees are expected to reduce length of stay as long as the reduction does not compromise quality of care. This expectation stems from an Awardee's internal cost reduction motive and from reported care redesigns, across Awardee hospitals that aimed to improve patient throughput and discharge efficiency.

BPCI Model 1 did not produce statistically significant changes in length of stay. The lack of a statistical impact seems contrary to Awardee focus on decreasing episode length of stay directly or indirectly through care redesigns that aim to improve patient throughput or increase efficiency in patient discharge processes (through improved coordination, for example). However, this lack of sizable finding may stem from the delayed appearance of effects from care redesigns[98] or from Awardee hospital efforts achieving minimal or varied changes in episode length of stay. Diving deeper into this estimate, Awardee specific DiD impact estimates indicate that 11 of the 24 hospital Awardees did achieve statistically significant decreases in episode length of stay ranging from decreases of 0.22 to 0.62 days (all statistically significant, $p < 0.01$). Of these 11 Awardees, 8 are in the Active cohort, and 5 of these 8 Awardees comprise the majority of the Expansive cohort. In other words, these Awardee hospital specific impact estimates correspond with hospital-wide care coordination redesign efforts indicative of Awardee hospitals in the Expansive cohort. Further, as previously noted, BPCI Model 1 Awardees previously in the PHC had indicated that they had already achieved reductions in patient length of stay (Section III.A). These hospitals were six of the nine Awardees in the exiting cohort and were all in the Targeted Care Redesign cohort.

[98] As previously reported (Section III.A.3), care redesigns were reported to take 3 to 6 months to implement.

Table 34: Episode Weighted, Unadjusted and Adjusted Episode Length of Stay for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean (days)**	6.62	7.29	6.62	7.24	6.63	7.24	6.58	7.23
	95% CI	[6.60, 6.64]	[7.11, 7.47]	[6.59, 6.65]	[7.06, 7.41]	[6.60, 6.67]	[7.06, 7.41]	[6.52, 6.64]	[7.06, 7.41]
	N	282,895		149,233		118,799		30,434	
Comparison	**Mean (days)**	6.42	6.23	6.47***	6.22	6.47***	6.23	6.44	6.22
	95% CI	[6.41, 6.43]	[6.15, 6.31]	[6.45, 6.48]	[6.15, 6.30]	[6.46, 6.49]	[6.15, 6.30]	[6.41, 6.48]	[6.14, 6.30]
	N	892,496		464,815		372,482		92,333	
	Active Cohort								
BPCI	**Mean (days)**	6.63	6.98	6.59	6.92	6.60	6.92	6.54*	6.91
	95% CI	[6.60, 6.66]	[6.83, 7.14]	[6.55, 6.64]	[6.76, 7.07]	[6.55, 6.65]	[6.77, 7.07]	[6.46, 6.63]	[6.75, 7.06]
	N	144,067		74,042		59,129		14,913	
Comparison	**Mean (days)**	6.41	6.35	6.46***	6.34	6.47***	6.35	6.42	6.33
	95% CI	[6.40, 6.43]	[6.26, 6.43]	[6.44, 6.48]	[6.26, 6.43]	[6.45, 6.50]	[6.27, 6.43]	[6.38, 6.46]	[6.25, 6.42]
	N	571,049		298,017		238,789		59,228	
	Exiting Cohort								
BPCI	**Mean (days)**	6.61	7.35	6.65	7.31	6.66*	7.32	6.61	7.30
	95% CI	[6.58, 6.64]	[7.12, 7.58]	[6.61, 6.69]	[7.09, 7.54]	[6.61, 6.70]	[7.10, 7.54]	[6.52, 6.69]	[7.08, 7.53]
	N	138,828		75,191		59,670		15,521	
Comparison	**Mean (days)**	6.46	6.23	6.48	6.20	6.49*	6.20	6.46	6.19
	95% CI	[6.44, 6.48]	[6.12, 6.34]	[6.46, 6.51]	[6.10, 6.30]	[6.46, 6.52]	[6.10, 6.30]	[6.41, 6.51]	[6.08, 6.29]
	N	403,966		209,467		167,988		41,479	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 35: Unweighted, Unadjusted Episode Length of Stay for Full, Active, and Exiting Cohorts [+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Full Cohort				
BPCI	**Mean (days)**	6.37	6.42***	6.45***	6.32*
	95% CI	[6.08, 6.66]	[6.11, 6.73]	[6.13, 6.76]	[6.03, 6.60]
	N	282,895	149,233	118,799	30,434
Comparison	**Mean (days)**	6.09	6.28***	6.31***	6.16***
	95% CI	[5.79, 6.40]	[5.96, 6.59]	[5.99, 6.62]	[5.86, 6.46]
	N	892,496	464,815	372,482	92,333
	Active Cohort				
BPCI	**Mean (days)**	6.30	6.33	6.37***	6.18***
	95% CI	[6.01, 6.59]	[6.00, 6.65]	[6.04, 6.70]	[5.89, 6.48]
	N	144,067	74,042	59,129	14,913
Comparison	**Mean (days)**	6.03	6.28***	6.32***	6.11***
	95% CI	[5.73, 6.33]	[5.98, 6.58]	[6.02, 6.62]	[5.82, 6.40]
	N	571,049	298,017	238,789	59,228
	Exiting Cohort				
BPCI	**Mean (days)**	6.50	6.56***	6.57***	6.54
	95% CI	[6.21, 6.78]	[6.28, 6.85]	[6.28, 6.86]	[6.28, 6.80]
	N	138,828	75,191	59,670	15,521
Comparison	**Mean (days)**	6.28	6.31***	6.33***	6.27
	95% CI	[5.95, 6.60]	[5.98, 6.65]	[5.99, 6.66]	[5.97, 6.58]
	N	403,966	209,467	167,988	41,479

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

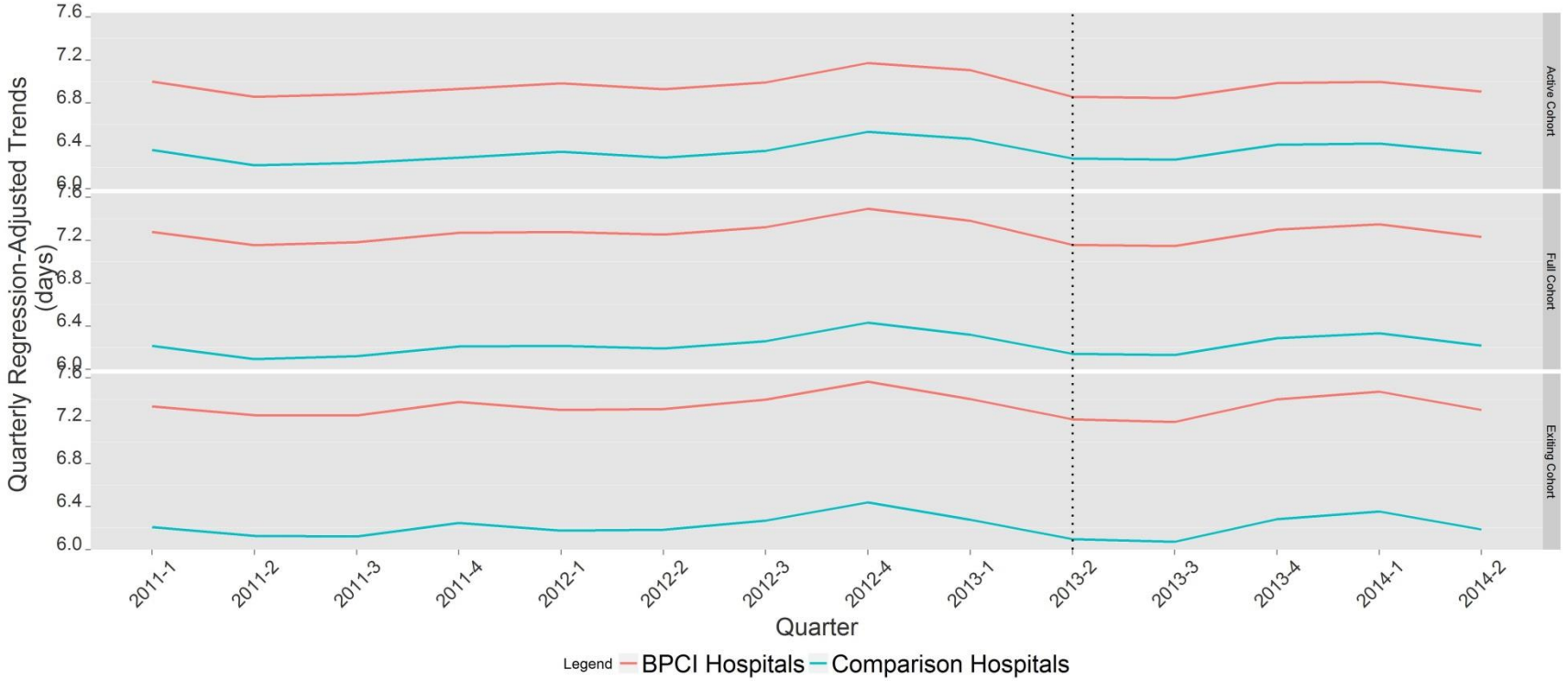
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 36: Since BPCI Inception DiD Estimates for Episode Length of Stay for Full, Active, and Exiting Cohort[+]

	Full Cohort	Active Cohort	Exiting Cohort
DiD (days)	-0.05	-0.06	-0.01
95% CI	(-0.12, 0.02)	(-0.15, 0.02)	(-0.11, 0.09)
Sample Size	1,789,437	1,087,174	827,451

