

Bundled Payments for Care Improvement Advanced Model

Final Evaluation Report Appendices



August 2026

Prepared for: Centers for Medicare & Medicaid Services

Submitted by: The Lewin Group, with our partners Abt
Global Inc., General Dynamics Information
Technology (GDIT), and Telligen

Prepared for:

Federal project officer

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

Submitted by:

The Lewin Group, with our partners Abt Global Inc., General Dynamics Information Technology (GDIT), and Telligen, including our clinical advisor Christine LaRocca, MD

The Lewin Group

3237 Airport Rd
La Crosse, WI 54603

Authors

Rachel Henke, Kelsey Bacon, Aylin Bradley, Micaela Brewington, Shannon Bicott, Amanda Christensen, Brady Durst, Kassandra Fate, Menbere Haile, Olivia Hayden, Nicholas Jones, Amanda Markovitz, Kathryn Martin, Isabella Martinez, Ian Morrall, Lorissa Pagán, Mary Beth Schaefer, Alexandra Stone, Sean Stuhlsatz

Acknowledgments

Christine LaRocca, MD¹ (clinical advisor) and supporting team members Ariana Ackerman, Wallace Ao, Projesh Ghosh, Grace Kim, Maxwell Mindock, Jonathan Porter, Steven Poskitt, Sabrina Ulsh, Matt Trombley,² Catherine Hersey,² Caroline Logan,² Sharmini Radakrishnan,² Becky Blystone,¹ Fatima Beltran,¹ Maria Joseph-King,³ Stephanie Shao,³ Srinivas Madasu,³ Debra Boston,³ Chris Alleman,³ Angela Sour³

¹ Telligen, ² Abt Global Inc., ³ General Dynamics Information Technology (GDIT)

This project was funded by the Centers for Medicare & Medicaid Services under Contract No. HHSM-500-2014-000331, Task Order 75FCMC18F0089. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. The Lewin Group assumes responsibility for the accuracy and completeness of the information contained in this report.

List of Appendices

Appendix A: Glossary of Terms and Acronyms	A-1
Appendix B: Clinical Episode Definitions	B-1
Appendix C: Methods	C-1
Appendix D: Patient Survey – Instrument	D-1
Appendix E: Supplemental Statistics on Comparison Groups	E-1
Appendix F: Participant Characteristics.....	F-1
Appendix G: Supplemental Sample Characteristics.....	G-1
Appendix H: Expenditure and Utilization Outcomes – Results ..	H-1
Appendix I: Medicare Program Savings – Results.....	I-1
Appendix J: Participant Reconciliation – Results	J-1
Appendix K: Claims-Based Quality Outcomes – Results.....	K-1
Appendix L: Patient Survey – Results	L-1

Appendix A: Glossary of Terms and Acronyms

The following tables provide definitions of the terms, abbreviations, and time periods that are used throughout the report and the accompanying appendices. For more information on these terms, refer to prior evaluation reports and appendices available on the BPCI Advanced Model web page: <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>. Exhibit A.4 provides a high-level overview of the topics discussed in all evaluation reports and appendices.

Exhibit A.1: Glossary of Terms

Name	Definition
90-Day Post-Discharge Period (PDP)	The 90 days following discharge from an inpatient stay or when an outpatient procedure is completed
Accountable Care Organization (ACO)	A group of doctors, hospitals, and other health care providers who work together to give coordinated, high-quality care to Medicare patients they serve.
Active Physician Group Practice (PGP)	A BPCI Advanced PGP participant with a participation agreement with CMS or included as a downstream episode initiator on a participation agreement between a convener and CMS that had at least one discharge attributed to the model between January 1, 2013, and December 31, 2025.
Administrative Quality Measures Set	A set of six exclusively claims-based quality measures collected directly by CMS. Two measures included in the set, All-Cause Hospital Readmission and Advanced Care Plan, applied to all clinical episodes, and up to two additional measures could apply to each clinical episode. For Model Year 4 (2021) and later, participants could choose to be evaluated based on either the Administrative Quality Measures Set or the Alternate Quality Measures Set.
Advanced Alternative Payment Model (Advanced APM)	Approaches to paying for health care that incentivize quality and value. Advanced APMs are a subset of APMs that meet these three criteria: - Requires participants to use certified electronic health record technology - Provides payment for covered professional services based on quality measures comparable to those used in the Merit-based Incentive Payment System Quality performance category - Either (1) is a Medical Home Model expanded under Center for Medicare and Medicaid Innovation (Innovation Center) authority or (2) requires participants to bear significant financial risk A benefit of a model's designation as an Advanced APM is that, under the Quality Payment Program, eligible clinicians who participate in an Advanced APM and meet certain Medicare patient/payment thresholds may qualify as a Qualifying APM Participant. They are then exempt from Merit-based Incentive Payment System reporting requirements and payment adjustments and receive a bonus incentive (prior to 2024) or a higher physician fee schedule update (2024 and beyond).
Alternate Quality Measures Set	A set of quality measures introduced in Model Year 4 (2021) that included a combination of up to five claims-based, hospital-based, and registry-based measures for each clinical episode. Two measures included in the set, All-Cause Hospital Readmission and Advanced Care Plan, applied to all clinical episodes, and up to three additional measures could apply to each clinical episode. Participants could choose to be evaluated based on either the Administrative Quality Measures Set or the Alternate Quality Measures Set.
Anchor Procedure	The hospital outpatient procedure that triggered the start of an outpatient episode.
Anchor Stay	The hospital inpatient stay that triggered the start of an inpatient episode.

Name	Definition
Benchmark Price	The price CMS calculated using claims-based historical data and risk adjustment models to account for variation in the Clinical Episode's standardized amounts. Starting in Model Year 4 (2021), the Benchmark Price during the Performance Period also accounted for realized national trends in the Performance Period that are driven by unanticipated, systematic factors.
Bundled Payments for Care Improvement Advanced (BPCI Advanced) Model	A voluntary value-based payment model from the CMS Innovation Center that tested whether linking payments for a clinical episode can reduce Medicare expenditures while maintaining or improving the quality of care.
BPCI Advanced Database	A database where CMS stored secure, frequently updated data about BPCI Advanced participants and episodes, from which the evaluation team could process various reports at any time.
Caregiver	A person that provided physical, medical, or household support or care to the patient during the episode.
Case-Mix Adjustment	A component that adjusted for varying levels of severity in a hospital's patient case mix that are outside the hospital's control. This value is only representative of the spending conditional on hospital patient case mix.
Clinical Episode	A type of episode of care (meaning an anchor stay or procedure and the 90-day post-discharge period), defined by certain Medicare Severity-Diagnosis Related Groups (MS-DRGs) or Healthcare Common Procedure Coding System (HCPCS) codes. See Appendix B for a complete list of MS-DRG and HCPCS codes corresponding to each clinical episode. For Model Years 1 and 2 (2018 and 2019), there were 32 clinical episodes, and for Model Years 3 through 8 (2020 through 2025), there were 34 clinical episodes. During Model Years 1 through 3, participants chose to participate in one or more clinical episodes. In Model Year 4 and later, participants were required to choose one or more clinical episode service line groups (see below).
Clinical Episode Service Line Group (CESLG)	A collection of clinical episodes sorted into clinically related groups. For Model Year 4 and later, participants were required to select CESLGs instead of one or more clinical episodes. Participants were not required to participate in clinical episodes within a CESLG that did not meet the minimum volume threshold during the baseline period.
Clinically Integrated Network	A legal entity comprising a group of health care providers working together to deliver efficient and coordinated care to patients and decrease the cost of care.
CMS Discount	A discount applied to the benchmark price to calculate the target price for each clinical episode for each hospital or PGP. The CMS discount was 3% for all episodes in Model Years 1 through 5. In Model Year 6 and later, the CMS discount was 2% for medical episodes and 3% for surgical episodes.
Composite Quality Score (CQS)	An aggregate quality score determined by calculating a score for each quality measure at the clinical episode level. Scores were scaled across the clinical episodes attributed to a specific hospital or PGP and weighted based on clinical episode volume. The CQS was used to adjust positive or negative total reconciliation amounts.
Consultant	In the context of BPCI Advanced, a nonparticipating organization that provided strategic guidance on participation and clinical episode selection, data analytic support, or other services to help episode initiators succeed in the model.
Convener Participant	A type of participant with at least one downstream episode initiator that facilitated coordination among downstream episode initiators and other providers and bore and apportioned financial risk. A convener participant may or may not have been a Medicare provider or initiated episodes. For Model Year 7 (2024), the following eligible entities could apply to participate in BPCI Advanced as new Convener Participants: Medicare-enrolled providers or suppliers, ACOs, or Active Convener Participants in Model Year 6 (2023) that are not Medicare-enrolled providers or suppliers, or ACOs.

Name	Definition
Difference-in-differences (DiD)	A statistical technique used in the BPCI Advanced evaluation to study the effect of an independent variable on an outcome, comparing the change over time in outcomes for the treatment group with the change for the control group. Referred to as “Impact Estimates” within the appendices.
Downstream Episode Initiator	A hospital or PGP that is associated with a convener participant and initiates episodes. Downstream episode initiators do not bear financial risk directly with CMS.
Episode Initiator	The Medicare-enrolled hospital or PGP participating in the BPCI Advanced model as a participant or a downstream episode initiator that could trigger episodes.
Episode of Care	The anchor stay or procedure plus the 90-day post-discharge or post-procedure period.
First Admission	The single episode retained for a patient when multiple eligible discharges or procedures occur during the 90-day post-anchor period, following the BPCI Advanced Model specifications for Model Year 4 and extending through Model Year 8.
Generative Artificial Intelligence (AI)	A form of AI that uses neural networks and deep learning algorithms to identify patterns within existing data as a basis for synthesizing knowledge to create original outputs.
Initial Model Design Period	Time period starting in Model Year 1 and extending through Model Year 3.
Key Informant Interview (KII)	A focused qualitative interview with an individual or individuals with firsthand knowledge about the topic area of interest. The evaluation team conducted KIIs to better understand a localized topic area as it related to the impact BPCI Advanced had on quality of care, utilization, or expenditures.
Major Changes Period	Time period starting in Model Year 4 and extending through Model Year 6.
Medical Episode	An episode with an MS-DRG or HCPCS code within one of the four medical clinical episode service line groups: <i>cardiac care</i> , <i>gastrointestinal care</i> , <i>medical and critical care</i> , and <i>neurological care</i> . See Appendix B for a complete list of MS-DRGs, HCPCS codes, and clinical episodes.
Model Extension Period	Time period starting in Model Year 7 and extending through Model Year 8.
Net Payment Reconciliation Amount (NPRA)	A positive total reconciliation amount paid to a participant by CMS, which included adjustments based on quality performance. NPRA was earned when the aggregate Medicare allowed amounts for a participant’s clinical episodes were lower than the target price for the clinical episodes and remained positive after adjustments for quality performance.
Non-Convener Participant	An individual hospital or PGP that assumed financial risk for clinical episodes. Non-convener participants were also episode initiators.
Participation Agreement	A formal agreement signed upon model entry between CMS and either the non-convener participant or convener participant that they agree to bear financial risk for clinical episodes under BPCI Advanced.
Patient Optimization	A practice in which providers identify and mitigate medical risk factors prior to surgical procedures to minimize complications and improve recovery.
Patient Survey	A survey that explores differences in patient care experiences and functional outcomes between Medicare patients cared for by BPCI Advanced providers and similar patients whose providers did not participate in BPCI Advanced. The patient survey collected information on a set of patient outcomes related to functional status, health care experience, and satisfaction with care and recovery.
Peer Group Trend Factor Adjustment	Adjusted for realized peer group trends driven by unanticipated, systematic factors by recentering the benchmark price around realized performance period clinical episode spending within each peer group nationally.