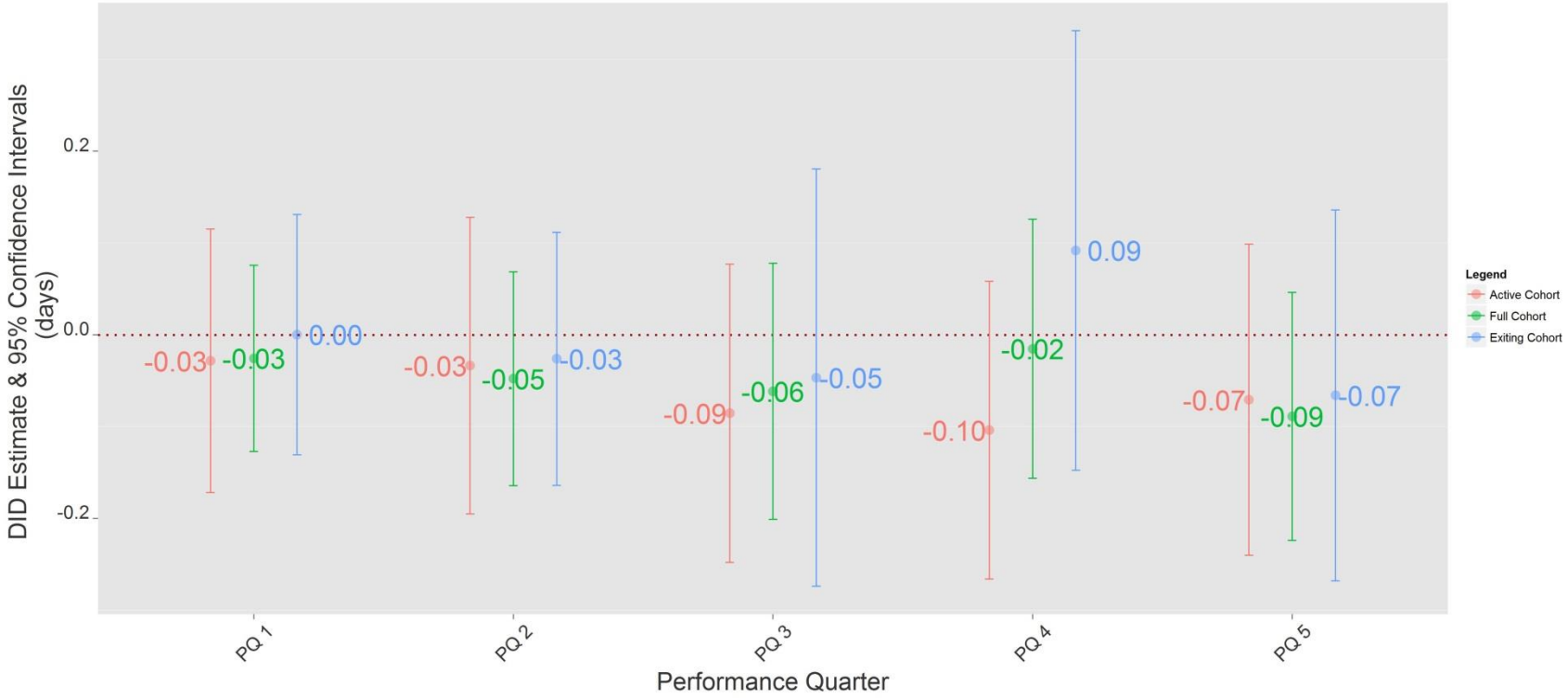
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 20: Adjusted Quarterly Trends for Episode Length of Stay for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 21: Performance Quarter DiD Estimates for Episode Length of Stay for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 37: Episode Weighted, Unadjusted and Adjusted Episode Length of Stay for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Expansive Cohort								
BPCI	**Mean (days)**	6.28	6.79	6.15***	6.63	6.16***	6.62	6.11***	6.67
	95% CI	[6.25, 6.32]	[6.64, 6.93]	[6.09, 6.21]	[6.47, 6.79]	[6.09, 6.23]	[6.46, 6.78]	[6.00, 6.21]	[6.49, 6.85]
	N	62,785		31,523		25,145		6,378	
Comparison	**Mean (days)**	6.11	6.01	6.19***	6.02	6.19***	6.01	6.19**	6.06
	95% CI	[6.09, 6.13]	[5.93, 6.10]	[6.17, 6.22]	[5.92, 6.12]	[6.16, 6.22]	[5.92, 6.10]	[6.13, 6.25]	[5.93, 6.19]
	N	251,960		130,432		104,241		26,191	
	Targeted Cohort								
BPCI	**Mean (days)**	6.71	7.49	6.75	7.46	6.76*	7.47	6.70	7.44
	95% CI	[6.69, 6.74]	[7.27, 7.70]	[6.71, 6.78]	[7.25, 7.67]	[6.72, 6.80]	[7.25, 7.68]	[6.63, 6.77]	[7.23, 7.65]
	N	220,110		117,710		93,654		24,056	
Comparison	**Mean (days)**	6.54	6.31	6.58***	6.31	6.59***	6.31	6.55	6.28
	95% CI	[6.53, 6.56]	[6.21, 6.40]	[6.56, 6.60]	[6.21, 6.40]	[6.57, 6.61]	[6.22, 6.40]	[6.51, 6.59]	[6.19, 6.37]
	N	668,503		348,605		279,659		68,946	
	PHC Cohort								
BPCI	**Mean (days)**	6.63	7.33	6.72***	7.36	6.71**	7.35	6.76**	7.37
	95% CI	[6.59, 6.66]	[7.12, 7.54]	[6.67, 6.76]	[7.16, 7.56]	[6.65, 6.76]	[7.16, 7.55]	[6.66, 6.86]	[7.18, 7.57]
	N	96,243		49,876		39,712		10,164	
Comparison	**Mean (days)**	6.43	6.22	6.50***	6.20	6.50***	6.20	6.50**	6.22
	95% CI	[6.41, 6.45]	[6.10, 6.33]	[6.47, 6.53]	[6.10, 6.30]	[6.47, 6.53]	[6.10, 6.30]	[6.44, 6.57]	[6.12, 6.31]
	N	271,554		140,080		112,235		27,845	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 38: Unweighted, Unadjusted Episode Length of Stay for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Expansive Cohort				
BPCI	**Mean (days)**	5.94	6.02**	6.08***	5.80***
	95% CI	[5.67, 6.21]	[5.70, 6.33]	[5.74, 6.41]	[5.54, 6.06]
	N	62,785	31,523	25,145	6,378
Comparison	**Mean (days)**	5.69	6.08***	6.14***	5.86***
	95% CI	[5.40, 5.98]	[5.78, 6.39]	[5.84, 6.45]	[5.57, 6.16]
	N	251,960	130,432	104,241	26,191
	Targeted Cohort				
BPCI	**Mean (days)**	6.60	6.61	6.61	6.58
	95% CI	[6.30, 6.90]	[6.30, 6.91]	[6.30, 6.92]	[6.28, 6.87]
	N	220,110	117,710	93,654	24,056
Comparison	**Mean (days)**	6.35	6.40***	6.41***	6.35
	95% CI	[6.03, 6.67]	[6.08, 6.72]	[6.09, 6.74]	[6.05, 6.65]
	N	668,503	348,605	279,659	68,946
	PHC Cohort				
BPCI	**Mean (days)**	6.64	6.70*	6.69	6.73
	95% CI	[6.38, 6.91]	[6.43, 6.96]	[6.41, 6.96]	[6.48, 6.98]
	N	96,243	49,876	39,712	10,164
Comparison	**Mean (days)**	6.34	6.41***	6.41***	6.39
	95% CI	[6.04, 6.63]	[6.10, 6.71]	[6.10, 6.72]	[6.10, 6.68]
	N	271,554	140,080	112,235	27,845

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

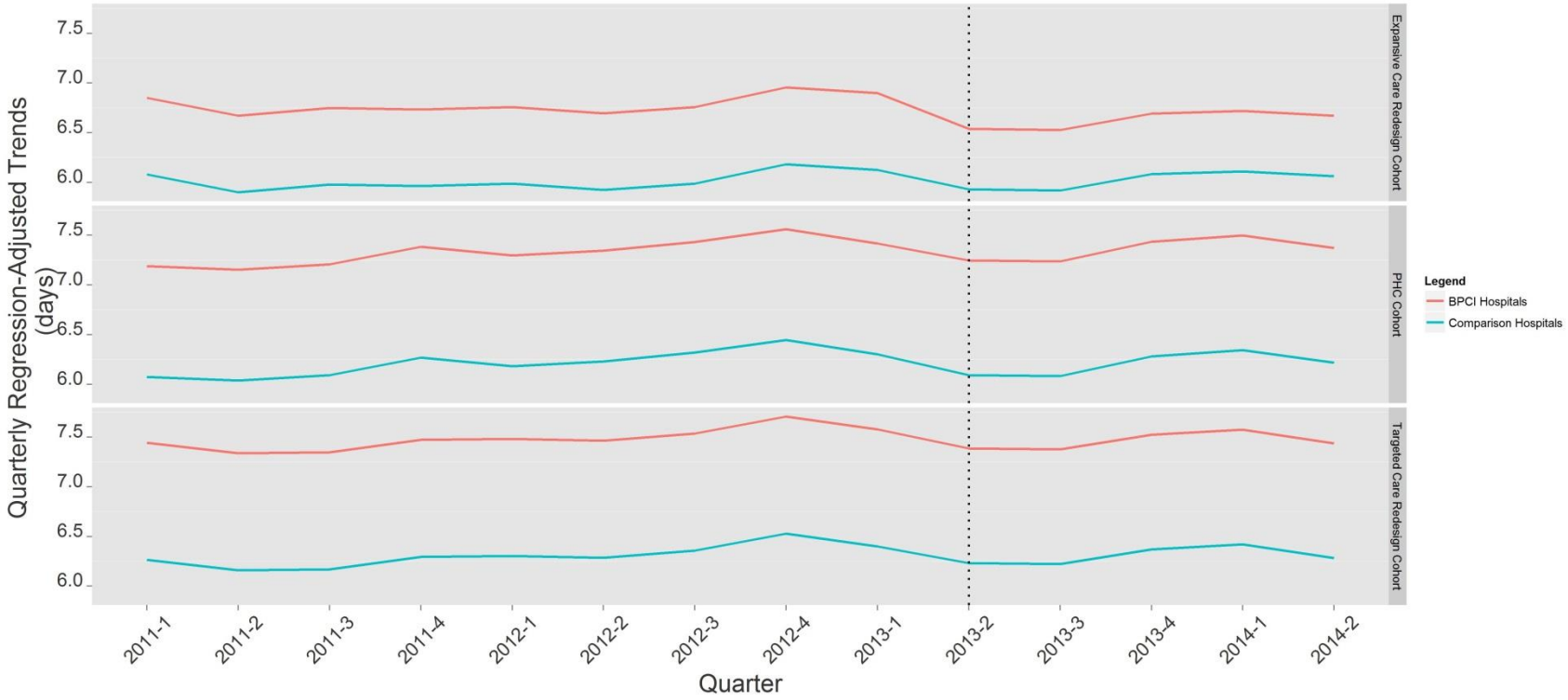
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 39: Since BPCI Inception DiD Estimates for Episode Length of Stay for Expansive, Targeted, and PHC Cohorts[+]

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD (days)	0.04	-0.16***	-0.02
95% CI	(-0.06, 0.14)	(-0.27, -0.05)	(-0.10, 0.06)
Sample Size	557,752	476,700	1,354,926