Name	Definition
Performance Period	A defined period during which episodes were initiated and all Medicare fee-for-service (FFS) payments aggregated for a specific clinical episode were attributed to a participant. The performance periods were used to determine reconciliation for episodes. Apart from the first performance period (October 1, 2018–December 31, 2018), performance periods ran from January 1 through June 30 and July 1 through December 31. The BPCI Advanced Model included 14 performance periods, running through December 31, 2025.
Post-Acute Care (PAC)	All care services received by a patient after discharge from a hospital stay or procedure. Includes care from the PAC provider (skilled nursing facility, inpatient rehabilitation facility, long-term care hospital, or home health agency).
Reconciliation	The semi-annual process where CMS determined whether participants were eligible for an NPRA payment or were required to pay back money to CMS. CMS compared the aggregate Medicare FFS allowed amounts for all items and services included in episodes attributed to a participant against the target price for those episodes and adjusted for performance on quality. If the aggregate amounts were lower than the target price (after accounting for the CQS and stop-gain/stop-loss adjustments), the participant was eligible to receive an NPRA payment from CMS. If the aggregate amounts were higher than the target price (after accounting for the CQS and stop-gain/stop-loss adjustments), they were required to pay a repayment amount to CMS.
Reconciliation Payment	The amount CMS had to pay participants when aggregate Medicare allowed amounts for clinical episodes that the participant had selected, including an adjustment for the CQS, were lower than the target price for such clinical episodes.
Repayment Amount	The amount participants had to pay CMS when aggregate Medicare allowed amounts for clinical episodes that the participant had selected, including an adjustment for the CQS, were higher than the target price for such clinical episodes.
Retrospective Trend Adjustment	A retrospective adjustment that was designed to help account for unanticipated systematic factors (such as payment system reforms) common to all hospitals (including nonparticipant hospitals) in the peer group during the performance period and was applied to the final target price. Adjustments were applied to each peer group and used observed episode spending.
Rural	A hospital was considered rural based on the CBSA Urban Rural Indicator field from the Provider of Services (POS) file. This variable indicates whether the county in which the hospital is located is defined as urban or rural. See Appendix C for more details on the POS data source.
Safety-Net Hospital	A facility was considered a safety-net facility based on a list from CMS, as described in Appendix C . Facilities were identified when their patient mix of beneficiaries with dual eligibility or Part D Low Income Subsidy exceeds the 75th percentile threshold for all congruent facilities that bill Medicare.
Surgical Episode	An episode with an MS-DRG or HCPCS code within one of the four surgical clinical episode service line groups: <i>cardiac procedures</i> , <i>gastrointestinal surgery</i> , <i>orthopedics</i> , and <i>spinal procedures</i> . See Appendix B for a complete list of MS-DRGs, HCPCS codes, and clinical episodes.
Target Price	The benchmark price with the CMS discount (in Model Years 6–8, the discount was 2% for medical episodes and 3% for surgical episodes; in Model Years 1–5, the discount was a flat 3% for all clinical episodes) applied for each combination of episode initiator and clinical episode. It was used to determine whether participants were eligible for NPRA or required to pay a repayment amount to CMS.

Note: See the first page of this appendix for data sources and more information.

Exhibit A.2: Acronym List

Acronym	Definition
ACO	Accountable Care Organization
ADI	Area Deprivation Index
AHRQ	Agency for Healthcare Research and Quality
AI	Artificial Intelligence
APM	Alternative Payment Model
BPCI	Bundled Payments for Care Improvement
BPCI Advanced	Bundled Payments for Care Improvement Advanced
BPCI-C	Bundled Payments for Care Improvement Initiative
BPID	Bundled Payments for Care Improvement Advanced Participant Identifier
CAHPS	Consumer Assessment of Healthcare Providers and Systems
CARE	Continuity Assessment Record and Evaluation
CCN	CMS Certification Number
CEN	Clinical Expert Network
CESLG	Clinical Episode Service Line Group
CI	Confidence Interval
CJR Model	Comprehensive Care for Joint Replacement Model
CMMI	Center for Medicare & Medicaid Innovation
CMS	Centers for Medicare & Medicaid Services
COVID-19	Coronavirus Disease 2019
CQS	Composite Quality Score
DiD	Difference-in-Differences
ED	Emergency Department
ESRD	End-Stage Renal Disease
FFS	Fee-for-Service
FY	Fiscal Year
GDIT	General Dynamics Information Technology
HCAHPS	Hospital Consumer Assessment of Healthcare Providers and Systems
HCC	Hierarchical Condition Category
HCPCS	Healthcare Common Procedure Coding System
HH	Home Health
ICD	International Classification of Diseases
IMC	Implementation and Monitoring Contractor
IPPS	Inpatient Prospective Payment System
IRF	Inpatient Rehabilitation Facility
ITT	Intention to Treat
LASSO	Least Absolute Shrinkage and Selection Operator
LCI	Lower Confidence Interval
LTCH	Long-Term Care Hospital
MBSF	Master Beneficiary Summary File
MJRLE	Major Joint Replacement of the Lower Extremity
MS-DRG	Medicare Severity-Diagnosis Related Group

Acronym	Definition
MY	Model Year MY1 = October 1, 2018–December 31, 2018 MY2 = January 1, 2019–December 31, 2019 MY3 = January 1, 2020–December 31, 2020 MY4 = January 1, 2021–December 31, 2021 MY5 = January 1, 2022–December 31, 2022 MY6 = January 1, 2023–December 31, 2023 MY7 = January 1, 2024–December 31, 2024 MY8 = January 1, 2025–December 31, 2025
N/A	Not Applicable
NPI	National Provider Identifier
NPRA	Net Payment Reconciliation Amount
PAC	Post-Acute Care
PDP	Post-Discharge Period
PGP	Physician Group Practice
POS	Provider of Services
PP	Percentage Point(s)
QP	Qualifying APM Participant
QR	Quick Response
RIF	Research Identifiable Files
SMD	Standardized Mean Difference
SNF	Skilled Nursing Facility
TEAM	Transforming Episode Accountability Model
TIN	Taxpayer Identification Number
UCI	Upper Confidence Interval

Note: See the first page of this appendix for data sources and more information.

Exhibit A.3: Impact Analysis Baseline and Intervention Periods

Model Period	Model Years	Baseline Period	Intervention Period
Initial Model Design	MY1–2	April 1, 2013–December 31, 2017	October 1, 2018–December 31, 2019
	MY3	April 1, 2013–December 31, 2017	January 1, 2020–December 31, 2020
Updated Design	MY4–6	January 1, 2015–September 30, 2018	January 1, 2021–December 31, 2023
Model Extension	MY7–8	January 1, 2016–September 30, 2018	January 1, 2024–April 30, 2025

Note: See the first page of this appendix for data sources and more information.

Exhibit A.4: Annual Evaluation Report Topics

Topic	Description	Evaluation Report						
		First	Second	Third	Fourth	Fifth	Sixth	Seventh
Model Years Covered	Indicates which model years are included for findings in each report.	MY1 (2018)	MY1–2 (2018– 2019)	MY1–3 (2018– 2020)	MY3 (2020)	MY4–5 (2021– 2022)	MY4–6 (2021– 2023)	MY4–8 (2021– 2025)
ACO Overlap	Examination of how BPCI Advanced intersected with ACOs through descriptive analyses, comparative impact estimates for ACO- versus non-ACO-aligned patients, and key informant interviews capturing participant perspectives.					X	X	
Care Transformation and Care Redesign	Summarizes the operational or clinical changes participants made in response to the model. Examples include care coordination, discharge planning, and use of CMS data.		X			X	X	X
Connections to Primary Care	Examines whether the model influenced use of primary care services, through impact analyses of primary care use outcomes and key informant interviews capturing participant perspectives.						X	
COVID-19 Analyses	Captures analyses and discussion related to the COVID-19 public health emergency, including methodological changes to the model in response to COVID-19, and the implications for model impacts, performance, and interpretation of findings.			X	X	X		
Clinical Episode Selection	Addresses which clinical episodes participants selected and how those choices shaped model participation.	X	X	X	X	X		X
Historically Underserved Populations Analyses	Examines how outcomes or impacts differ for populations that have been historically underserved. Includes stratified analyses by patient subgroups and key informant interviews capturing participant perspectives.				X	X	X	
Model Reach	Summarizes the scale of the model, such as hospital participation rates, number of clinicians or organizations involved, and the share of eligible episodes or beneficiaries affected.	X	X			X	X	X

Topic	Description	Evaluation Report						
		First	Second	Third	Fourth	Fifth	Sixth	Seventh
Outcomes: Expenditures, Utilization, and Medicare Program Savings	Impact of the model on episode payments, utilization, and savings for Medicare. It captures both episode-level payment changes and program-level financial effects.		X	X	X	X	X	X
Outcomes: Patient Reported	Focuses on outcomes reported directly by patients via a survey, including functional status, care experience, and satisfaction.		X		X	X	X	X
Outcomes: Quality	Assesses whether quality of care changed under the model. Includes mortality and readmission rates.		X	X	X	X	X	X
Participant Characteristics	Describes the organizations participating in the model: hospitals, physician group practices, and conveners. Includes provider attributes and how participating entities differed over time.	X	X	X	X	X	X	X
Participation Decisions	Examines why organizations joined, remained in, or exited the model and how actively they stayed engaged. Reflects participation patterns and retention over time.	X	X			X	X	X
Patient Mix	Examines differences in the types of patients included in episodes, such as clinical complexity, demographic composition, or other case-mix characteristics.			X	X	X	X	X

Note: See the first page of this appendix for data sources and more information.