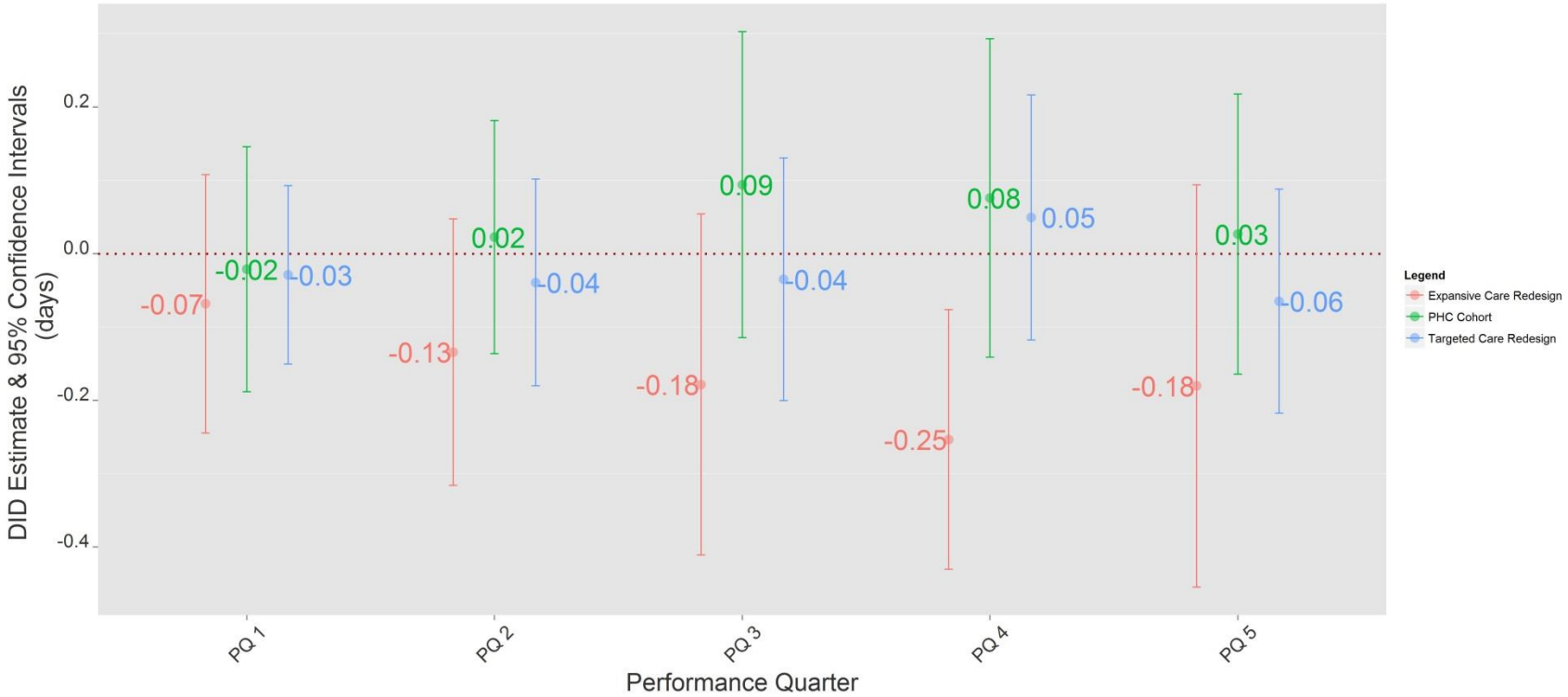
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 22: Adjusted Quarterly Trends for Episode Length of Stay for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 23: Performance Quarter DiD Estimates for Episode Length of Stay for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

III.B.3. Medicare Payments

A BPCI Model 1 design component, the IPPS discount, provides Medicare with risk-free savings on its payments to Awardee hospitals. This discount may lead to further savings to Medicare through Awardee efficiency gains in care delivery, initiated by the Awardee to engender internal cost savings that may offset discounted Medicare revenue.

Medicare payments[99] are assessed over two time periods, during and up to 30 days after an episode, by examining:

- Total Medicare episode payments and its components such as hospital and physician components during the episode.
- 30-day post-episode Medicare payments and preliminary analyses on post-episode care components, such as Medicare payments to and utilization of SNFs.

These payment measures are not standardized allowed amounts but actual payments from Medicare to providers.[100] Consequently, these payments include the Hospital Readmission Reduction Act payment adjustments that occurred during the five-quarter performance period. It is possible that the impact on Medicare payments may not be wholly attributable to BPCI Model 1. Future work will examine standardized allowed Medicare payment amounts that will provide a truer assessment of BPCI Model 1's effect on payments.

Measures in this domain are expressed as continuous dollar values that indicate average (total) Medicare payments over episode or post-episode periods. Appendix A provides measure specifications. Appendix B presents data for these and related measures.

The remainder of Section III.B.3 first provides a high-level overview of measures in the Medicare payment domain (Section III.B.3.1) and then presents individual measure result overviews, analysis, and data (Section III.B.3.1.1 and III.B.3.1.2). Findings focus on changes from Baseline to Since BPCI Inception periods for unadjusted and adjusted statistics and DiD impact estimates. Additionally, Awardee-level data may be presented to examine whether particular hospitals within a cohort are driving results.

III.B.3.1. Domain-Wide Overview of Results

On average, total Medicare episode payments for Awardee and comparison hospitals increased from Baseline to the Since BPCI Inception period. However, this increase was greater for comparison hospitals than for Awardee hospitals. DiD impact estimates indicate an average statistically significant decrease of \$123 ($p < 0.01$) in total Medicare episode payments over the model performance period relative to Baseline and comparison hospitals. Deeper analysis of this impact estimate shows that approximately \$83 ($p < 0.1$) is attributable to Medicare episode payments to hospitals (Table 72 in Appendix B) of which approximately \$30 may stem from the Model 1 IPPS discount on Medicare payments to Awardee hospitals. The remaining \$40 ($p < 0.01$) of the \$123 impact estimate stems from physician (or other provider) billing during the episode.

[99] Both payment measures are transformed via winsorising to limit extreme payment values in the 99th percentile.
[100] Standardized allowed amounts will be assessed in the next annual report.

Despite the overall *positivity* of this result—in that a statistically significant decrease to total Medicare episode payments is noted—a concerning aspect is that these results are driven by Awardees that terminated their Awardee Agreements (–$219, $p < 0.01$). Indeed, the hospital-only Medicare episode payment impact for Active cohort Awardees was not statistically significant at +$11, while the Exiting cohort Awardee impact was –$171 ($p < 0.05$). Interestingly, non-hospital Medicare episode payment impacts were relatively similar across Exiting and Active Awardees (~$40, $p < 0.01$; Table 78 in Appendix B).

The PHC cohort Awardees contributed to the Exiting cohort Awardee impact on total Medicare episode payments (–$152, $p < 0.05$). There was a large difference in impact estimates when stratifying Awardee hospitals by care redesign cohort, with the Targeted Care Redesign Awardees having an impact of –$161 ($p < 0.01$) on Medicare episode payments.

This differential between Active and Exiting Awardees is not unexpected. Exit interview data indicated that terminations were due, in part, to not *needing* Model 1 incentive components such as gainsharing (at least not at the expense of the IPPS discount). In other words, some of these terminating Awardees believed they had progressed in their endeavors to an extent where they need not "give back" a percent of their IPPS revenue. Further, six of the nine Exiting Awardees had recent experience from a similar CMS model, the PHC. Conversely, Active Awardees still believed they had enough to gain from this model in terms of clinician and hospital alignment. This sentiment may be indicative of a delay in the translation of Active Awardee activities to Model 1 goals.

Preliminary analysis of total post-episode Medicare payments indicates a less propitious Model 1 impact. As with Medicare payments over the episode, total post-episode Medicare payments (30-day window) increased from Baseline to the Since BPCI Inception period for Awardee and comparison hospital cohorts. These increases translated to an average (of total) post-episode Medicare payment impact estimate of +$95 ($p < 0.05$), relative to Baseline and comparison hospitals. Active Awardees influenced this program-wide result heavily with a post-episode Medicare payment impact estimate of +$248 ($p < 0.01$). Examination of unadjusted average Medicare payments across post-episode provider/service types for Awardee and comparison hospitals indicates that the adjusted DiD impact estimate of +$248 for Active cohorts may be primarily attributable to post-episode increases in Medicare payments to home health, skilled nursing, and other inpatient services/facilities over the model performance period, relative to Baseline and comparison hospitals. These relative payment increases for these facilities may be indicative of increased utilization or increased care intensity (and corresponding cost) or geographical payment differences (e.g., wage index differentials). Examination of unadjusted utilization rates for these facilities indicates no noteworthy changes from Baseline or relative to comparison hospital patients. Analysis of this measure is not complete and should be interpreted with extreme caution. This measure will be explored further in the 2015 Annual Report.

III.B.3.1.1. Total Medicare Episode Payments

III.B.3.1.1.1. Overview of Results

Table 40 presents total Medicare episode payments as unadjusted and adjusted averages in dollars for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. This table also includes these statistics for Active and Exiting cohorts.

Full cohort unadjusted statistics show an increase in average Medicare episode payments of $509 and $500 for Model 1 and comparison hospitals, respectively (Since BPCI Inception vs. Baseline; $p < 0.01$). After adjusting for patient and episode characteristics, these differentials reduced to increases of $85 and $209. Active cohort Awardee and comparison hospitals had lower increases in unadjusted means ($401 and $427, respectively) and respective adjusted differences of $159 and 184 when comparing Since BPCI Inception to Baseline. Exiting cohort Awardee and comparison hospitals had unadjusted average increases from Baseline of $582 and $588 per episode (Since BPCI Inception vs. Baseline; $p < 0.01$). Adjusted, these increases reduced to $12 and $231 for Awardee and comparison hospitals, respectively.

Table 41 presents unadjusted, unweighted average Medicare episode payments in dollars for Full, Active, and Exiting cohorts. The data indicate that some (Awardee) hospitals may be driving the unadjusted statistics in the Full and Active cohorts to differ from unadjusted, weighted statistics presented in Table 40. However, examination of hospital specific trends (not shown) indicates that these discrepancies are due to overall variation in payments across hospitals, not a subset of outlier Awardees.

Table 42 presents Full, Active, and Exiting cohort DiD estimates. These estimates are the impact estimates of BPCI Model 1 when compared to comparison hospitals for Baseline and Since BPCI Inception periods. The Full cohort estimate is an average reduction in Medicare episode payments of –$123 ($p < 0.01$). Subcohort analysis shows that estimate is driven by Exiting Awardees (–$219, $p < 0.01$), while Awardees still active had no statistically significant impacts on this measure. Figure 25 breaks these impact estimates down by PQ.

Table 43 presents total Medicare episode payments as unadjusted and adjusted averages in dollars for PHC, Expansive, and Targeted BPCI and comparison cohorts over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. Generally, Awardee and comparison hospitals across these cohorts show increases in average (total) Medicare episode payments in the Since BPCI Inception period relative to Baseline. Table 44 presents unadjusted, unweighted Medicare episode payments as averages for Expansive and Targeted Care Redesign and PHC cohorts. The data indicate that discrepancies between these unadjusted and unweighted averages and those in Table 43 are due to overall variation in payments across hospitals. . Table 45 presents impact estimates for these cohorts. Both the Targeted Care Redesign and PHC cohort show statistically significant impact decreases of $161 ($p < 0.01$) and $153 ($p < 0.05$), respectively. This result is not surprising, as all Exiting Awardees (and their comparison hospitals) are in the Targeted cohort and account for 56 percent of Awardees in the Targeted cohort. Further, the PHC cohort is composed solely from the Exiting cohort.

III.B.3.1.1.2. Analysis of Measure Findings – Total Medicare Episode Payments

Overall, the average Medicare episode payments increased over time. However, a reductive impact of $123 dollars per episode among Awardees is noted. This result stems from increases across comparison hospitals outpacing Awardee increases in this measure. It implies a decreased differential over time in Medicare episode payments relative to the counterfactual (i.e., what payments would have been in the absence of BPCI Model 1).

There is evidence that the model is reaching one goal of decreasing Medicare payments. Deeper analysis of this impact estimate shows that approximately $83 is attributable to Medicare payments to hospitals (Appendix B, Table 72), of which approximately $30 stems from the IPPS discount on Medicare payments to Awardee hospitals.[101] The remaining $40 (of the $123 impact) stems from physician (or other provider) billing during the episode.

Despite the overall *positivity* of this result—in that a statistically significant decrease to Medicare episode payments is noted—a concerning aspect is that these results are driven by Awardees that terminated from the model (–$219, $p < 0.01$). Indeed, hospital-only Medicare episode payment impacts for Active Awardees was not statistically significant at +$11, while the Exiting Awardee impact was at –$171 ($p < 0.05$). Interestingly, non-hospital Medicare episode payment impacts were relatively similar across Exiting and Active Awardees (Appendix B, Table 78).

The PHC cohort contributed to the Exiting cohort impact on Medicare episode payments (–$152, $p < 0.05$). There was a large difference in impact estimates when stratifying Awardee hospitals by care redesign cohort, with the Targeted Care Redesign Awardees having an impact of –$161 ($p < 0.01$) on Medicare episode payments.

This differential between Active and Exiting Awardees is not unexpected. Exit interview data indicated that terminations were due, in part, to not needing Model 1 incentivizing components such as gainsharing, at least not at the expense of the IPPS discount. In other words, some of these terminating Awardees believed they had progressed in their endeavors to an extent where they need not "give back" a percent of their IPPS revenue. Further, six of the nine Exiting Awardees had experience from a similar, recent CMS model (PHC Awardees). Conversely, Active Awardees still believed they had enough to gain from this model in terms of clinician and hospital alignment to remain active in BPCI Model 1.

[101] This estimate comes from MAC processing instructions that apply the IPPS discount to Medicare IPPS operating payment portions.

Table 40: Episode Weighted, Unadjusted and Adjusted Total Medicare Episode Payments Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean $)**	12,925	13,401	13,434***	13,486	13,391***	13,464	13,600***	13,576**
	95% CI	[12,875, 12,974]	[13,307, 13,494]	[13,364, 13,503]	[13,376, 13,597]	[13,313, 13,469]	[13,355, 13,573]	[13,447, 13,754]	[13,460, 13,691]
	N	270,986		143,008		113,822		29,186	
Comparison	**Mean ($)**	12,329	11,370	12,829***	11,579***	12,808***	11,556***	12,913***	11,668***
	95% CI	[12,301, 12,356]	[11,308, 11,432]	[12,791, 12,867]	[11,501, 11,657]	[12,766, 12,851]	[11,482, 11,631]	[12,831, 12,995]	[11,579, 11,758]
	N	858,480		447,497		358,458		89,039	
	Active Cohort								
BPCI	**Mean ($)**	12,340	13,209	12,741***	13,368	12,726***	13,353	12,798***	13,425*
	95% CI	[12,274, 12,406]	[13,072, 13,346]	[12,645, 12,836]	[13,212, 13,523]	[12,620, 12,833]	[13,202, 13,504]	[12,586, 13,011]	[13,254, 13,596]
	N	138,045		70,833		56,550		14,283	
Comparison	**Mean ($)**	12,551	11,529	12,978***	11,713***	12,973***	11,699***	12,998***	11,771***
	95% CI	[12,517, 12,586]	[11,448, 11,609]	[12,929, 13,026]	[11,617, 11,809]	[12,918, 13,028]	[11,609, 11,789]	[12,896, 13,099]	[11,651, 11,891]
	N	549,755		287,143		229,963		57,180	
	Exiting Cohort								
BPCI	**Mean ($)**	13,532	11,068	14,114***	11,080	14,048***	11,048	14,369***	11,208
	95% CI	[13,458, 13,606]	[10,886, 11,250]	[14,013, 14,215]	[10,876, 11,284]	[13,934, 14,161]	[10,844, 11,251]	[14,150, 14,588]	[11,000, 11,416]
	N	132,941		72,175		57,272		14,903	
Comparison	**Mean ($)**	12,482	12,498	13,070***	12,729***	13,042***	12,697**	13,186***	12,857***
	95% CI	[12,439, 12,525]	[12,397, 12,598]	[13,009, 13,131]	[12,606, 12,852]	[12,972, 13,111]	[12,575, 12,819]	[13,057, 13,315]	[12,732, 12,982]
	N	388,147		201,531		161,579		39,952	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 41: Unweighted, Unadjusted Total Medicare Episode Payments for Full, Active, and Exiting Cohorts [+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Full Cohort				
BPCI	**Mean ($)**	12,420	12,762***	12,737***	12,858***
	95% CI	[11,754, 13,086]	[12,070, 13,453]	[12,040, 13,433]	[12,186, 13,530]
	N	270,986	143,008	113,822	29,186
Comparison	**Mean ($)**	11,939	12,431***	12,406***	12,526***
	95% CI	[11,216, 12,662]	[11,719, 13,142]	[11,694, 13,117]	[11,814, 13,238]
	N	858,480	447,497	358,458	89,039
	Active Cohort				
BPCI	**Mean ($)**	11,711	12,017***	12,006***	12,062***
	95% CI	[11,059, 12,364]	[11,330, 12,705]	[11,313, 12,698]	[11,392, 12,732]
	N	138,045	70,833	56,550	14,283
Comparison	**Mean ($)**	12,040	12,460***	12,448***	12,505***
	95% CI	[11,320, 12,760]	[11,772, 13,148]	[11,762, 13,134]	[11,810, 13,201]
	N	549,755	287,143	229,963	57,180
	Exiting Cohort				
BPCI	**Mean ($)**	13,627	13,953***	13,894***	14,185***
	95% CI	[12,939, 14,315]	[13,255, 14,651]	[13,191, 14,598]	[13,509, 14,862]
	N	132,941	72,175	57,272	14,903
Comparison	**Mean ($)**	12,177	12,762***	12,730***	12,890***
	95% CI	[11,451, 12,903]	[12,006, 13,518]	[11,968, 13,492]	[12,157, 13,624]
	N	388,147	201,531	161,579	39,952

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

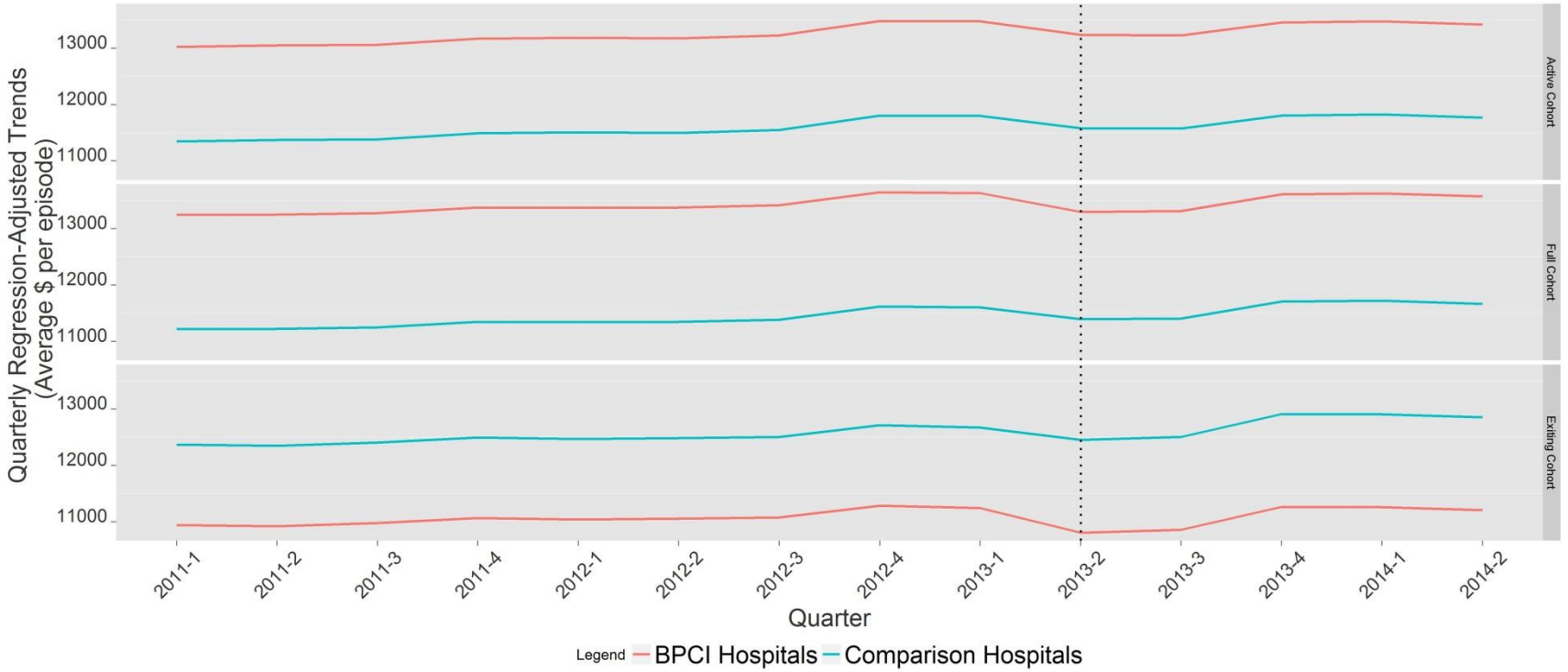
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 42: Since BPCI Inception DiD Estimates Total Medicare Episode Payments for Full, Active, and Exiting Cohort+

	Full Cohort	Active Cohort	Exiting Cohort
DiD ($)	-123.10***	-26.14	-219.26***
95% CI	(-213.90, -32.31)	(-134.70, 82.43)	(-355.21, -83.30)
Sample Size	1,685,572	1,024,861	778,899