Appendix B: Clinical Episode Definitions

The following tables display the BPCI Advanced clinical episode by clinical episode service line group (CESLG) (Exhibit B.1), as well as the Healthcare Common Procedure Coding System (HCPCS) and Medicare Severity-Diagnosis Related Group (MS-DRG) codes that trigger each clinical episode (Exhibits B.2–B.6).

Note that MS-DRG, HCPCS, and procedure codes are subject to change across model years. CMS revises MS-DRG definitions from year to year. We reported Model Years 4 and 5 BPCI Advanced inpatient and multi-setting clinical episodes and MS-DRGs in the BPCI Advanced Fifth and Sixth Evaluation Reports, respectively.¹ In Exhibit B.2, we report the Model Year 6 BPCI Advanced inpatient and multi-setting clinical episodes and MS-DRGs, and in Exhibit B.3 we report the Model Years 7 and 8 BPCI Advanced inpatient and multi-setting clinical episodes and MS-DRGs. For a full list of Model Years 6 through 8 trigger HCPCS codes, see Exhibit B.4.

Certain clinical episodes were defined by unique triggering rules. Endovascular cardiac valve replacement episodes were triggered by claims that had both one of the corresponding MS-DRG codes (see Exhibits B.2–B.3) and at least one procedure code from Exhibit B.5. Two clinical episodes were multi-setting, where the clinical episode included both inpatient and outpatient episodes (major joint replacement of the lower extremity and, beginning in Model Year 6, major joint replacement of the upper extremity). For both of these multi-setting clinical episodes, inpatient episodes required the appropriate MS-DRG (Exhibit B.2–B.3). For major joint replacement of the lower extremity, outpatient episodes required the corresponding trigger HCPCS code (Exhibit B.4), while for major joint replacement of the upper extremity, outpatient episodes required a corresponding triggering procedure code (see Exhibit B.6).

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on methods. For details regarding previous model year clinical episode definitions, see the BPCI Advanced Evaluation First Evaluation Report through the Sixth Evaluation Report. All information, including clinical episodes, and MS-DRG and HCPCS codes, can also be found on the *Technical Resources* section of the *Participant Resources* page of the CMS BPCI Advanced Model website:

<https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced/participant-resources>

¹ All the BPCI Advanced Evaluation Reports and their appendices are available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Exhibit B.1: BPCI Advanced Evaluation CESLG Definitions

Clinical Episode Type	CESLG	Clinical Episode
Medical	Cardiac Care	- Acute Myocardial Infarction - Cardiac Arrhythmia - Congestive Heart Failure
	Gastrointestinal Care	- Disorders of Liver Except Malignancy, Cirrhosis, or Alcoholic Hepatitis - Gastrointestinal Hemorrhage - Gastrointestinal Obstruction - Inflammatory Bowel Disease
	Medical & Critical Care	- Cellulitis - Chronic Obstructive Pulmonary Disease, Bronchitis, Asthma - Renal Failure - Sepsis - Simple Pneumonia and Respiratory Infections - Urinary Tract Infection
	Neurological Care	- Seizures - Stroke
Surgical	Cardiac Procedures	- Cardiac Defibrillator (Inpatient) - Cardiac Defibrillator (Outpatient) - Cardiac Valve - Coronary Artery Bypass Graft - Endovascular Cardiac Valve Replacement - Pacemaker - Percutaneous Coronary Intervention (Inpatient) - Percutaneous Coronary Intervention (Outpatient)
	Gastrointestinal Surgery	- Bariatric Surgery - Major Bowel Procedure
	Orthopedics	- Double Joint Replacement of the Lower Extremity - Fractures of the Femur and Hip or Pelvis^a - Hip and Femur Procedures Except Major Joint - Lower Extremity and Humerus Procedures Except Hip, Foot, Femur - Major Joint Replacement of the Lower Extremity - Major Joint Replacement of the Upper Extremity
	Spinal Procedures	- Back and Neck Except Spinal Fusion (Inpatient) - Back and Neck Except Spinal Fusion (Outpatient) - Spinal Fusion

Note: See the first page of this appendix for data sources and more information.

^a The BPCI Advanced Evaluation's classification of this clinical episode as surgical differs from the [BPCI Advanced Model Pricing Methodology and Clinical Episode Construction Specifications](#), which classify fractures of the femur and hip or pelvis as a medical clinical episode. For evaluation purposes, each CESLG and all corresponding episodes have been aligned as either medical or surgical.

Exhibit B.2: BPCI Advanced Inpatient and Multi-setting Clinical Episodes and MS-DRGs, Model Year 6 (Jan. 2023–Dec. 2023)

Clinical Episode	MS-DRG Trigger Codes
Acute Myocardial Infarction	280–282
Back and Neck Except Spinal Fusion (Inpatient)	518–520
Bariatric Surgery	619–621
Cardiac Arrhythmia	308–310
Cardiac Defibrillator (Inpatient)	222–227, 275–277
Cardiac Valve	212, 216–221
Cellulitis	602–603
Chronic Obstructive Pulmonary Disease, Bronchitis, Asthma	190–192, 202–203
Congestive Heart Failure	291–293
Coronary Artery Bypass Graft	231–236
Disorders of Liver Except Malignancy, Cirrhosis, or Alcoholic	441–443
Double Joint Replacement of the Lower Extremity	461–462
Endovascular Cardiac Valve Replacement	266–267
Fractures of the Femur and Hip or Pelvis	533–536
Gastrointestinal Hemorrhage	377–379
Gastrointestinal Obstruction	388–390
Hip and Femur Procedures Except Major Joint	480–482
Inflammatory Bowel Disease	385–387
Lower Extremity and Humerus Procedure Except Hip, Foot, Femur	492–494
Major Bowel Procedure	329–331
Major Joint Replacement of the Lower Extremity	469–470, 521–522
Major Joint Replacement of the Upper Extremity	483
Pacemaker	242–244
Percutaneous Coronary Intervention (Inpatient)	246–251, 321–322
Renal Failure	682–684
Seizures	100–101
Sepsis	870–872
Simple Pneumonia and Respiratory Infections	177–179, 193–195
Spinal Fusion	453–455, 459–460, 471–473
Stroke	061–066
Urinary Tract Infection	689–690

Note: See the first page of this appendix for data sources and more information.

Exhibit B.3: BPCI Advanced Inpatient and Multi-setting Clinical Episodes and MS-DRGs, Model Years 7–8 (Jan. 2024–Oct. 2025)

Clinical Episode	MS-DRG Trigger Codes
Acute Myocardial Infarction	280–282
Back and Neck Except Spinal Fusion (Inpatient)	518–520
Bariatric Surgery	619–621
Cardiac Arrhythmia	308–310
Cardiac Defibrillator (Inpatient)	275–277
Cardiac Valve	212, 216–221
Cellulitis	602–603
Chronic Obstructive Pulmonary Disease, Bronchitis, Asthma	190–192, 202–203
Congestive Heart Failure	291–293
Coronary Artery Bypass Graft	231–236
Disorders of Liver Except Malignancy, Cirrhosis, or Alcoholic	441–443
Double Joint Replacement of the Lower Extremity	461–462
Endovascular Cardiac Valve Replacement	266–267
Fractures of the Femur and Hip or Pelvis	533–536
Gastrointestinal Hemorrhage	377–379
Gastrointestinal Obstruction	388–390
Hip and Femur Procedures Except Major Joint	480–482
Inflammatory Bowel Disease	385–387
Lower Extremity and Humerus Procedure Except Hip, Foot, Femur	492–494
Major Bowel Procedure	329–331
Major Joint Replacement of the Lower Extremity	469–470, 521–522
Major Joint Replacement of the Upper Extremity	483
Pacemaker	242–244
Percutaneous Coronary Intervention (Inpatient)	250–251, 318 ^a , 321–322, 359 ^a –360 ^a
Renal Failure	682–684
Seizures	100–101
Sepsis	870–872
Simple Pneumonia and Respiratory Infections	177–179, 193–195
Spinal Fusion	402, 426–430, 447–448, 450–451, 453 ^b –455 ^b , 459 ^b –460 ^b , 471–473
Stroke	061–066

Clinical Episode	MS-DRG Trigger Codes
Urinary Tract Infection	689–690

Note: See the first page of this appendix for data sources and more information.

^a MS-DRG trigger code only present for Model Year 8 (2025).

^b MS-DRG trigger code only present for Model Year 7 (2024).

Exhibit B.4: BPCI Advanced Outpatient and Multi-setting Clinical Episodes and HCPCS Codes, Model Years 6–8 (Jan. 2023–Oct. 2025)

Clinical Episode	HCPCS Trigger Code
Back and Neck Except Spinal Fusion	62287, 63005, 63011–63012, 63017, 63030, 63040, 63042, 63045–63047, 63056, 63075
Cardiac Defibrillator (Outpatient)	33249, 33262–33264, 33270
Percutaneous Coronary Intervention (Outpatient)	92920, 92924, 92928, 92933, 92937, 92943, C9600, C9602, C9604, C9607
Major Joint Replacement of the Lower Extremity	27447, 27130
Major Joint Replacement of the Upper Extremity	23472

Note: See the first page of this appendix for data sources and more information.

Exhibit B.5: Procedure Codes for Endovascular Cardiac Valve Replacement, Model Years 6–8 (Jan. 2023–Oct. 2025)

Procedure Code	ICD-9/ICD-10
3505	ICD-9
3506	ICD-9
02RF37H	ICD-10
02RF37Z	ICD-10
02RF38H	ICD-10
02RF38N	ICD-10
02RF38Z	ICD-10
02RF3JH	ICD-10
02RF3JZ	ICD-10
02RF3KH	ICD-10
02RF3KZ	ICD-10
X2RF332	ICD-10

Note: See the first page of this appendix for data sources and more information.