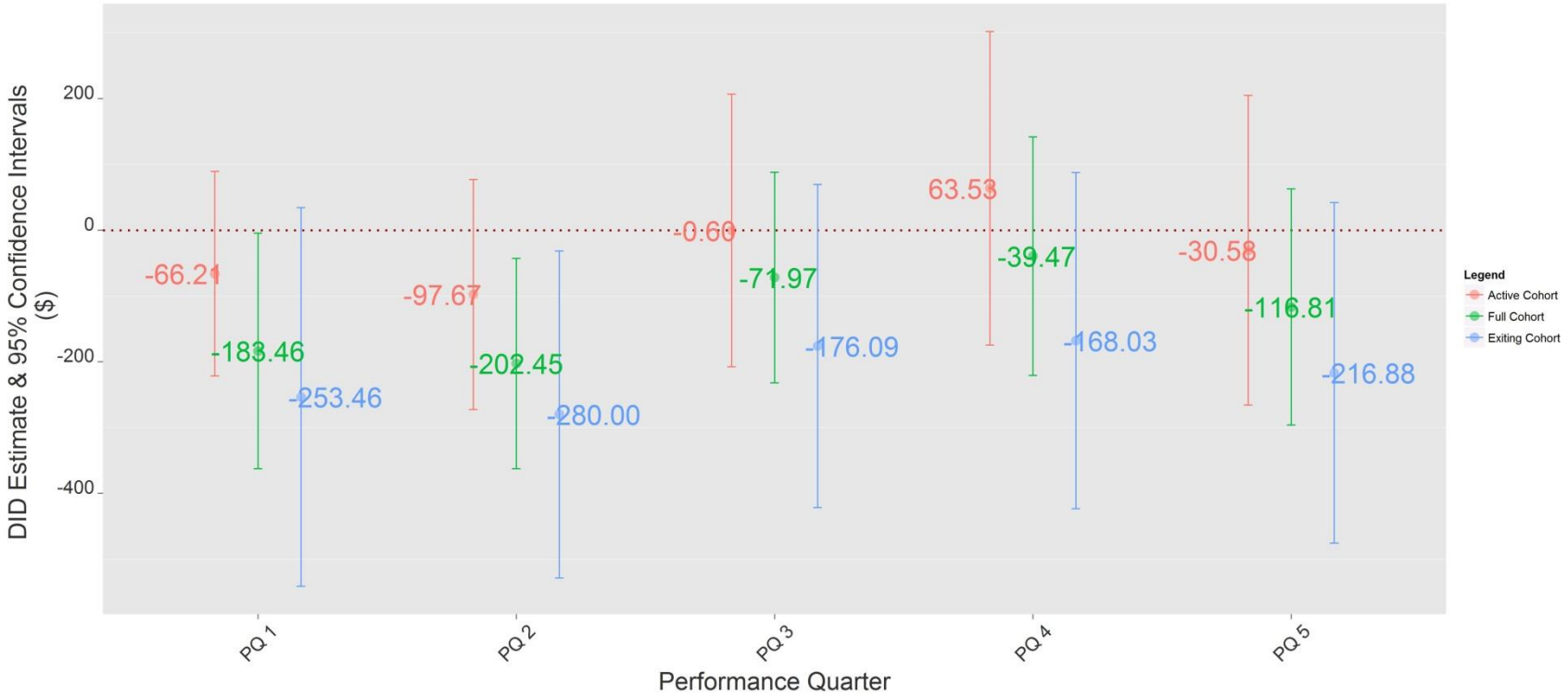
+ Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 24: Adjusted Quarterly Trends for Total Medicare Episode Payment for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 25: Performance Quarter DiD Estimates for Total Medicare Episode Payment for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 43: Episode Weighted, Unadjusted and Adjusted Total Medicare Episode Payments for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Expansive Cohort								
BPCI	**Mean ($)**	10,735	12,353	10,952***	12,463	10,958***	12,453	10,927*	12,505
	95% CI	[10,662, 10,808]	[12,225, 12,482]	[10,850, 11,054]	[12,298, 12,629]	[10,841, 11,075]	[12,289, 12,617]	[10,726, 11,128]	[12,333, 12,677]
	N	60,363		30,298		24,159		6,139	
Comparison	**Mean ($)**	11,326	10,337	11,740***	10,502**	11,734***	10,492**	11,761***	10,543**
	95% CI	[11,284, 11,368]	[10,243, 10,431]	[11,682, 11,798]	[10,382, 10,622]	[11,669, 11,800]	[10,375, 10,608]	[11,636, 11,886]	[10,408, 10,679]
	N	242,999		125,624		100,333		25,291	
	Targeted Cohort								
BPCI	**Mean ($)**	13,552	11,203	14,101***	11,277	14,047***	11,252	14,312***	11,377
	95% CI	[13,492, 13,612]	[11,030, 11,375]	[14,018, 14,184]	[11,093, 11,462]	[13,953, 14,140]	[11,069, 11,435]	[14,127, 14,498]	[11,186, 11,568]
	N	210,623		112,710		89,663		23,047	
Comparison	**Mean ($)**	12,698	12,603	13,226***	12,839***	13,200***	12,814***	13,334***	12,939***
	95% CI	[12,665, 12,732]	[12,515, 12,692]	[13,180, 13,272]	[12,738, 12,940]	[13,147, 13,252]	[12,716, 12,912]	[13,235, 13,434]	[12,826, 13,051]
	N	642,471		335,544		269,103		66,441	
	PHC Cohort								
BPCI	**Mean ($)**	11,929	9,845	12,343***	9,848	12,282***	9,825	12,585***	9,943
	95% CI	[11,856, 12,001]	[9,661, 10,029]	[12,242, 12,445]	[9,637, 10,060]	[12,168, 12,396]	[9,616, 10,033]	[12,358, 12,811]	[9,720, 10,167]
	N	92,083		47,831		38,093		9,738	
Comparison	**Mean ($)**	11,338	11,399	11,985***	11,555*	11,940***	11,532	12,164***	11,650***
	95% CI	[11,292, 11,384]	[11,290, 11,509]	[11,921, 12,049]	[11,411, 11,700]	[11,869, 12,012]	[11,390, 11,673]	[12,020, 12,307]	[11,495, 11,806]
	N	260,565		134,597		107,823		26,774	

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 44: Unweighted, Unadjusted Total Medicare Episode Payments for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline	Since BPCI Inception	BPCI – Year 1	BPCI – PQ 5
	Expansive Cohort				
BPCI	**Mean ($)**	10,742	10,811	10,813	10,802
	95% CI	[10,208, 11,276]	[10,224, 11,397]	[10,207, 11,419]	[10,286, 11,319]
	N	60,363	30,298	24,159	6,139
Comparison	**Mean ($)**	11,370	11,702***	11,674***	11,806***
	95% CI	[10,644, 12,095]	[11,028, 12,377]	[11,008, 12,339]	[11,102, 12,510]
	N	242,999	125,624	100,333	25,291
	Targeted Cohort				
BPCI	**Mean ($)**	13,294	13,664***	13,608***	13,886***
	95% CI	[12,559, 14,028]	[12,924, 14,404]	[12,871, 14,346]	[13,136, 14,636]
	N	210,623	112,710	89,663	23,047
Comparison	**Mean ($)**	12,281	12,812***	12,782***	12,928***
	95% CI	[11,559, 13,003]	[12,083, 13,540]	[12,050, 13,515]	[12,215, 13,642]
	N	642,471	335,544	269,103	66,441
	PHC Cohort				
BPCI	**Mean ($)**	11,657	11,950***	11,902***	12,142***
	95% CI	[11,133, 12,181]	[11,412, 12,489]	[11,359, 12,446]	[11,623, 12,661]
	N	92,083	47,831	38,093	9,738
Comparison	**Mean ($)**	10,978	11,593***	11,556***	11,740***
	95% CI	[10,395, 11,561]	[10,974, 12,211]	[10,934, 12,178]	[11,136, 12,344]
	N	260,565	134,597	107,823	26,774

[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims. Unlike weighted statistics, cohort level statistics consider all hospitals as having equal weight.

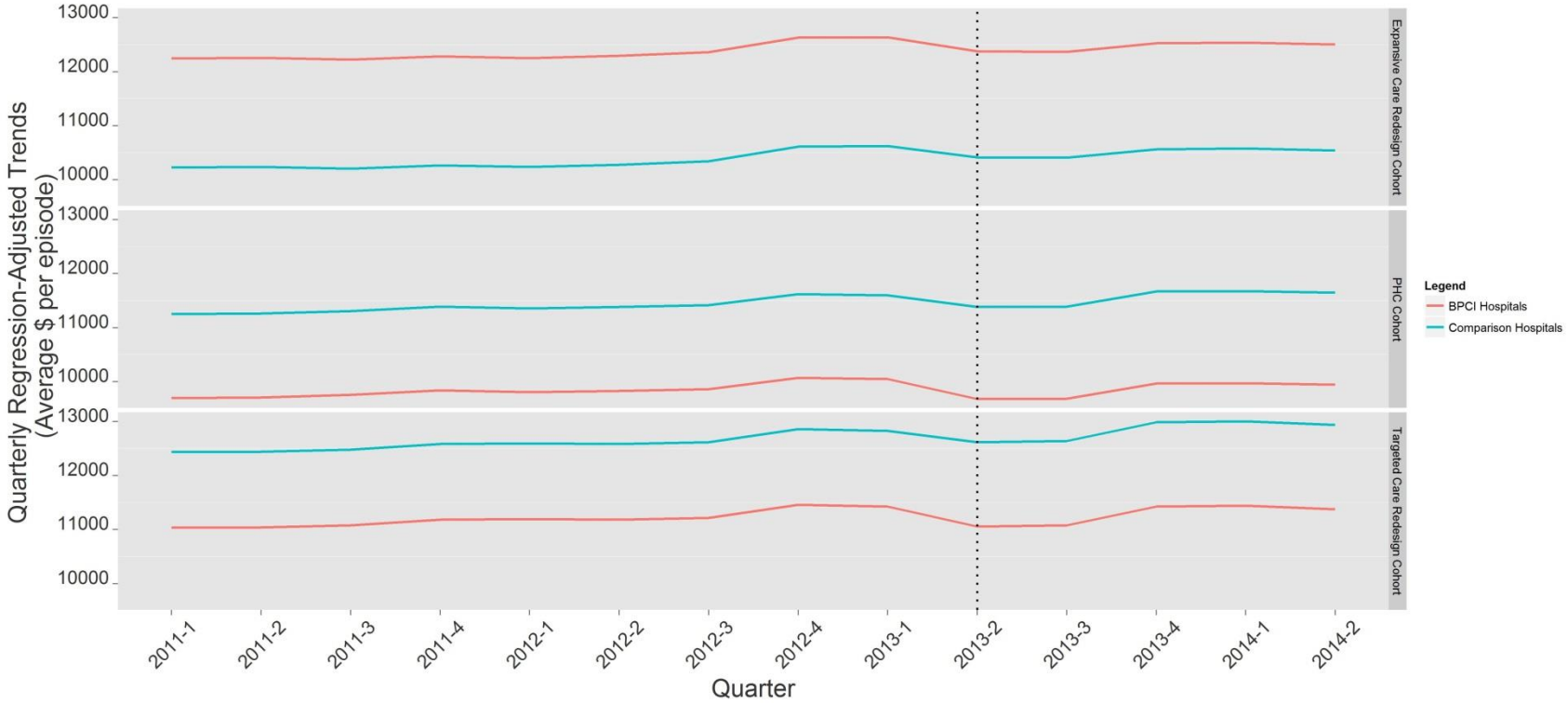
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 45: Since BPCI Inception DiD Estimates for Total Medicare Episode Payments for Expansive, Targeted, and PHC Cohorts+

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD ($)	-152.66**	-55.32	-160.89***
95% CI	(-289.79, -15.52)	(-216.18, 105.54)	(-265.58, -56.21)
Sample Size	524,374	450,099	1,275,320