Exhibit B.6: Procedure Codes for Major Joint Replacement of the Upper Extremity, Model Years 6–8 (Jan. 2023–Oct. 2025)

Procedure Code	Description
0RRE07Z	Replacement of Right Sternoclavicular Joint with Autologous Tissue Substitute, Open Approach
0RRE0JZ	Replacement of Right Sternoclavicular Joint with Synthetic Substitute, Open Approach
0RRE0KZ	Replacement of Right Sternoclavicular Joint with Nonautologous Tissue Substitute, Open Approach
0RRF07Z	Replacement of Left Sternoclavicular Joint with Autologous Tissue Substitute, Open Approach

Procedure Code	Description
0RRF0JZ	Replacement of Left Sternoclavicular Joint with Synthetic Substitute, Open Approach
0RRF0KZ	Replacement of Left Sternoclavicular Joint with Nonautologous Tissue Substitute, Open Approach
0RRG07Z	Replacement of Right Acromioclavicular Joint with Autologous Tissue Substitute, Open Approach
0RRG0JZ	Replacement of Right Acromioclavicular Joint with Synthetic Substitute, Open Approach
0RRG0KZ	Replacement of Right Acromioclavicular Joint with Nonautologous Tissue Substitute, Open Approach
0RRH07Z	Replacement of Left Acromioclavicular Joint with Autologous Tissue Substitute, Open Approach
0RRH0JZ	Replacement of Left Acromioclavicular Joint with Synthetic Substitute, Open Approach
0RRH0KZ	Replacement of Left Acromioclavicular Joint with Nonautologous Tissue Substitute, Open Approach
0RRJ00Z	Replacement of Right Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Open Approach
0RRJ07Z	Replacement of Right Shoulder Joint with Autologous Tissue Substitute, Open Approach
0RRJ0J6	Replacement of Right Shoulder Joint with Synthetic Substitute, Humeral Surface, Open Approach
0RRJ0J7	Replacement of Right Shoulder Joint with Synthetic Substitute, Glenoid Surface, Open Approach
0RRJ0JZ	Replacement of Right Shoulder Joint with Synthetic Substitute, Open Approach
0RRJ0KZ	Replacement of Right Shoulder Joint with Nonautologous Tissue Substitute, Open Approach
0RRK00Z	Replacement of Left Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Open Approach
0RRK07Z	Replacement of Left Shoulder Joint with Autologous Tissue Substitute, Open Approach
0RRK0J6	Replacement of Left Shoulder Joint with Synthetic Substitute, Humeral Surface, Open Approach
0RRK0J7	Replacement of Left Shoulder Joint with Synthetic Substitute, Glenoid Surface, Open Approach
0RRK0JZ	Replacement of Left Shoulder Joint with Synthetic Substitute, Open Approach
0RRK0KZ	Replacement of Left Shoulder Joint with Nonautologous Tissue Substitute, Open Approach
0XM00ZZ	Reattachment of Right Forequarter, Open Approach
0XM10ZZ	Reattachment of Left Forequarter, Open Approach
0XM20ZZ	Reattachment of Right Shoulder Region, Open Approach
0XM30ZZ	Reattachment of Left Shoulder Region, Open Approach
0XM40ZZ	Reattachment of Right Axilla, Open Approach
0XM50ZZ	Reattachment of Left Axilla, Open Approach
0XM60ZZ	Reattachment of Right Upper Extremity, Open Approach
0XM70ZZ	Reattachment of Left Upper Extremity, Open Approach
0XM80ZZ	Reattachment of Right Upper Arm, Open Approach
0XM90ZZ	Reattachment of Left Upper Arm, Open Approach
0RRJ008 ^a	Replacement of Right Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Subscapularis-sparing Technique, Open Approach
0RRJ0J8 ^a	Replacement of Right Shoulder Joint with Synthetic Substitute, Subscapularis-sparing Technique, Open Approach
0RRK008 ^a	Replacement of Left Shoulder Joint with Reverse Ball and Socket Synthetic Substitute, Subscapularis-sparing Technique, Open Approach
0RRK0J8 ^a	Replacement of Left Shoulder Joint with Synthetic Substitute, Subscapularis-sparing Technique, Open Approach

Note: See the first page of this appendix for data sources and more information.

^a MS-DRG trigger code only present for Model Year 8 (2025).

Appendix C: Methods

This appendix includes details on data sources and methods used for analyses in the Final Evaluation Report. It has four main sections: Primary and Secondary Data Sources, Qualitative Methods, Patient Survey Methods, and Quantitative Methods for Secondary Data.

To evaluate the BPCI Advanced Model, we used a mixed-methods approach, meaning that we combined quantitative and qualitative methods to analyze both primary and secondary data. Qualitative methods provide deep insights into providers' response to the model and the on-the-ground realities of model implementation. Quantitative methods can precisely quantify the magnitude of effects and provide confidence intervals indicating the uncertainty of conclusions. Combining quantitative and qualitative methods allowed us to substantiate conclusions with multiple sources of evidence, contextualize findings more broadly, and investigate a wider range of facets of the model.

In addition to presenting new analyses, the Final Evaluation Report discusses results from past reports and builds on methodologies presented in those reports. Please see the corresponding methods appendices from prior reports for details on those methodologies, available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix.

A. Primary and Secondary Data Sources

This section describes the primary and secondary data sources that we used for qualitative and quantitative analyses. Exhibit C.1 lists the data sources and their uses in the evaluation report. We used primary data sources to describe BPCI Advanced patient and participant experiences in the model. We used provider-level data sources to identify and describe BPCI Advanced participating providers and the comparison group; we modified the list of comparison providers selected in Model Year 4 (2021) to create the comparison group for pooled Model Year 4 through 6 analyses.¹ We also used provider-level data sources to run various descriptive analyses and supplement our claims data for impact evaluation.

To assess participants' financial performance in the model and estimate Medicare program savings, we used reconciliation data from the BPCI Advanced Implementation and Monitoring Contractor (IMC).