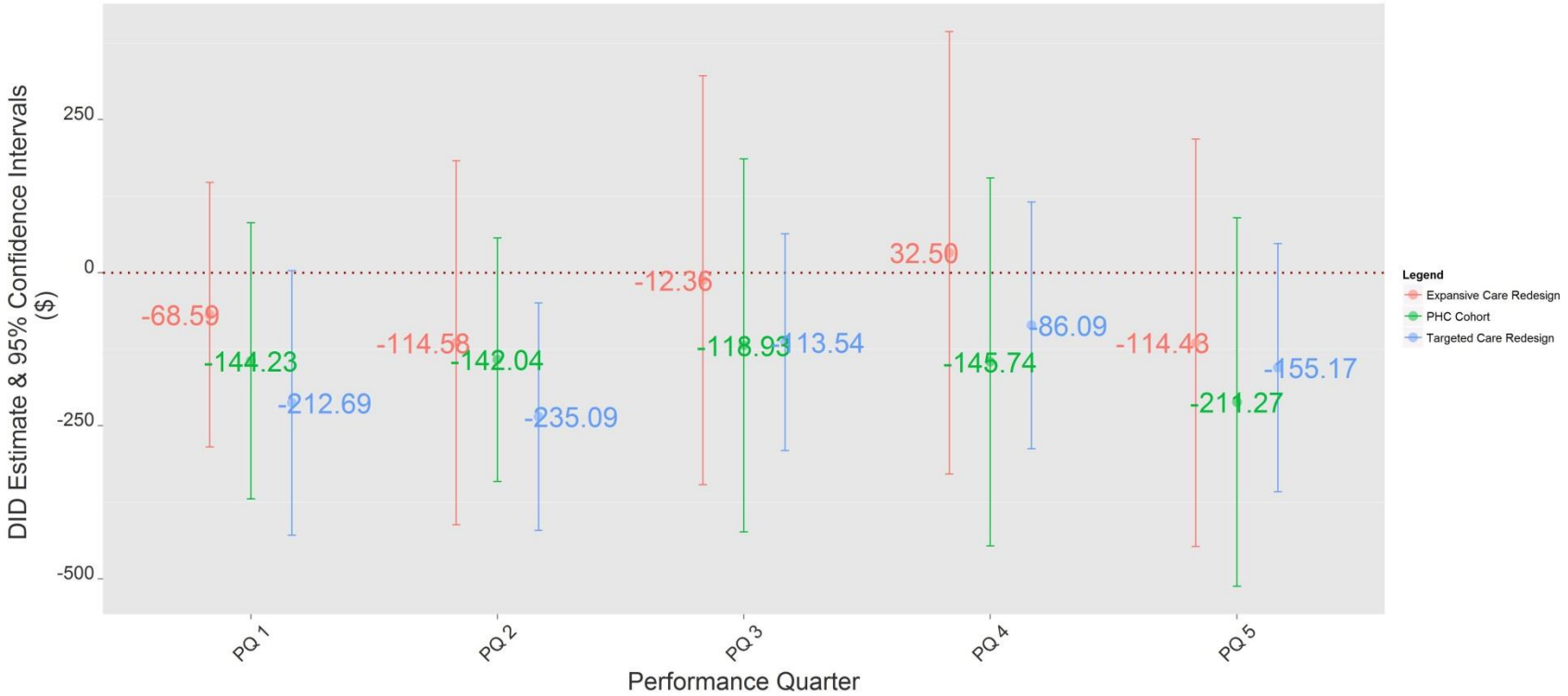
+ Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 26: Adjusted Quarterly Trends for Total Medicare Episode Payment for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 27: Performance Quarter DiD Estimates for Total Medicare Episode Payment for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

III.B.3.1.2. 30-Day Post-Episode Medicare Payments

III.B.3.1.2.1. Overview of Results

Table 46 presents total Medicare payments paid within 30 days after an episode as unadjusted and adjusted averages for Full cohort Awardee and comparison hospitals over Baseline, Since BPCI Inception, BPCI – Year 1, and BPCI – PQ 5. This table also includes these payments for Active and Exiting cohorts. Full cohort unadjusted averages show an increase in post-episode Medicare payments of $266 and $126 for Model 1 and comparison hospitals, respectively (Since BPCI Inception vs. Baseline; $p < 0.01$). After adjusting for patient and episode characteristics, these differentials change to increases of $264 and $169. Active cohort unadjusted averages show an increase in post-episode Medicare payments of $358 and $105 for Model 1 and comparison hospitals, respectively (Since BPCI Inception vs. Baseline; $p < 0.01$), with adjusted averages respectively showing increases of $395 and $147 when comparing Since BPCI Inception to Baseline. Exiting cohort unadjusted averages show an increase in post-episode Medicare payments of $172 and $162 for Model 1 and comparison hospitals, respectively (Since BPCI Inception vs. Baseline; $p < 0.01$), with adjusted averages showing respective increases of $150 and $196 when comparing Since BPCI Inception to Baseline.

Table 47 presents Full, Active, and Exiting cohort DiD estimates. These estimates are the impact estimates of BPCI Model 1 when compared to comparison hospitals for Baseline and Since BPCI Inception periods. The Full cohort estimate indicates an average increase in post-episode Medicare payments of $95 ($p < 0.05$). Subcohort analysis shows that estimate is driven by Active Awardees ($248, $p < 0.01$), while Exiting Awardees had no statistically significant impacts on this measure. Figure 30 breaks these estimates down by PQ and shows that Active Awardee impacts for this measure increased substantially in PQ 3 ($273, $p < 0.05$) to its highest point over these five PQs in PQ 4 ($472, $p < 0.05$) to a relatively lesser impact of $431 ($p < 0.05$) in PQ 5.

Tables 48 and 49 and Figure 32 present elevated changes and impact estimates for Expansive Awardees. This is expected, given Expansive Awardees are also Active Awardees.

III.B.3.1.2.2. Analysis of Measure Findings – 30-Day Post-Episode Medicare Payments

Full cohort analysis of unadjusted and adjusted 30-day post-episode payments indicates increases for Awardee and comparison hospitals from Baseline to Since BPCI Inception periods. Full cohort DiD results show that BPCI Model 1 is associated with a $95 ($p < 0.05$) increase in post-episode Medicare payments. Active Awardees influenced this program-wide result heavily with a post-episode Medicare payment impact of +$248 ($p < 0.01$).

Additional, preliminary analyses provide some insight. Examination of unadjusted average Medicare payments across post-episode provider/service types for Awardee and comparison hospitals indicates that the adjusted DiD impact estimate of +$248 for Active cohorts (unadjusted DiD impact +$253, Table 46) may be primarily attributable to HHA, SNF, and non-Awardee inpatient facilities.[102] Unadjusted DiD impacts for these facilities are +$37.51, +$59.18, and +$95.57, respectively.

[102] This includes rehabilitation or long-term care hospitals and excludes non-BPCI Model 1 Awardee or comparison hospitals.

Further, examination of the unadjusted average of any and all post-episode provider utilization indicates that the average number of post-episode Medicare claims, for those that had at least one post-episode claim, varied by less than half a claim on average between Baseline and Since BPCI Inception across all cohorts for Awardee and comparison hospitals. For example, average number of all Medicare claims during the post-episode period for Active cohort Awardees was 13.33 for the Baseline period and 13.55 over the Since BPCI Inception period, and Active cohort comparison hospitals had averages of 12.18 and 12.29 for these periods, respectively.[103] Utilization counts at the provider/service type level have comparable changes (in magnitude).

Taken together, a possible explanation is that minimal changes in utilization across service types are not driving the elevated Active cohort impact estimate for this measure but that resource intensity for some of these services (e.g., rehabilitation hospital or SNF) is higher, on average, for patients discharged from Awardee hospital and leading to higher average payments. This is not reported as a finding but as an avenue to explore for the 2015 Annual Report. Further, interpretation is cautioned, as this measure is composed of Medicare payments whose calculation varies across provider/service types and that are not standardized allowed amounts.

[103] The overwhelming majority of these services came from (clinician) provider and supplier services.

Table 46: Episode Weighted, Unadjusted and Adjusted 30-Day Post-Episode Medicare Payments for Full, Active, and Exiting Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Full Cohort								
BPCI	**Mean ($)**	9,747	9,710	10,013***	9,974**	10,006***	9,978**	10,041***	9,958**
	95% CI	[9,697, 9,796]	[9,553, 9,867]	[9,945, 10,082]	[9,806, 10,142]	[9,929, 10,083]	[9,810, 10,146]	[9,888, 10,194]	[9,788, 10,128]
	N	269,434		142,282		113,047		29,235	
Comparison	**Mean ($)**	8,531	7,787	8,657***	7,956***	8,696***	7,960***	8,503	7,940**
	95% CI	[8,504, 8,557]	[7,708, 7,866]	[8,622, 8,693]	[7,876, 8,036]	[8,656, 8,736]	[7,882, 8,038]	[8,426, 8,579]	[7,851, 8,030]
	N	852,081		444,136		355,440		88,696	
	Active Cohort								
BPCI	**Mean ($)**	9,653	9,479	10,011***	9,874***	10,002***	9,889***	10,046***	9,813**
	95% CI	[9,583, 9,724]	[9,295, 9,664]	[9,911, 10,111]	[9,671, 10,077]	[9,891, 10,114]	[9,687, 10,092]	[9,816, 10,276]	[9,607, 10,019]
	N	137,098		70,389		56,090		14,299	
Comparison	**Mean ($)**	8,750	8,070	8,855***	8,217**	8,910***	8,233**	8,636**	8,156
	95% CI	[8,717, 8,783]	[7,977, 8,163]	[8,810, 8,901]	[8,121, 8,314]	[8,859, 8,962]	[8,140, 8,325]	[8,539, 8,733]	[8,043, 8,269]
	N	545,364		285,128		227,891		57,237	
	Exiting Cohort								
BPCI	**Mean ($)**	9,843	9,754	10,015***	9,904	10,010***	9,894	10,036*	9,947
	95% CI	[9,774, 9,912]	[9,558, 9,949]	[9,921, 10,109]	[9,698, 10,110]	[9,904, 10,115]	[9,688, 10,099]	[9,832, 10,239]	[9,738, 10,155]
	N	132,336		71,893		56,957		14,936	
Comparison	**Mean ($)**	8,169	7,438	8,331***	7,634**	8,352***	7,623**	8,246	7,676***
	95% CI	[8,131, 8,207]	[7,331, 7,545]	[8,279, 8,384]	[7,525, 7,742]	[8,293, 8,412]	[7,517, 7,729]	[8,135, 8,357]	[7,559, 7,794]
	N	385,388		199,829		160,323		39,506	

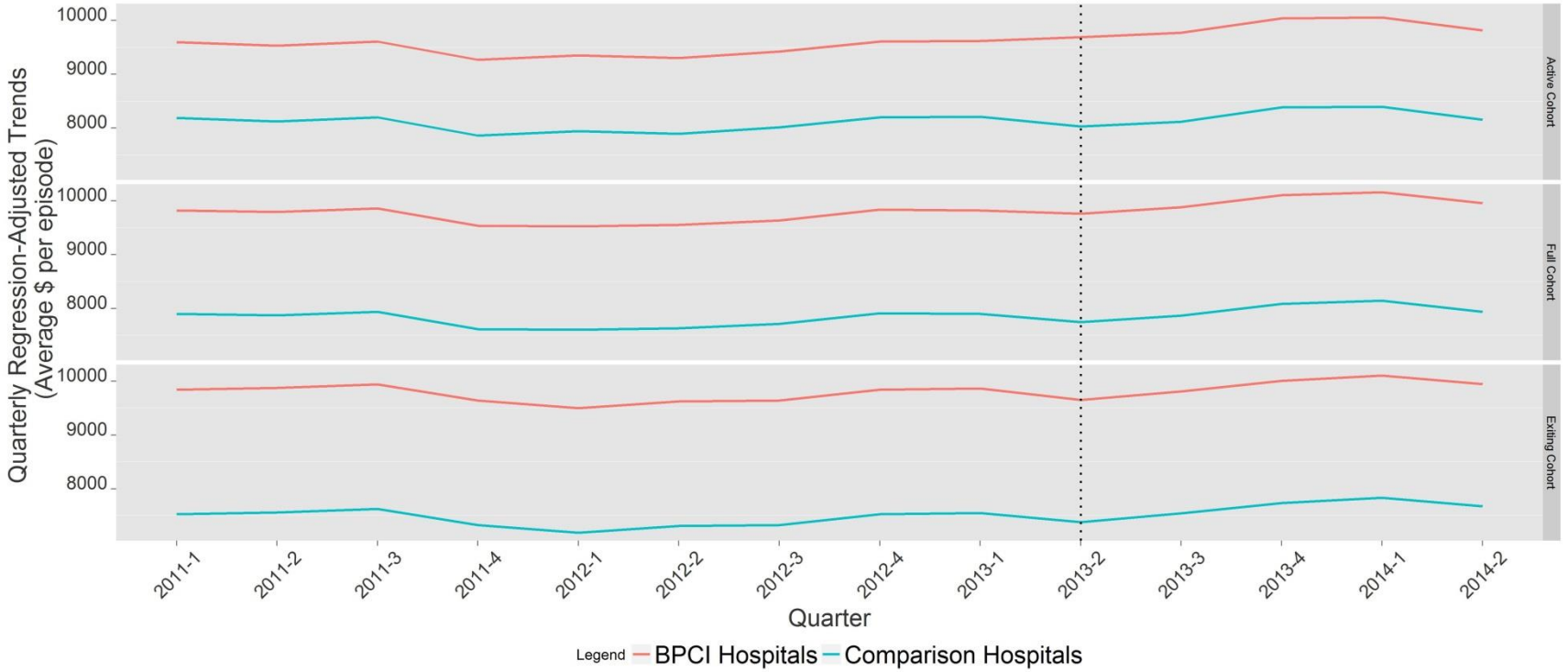
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 47: Since BPCI Inception DiD Estimates for 30-Day Post-Episode Medicare Payments for Full, Active, and Exiting Cohort[+]

	Full Cohort	Active Cohort	Exiting Cohort
DiD ($)	94.81**	247.63***	-45.00
95% CI	(5.04, 184.58)	(115.16, 380.10)	(-153.65, 63.65)
Sample Size	1,658,377	1,006,989	766,557

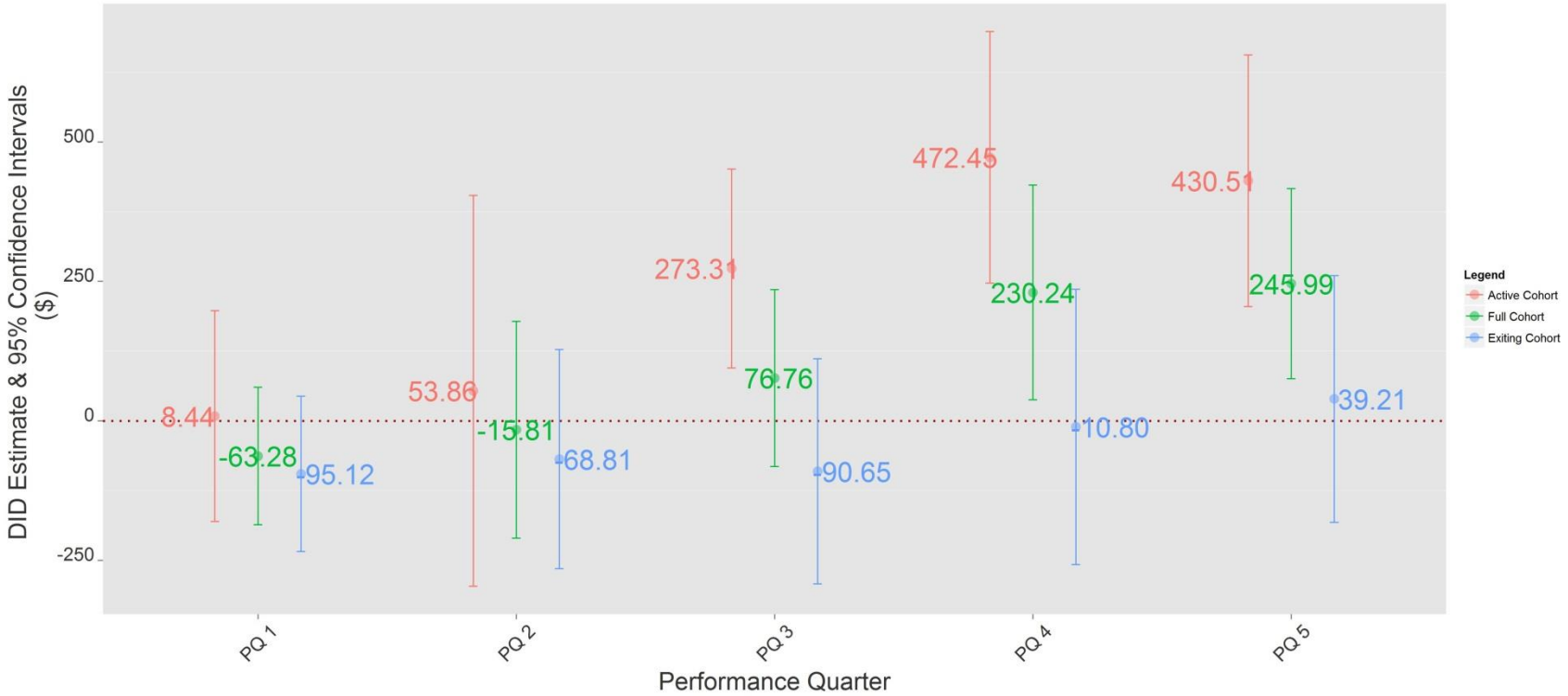
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 28: Adjusted Quarterly Trends for 30-Day Post-Episode Medicare Payments for Full, Active, and Exiting Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 29: Performance Quarter DiD Estimates for 30-Day Post-Episode Medicare Payments for Full, Active, and Exiting Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Table 48: Episode Weighted, Unadjusted and Adjusted 30-Day Post-Episode Medicare Payments for Expansive, Targeted, and PHC Cohorts[+]

Hospital Cohort	Measure	Baseline		Since BPCI Inception		BPCI – Year 1		BPCI – PQ 5	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
	Expansive Cohort								
BPCI	**Mean ($)**	8,944	8,324	9,229***	8,552	9,211***	8,572	9,300*	8,474
	95% CI	[8,849, 9,039]	[8,110, 8,538]	[9,087, 9,371]	[8,306, 8,798]	[9,058, 9,364]	[8,330, 8,814]	[8,944, 9,657]	[8,211, 8,737]
	N	60,007		30,124		23,983		6,141	
Comparison	**Mean ($)**	8,414	7,906	8,524***	8,042	8,600***	8,062*	8,224**	7,964
	95% CI	[8,367, 8,461]	[7,785, 8,027]	[8,459, 8,588]	[7,915, 8,169]	[8,526, 8,674]	[7,937, 8,187]	[8,088, 8,361]	[7,828, 8,100]
	N	241,140		124,831		99,435		25,396	
	Targeted Cohort								
BPCI	**Mean ($)**	9,977	10,125	10,224***	10,399*	10,220***	10,398*	10,238***	10,401*
	95% CI	[9,919, 10,034]	[9,932, 10,318]	[10,145, 10,302]	[10,195, 10,603]	[10,132, 10,308]	[10,195, 10,602]	[10,068, 10,407]	[10,196, 10,607]
	N	209,427		112,158		89,064		23,094	
Comparison	**Mean ($)**	8,620	7,779	8,748***	7,960**	8,774***	7,959***	8,643	7,962**
	95% CI	[8,589, 8,651]	[7,682, 7,876]	[8,706, 8,790]	[7,862, 8,057]	[8,726, 8,821]	[7,864, 8,054]	[8,552, 8,733]	[7,855, 8,070]
	N	637,686		332,937		266,898		66,039	
	PHC Cohort								
BPCI	**Mean ($)**	9,924	9,550	10,029	9,623	10,044	9,618	9,971	9,644
	95% CI	[9,844, 10,003]	[9,359, 9,740]	[9,916, 10,141]	[9,422, 9,824]	[9,917, 10,170]	[9,420, 9,815]	[9,728, 10,214]	[9,431, 9,857]
	N	91,710		47,693		37,906		9,787	
Comparison	**Mean ($)**	7,765	7,196	7,904***	7,357*	7,915***	7,351*	7,858	7,377**
	95% CI	[7,723, 7,808]	[7,081, 7,311]	[7,845, 7,962]	[7,241, 7,472]	[7,849, 7,981]	[7,240, 7,462]	[7,732, 7,984]	[7,246, 7,509]
	N	258,944		133,415		107,114		26,301	

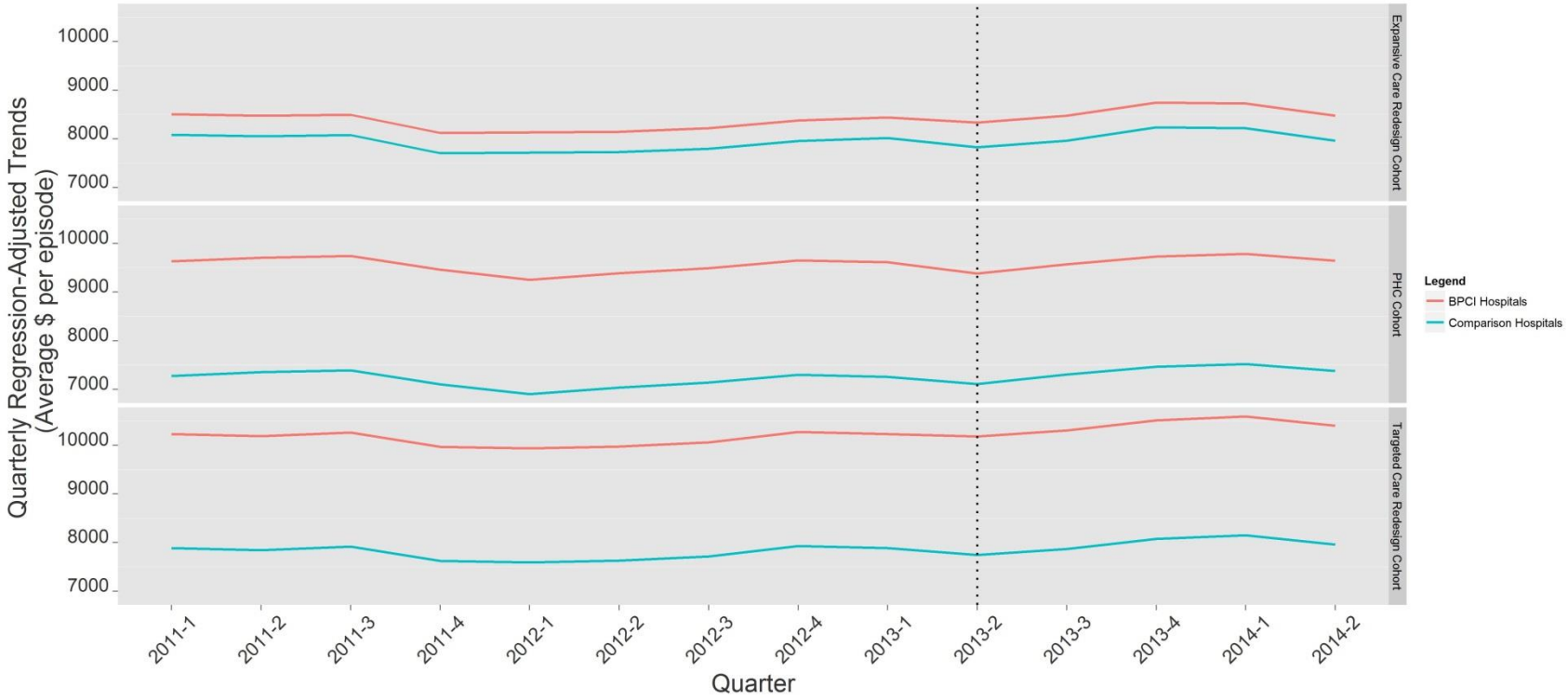
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels when program periods (Since BPCI Inception, Year 1, and PQ 5) are each compared to Baseline for adjusted or unadjusted statistics.

Table 49: Since BPCI Inception DiD Estimates for 30-Day Post-Episode Medicare Payments for Expansive, Targeted, and PHC Cohorts[+]

	PHC Cohort	Expansive Care Redesign	Targeted Care Redesign
DiD ($)	-86.85	92.64	92.83*
95% CI	(-206.01, 32.32)	(-99.94, 285.22)	(-10.14, 195.80)
Sample Size	517,453	443,166	1,254,603

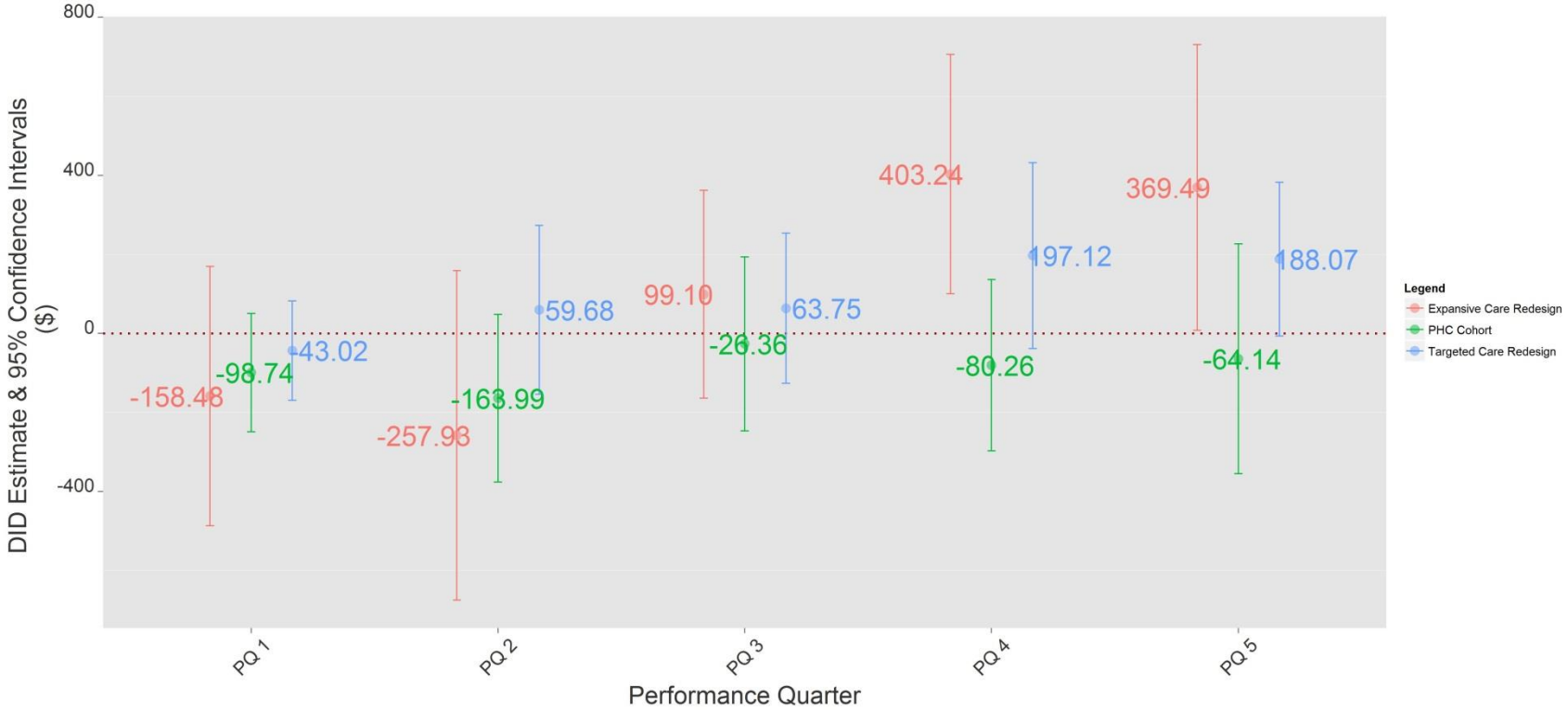
[+] Table 6 details Baseline, Since BPCI Inception, Year 1, and PQ 5 timeframes. Table 10 details cohort membership. Data source: Medicare claims.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ denote statistical significances levels.

Figure 30: Adjusted Quarterly Trends for 30-Day Post-Episode Medicare Payments for PHC, Expansive, and Targeted Cohorts*



* Dotted line indicates BPCI Model 1 start date. Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included, from January 1, 2011, through June 30, 2014.

Figure 31: Performance Quarter DiD Estimates for 30-Day Post-Episode Medicare Payments for PHC, Expansive, and Targeted Cohorts*



* Table 10 details cohort membership. Data source: Medicare claims. Fourteen quarters of data are included in regression, from January 1, 2011, through June 30, 2014.

Section IV. Discussion

The longevity of BPCI Model 1 relies on an acceptable level of Awardee participation, and its success relies on in-hospital clinician and non-clinician support. As previously noted, nine Awardee hospitals terminated their Awardee Agreement for the period of the analysis covered in this report, through June 30, 2014. As of March 2015, 3 additional Awardees have terminated their Awardee Agreements with CMS for a total of 12 of 24 Awardee withdrawals. Terminating Awardees repeatedly noted the financial burden of the IPPS discount as a motivating factor for termination. Some of these Awardees also noted difficulties in attributing internal hospital cost savings (or even realizing any) to BPCI Model 1 such that the IPPS discount was partially or fully offset. Many of the initial terminating Awardees did believe they had achieved sufficient progress in their care redesign implementation to proceed without model incentive components.

Model 1 does presume that enrolled practitioners—those with the potential to receive gainsharing from Awardees—are incentivized by gainsharing to carry out care redesigns and help affect efficiency gains at a hospital. Care redesigns, however, are not solely implemented by enrolled practitioners and are typically hospital-wide or multidepartmental endeavors. Thus, despite the lackluster physician enrollment or engagement reported by some Awardees (in Exiting and Active cohorts), non-enrolled physicians and non-physician clinical staff may aid the implementation of care redesigns and their effect on model goals. The extent to which this may occur is currently unknown. Future analyses will attempt to discern the extent by stratifying Model 1 impact effects by enrolled and non-enrolled physicians.

Impact estimates indicated that Medicare payment increases were *muted* (i.e., increased less than comparisons) for Awardees over the primary period of focus under this model, the inpatient stay (i.e., episode). Medicare payments to other providers after the episode period (e.g., physicians, nursing facilities, and rehabilitation hospitals) did increase relative to baseline and comparison hospitals. These Medicare payment findings provide interim insight on *potential* Model 1 effects. Payments over this post-episode period are monitored and compared to historical baselines by another CMS contractor, and information from CMS has indicated that Awardees did not exceed predetermined thresholds for payment increases over the 30-day post-episode period.

Resource utilization impact findings were mixed, as evidenced by increased likelihoods of ICU services and some decreases in episode length of stay across Awardees. Among Active cohort Awardees, increased mortality likelihoods for patients discharged from Awardee hospitals were noted. The elevated mortality impact estimate appeared to be attributable to outlier Awardee performance on this measure. Taken together, these impacts and increases in post-episode Medicare payments may be indicative of (unintended) adverse consequences of Model 1. However, such conclusions would be premature without additional analyses that incorporate patient health and experience data and other quality indicators. These data sources are expected in the 2015 Annual Report.

Aforementioned impact differentials between Active and Exiting cohort Awardees were not unexpected. As previously noted, exit interview data indicated that Awardee Agreement terminations were due, in part, to Awardee beliefs that their care redesigns achieved sufficient progress to continue without model incentivizing components or the Model 1 IPPS payment

discount. Six of the nine Exiting cohort Awardees had recent experience in implementing care redesigns and similar incentivizing mechanisms from another CMS model, the PHC. These six Awardees were among the earliest Awardees to terminate their agreement. Additional quarters of data will allow for inferences as to whether BPCI Model 1 had sustained impact on Exiting cohort Awardees. Conversely, Active Awardees still believed they had enough to gain from BPCI Model 1 in terms of improving clinician and hospital alignment towards care redesigns to remain active in BPCI Model 1. This sentiment may be indicative of a delay in the translation of Active cohort Awardee care redesign implementation effects to Model 1 goals.

Section V. Forthcoming Measures and Analyses

V.A. Domains and Measures

Table 50 presents measures for future analyses, expected in the 2015 Annual Report. These measures will provide a more comprehensive view of beneficiary outcomes and Awardee ability to achieve internal hospital cost savings.

Table 50: Domains and Measures in Future Reports

Domain	Measure
Care Processes	• Care coordination at discharge (PHES)
Medicare Payments and Internal Hospital Costs	• Internal hospital costs savings from enrolled physicians
Health Care Outcomes/ Patient Experience	• Functional status – mobility (PHES) • Functional status – pain intensity (PHES) • PAC assessment (MDS, HHA-OASIS, IRF-PAI)

V.B. Types of Analyses

In addition to measures noted above, forthcoming analyses aim to examine the following:

- *Enrolled physicians.* Differences across measures for physicians enrolled in BPCI Model 1 will be examined. Further, these differences will be compared to non-enrolled physician measure performance at Awardee hospitals. These comparisons will allow inference into whether impact results for an Awardee hospital reflect an "overpowering" of non-enrolled physician measure performance, or vice-versa.
- *Alternative groupings of care redesigns across Awardees.* The initial grouping of Awardees by care redesign types for this report might not have great specificity due to the variety of care redesigns pursued across Awardees. Further, Awardees may have changed their pursued care redesigns. Alternative models for care redesign classification will be assessed in 2015.
- *Testing for care shifting and other methods of "gaming" the system.* Physicians often have admissions privileges at multiple hospitals or can rely on transfers to keep lower-cost patients at BPCI hospitals. The availability of physician-level data will allow testing for various methods that could potentially bias results.