Similar to past reports, to study the impacts of the BPCI Advanced Model in Model Years 4 through 6, we used Medicare claims and enrollment data to construct episodes of care for BPCI Advanced and matched comparison providers. We also used Medicare claims and enrollment

¹ For details on the comparison groups constructed for previous evaluation reports (including those used to estimate impacts for Model Years 1 and 2 and Model Year 3), see the corresponding methods appendices. Past reports and their appendices are available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

data to create outcome measures and patient risk factors associated with the outcomes. To study the model's impacts in Model Years 7 and 8, we integrated multiple data sets supplied by the IMC. For the period from January 2016 to September 2022, which includes our baseline period (January 2016–September 2018), we used episode-level target price files that the IMC used to calculate preliminary target prices for participants for Model Years 4 and 7. For the period from October 2022 to April 2025, which includes our intervention period (January 2024–April 2025), we used episode-level monthly files, which included a national set of episodes and formed the basis of the files that the IMC provided to participants to support their participation in the model. Each monthly file covered 18 months of data, and we used multiple monthly files to assemble our analytic research dataset. To ensure data accuracy, we applied a set of predefined selection criteria to episodes within each input file, including restrictions intended to minimize episode overlap and align observation windows across datasets. We selected files to maximize the claims runout time for standardized payment calculations and applied a minimum 3-month claims runout window. When feasible, we supplemented these data with our existing claims data to improve their internal consistency and suitability for impact evaluation purposes. Limitations of this approach are noted in the relevant sections.

Exhibit C.1: Data Sources Used in the BPCI Advanced Final Evaluation Report

Type	Dataset Name	Date Range	Dataset Contents	Use
Primary Data Sources	Clinical Expert Network (CEN)	2021–2022	Notes from meetings and emails with clinical experts who provide insights into the selected 2021–2022 patient survey results, including the clinical significance of findings, potential mechanisms underlying less favorable functional status and care experience outcomes, possible unintended consequences of the BPCI Advanced model, differences between hospital- and PGP-initiated episodes, and considerations related to respondents with dual eligibility.	Used to further interpret patient survey findings, assess whether less favorable results may be clinically meaningful, identify additional data needs and considerations for future analysis, and inform evaluation conclusions related to care patterns, subgroup differences, and potential unintended consequences.
	Key Informant Interviews (KIIs)	2019–2024	Transcripts from semi-structured telephone interviews conducted with a sample of physician group practices (PGPs), hospitals, and conveners on specific topics of interest.	Used to understand participation decisions, clinical episode service line group (CESLG) selection, and other participant experiences with the model. KIIs during Model Years 6 (2023) and 7 (2024) focused on how Medicare Accountable Care Organization (ACO) participation affected participant experience in BPCI Advanced, participant strategies to coordinate with primary care providers, perceptions of the target pricing methodology, financial factors influencing continued participation in BPCI Advanced, and potential unintended consequences related to BPCI Advanced target pricing.
	Patient Survey	Surveys collected July 2024–May 2025 (Wave 5)	Surveys completed by Medicare beneficiaries or their proxies. Patients received surveys about 90 days after leaving the hospital.	Used to create patient-reported outcome measures, such as patient-reported change in functional status, care experiences, and satisfaction with recovery and care received.
	Site Visits	2019–2024	Interview transcripts from virtual and in-person site visits that cover a wide range of subjects related to a hospital, PGP, or convener's experience in BPCI Advanced.	Used to understand participation decisions, CESLG selection, and other participant experiences with the model, as well as patient, participant, or market factors that may influence variation in model effects. New topics covered in Model Year 7 included decisions to rejoin BPCI Advanced, anticipated participation in future bundled payment models, interpretation of patient survey results, and the sustainability of care redesign activities after BPCI Advanced has ended.

Type	Dataset Name	Date Range	Dataset Contents	Use
Provider-Level Secondary Data Sources	Agency for Healthcare Research and Quality (AHRQ) Compendium of U.S. Health Systems, Hospital Linkage File	2016–2018; 2020–2023	Data linking hospitals to health systems.	Used to identify whether a hospital is part of a health system.
	AHRQ Compendium of U.S. Health Systems, Group Practice Linkage File	2016–2018; 2020–2023	Data linking PGPs to health systems.	Used to identify whether a PGP is part of a health system.
	Area Health Resource File	2015–2019; 2024–2025	County-level data on population, environment, geography, health care facilities, and health care professionals.	Used to create baseline hospital and PGP characteristics for balance assessment; construction of hospital and PGP comparison groups for Model Years 4–6; and construction of balancing weights used for Model Year 7–8 analyses that used the national comparison group.
	CMS Master Data Management Data Beneficiary Extract	2015–2025	Information on ACO attribution for patients with BPCI Advanced–eligible discharges and procedures during the baseline and intervention periods.	Used for balance assessment and impact estimates.
	CMS BPCI Advanced Participants Database	2018–2025	Information compiled by CMS on BPCI Advanced participants, including participant name, CMS Certification Number (CCN), Taxpayer Identification Number (TIN), location, type (hospital, PGP, other), BPCI Advanced “role,” clinical episode(s) selected, BPCI Advanced participation start and end dates, and contact information.	Used to identify participants, hospital episode initiators, and PGP episode initiators participating in BPCI Advanced and the clinical episodes in which they participated. Also used to identify entities that retroactively withdrew from the model or applied but did not become a participant or an episode initiator.
	CMS BPCI Advanced Hospital Target Pricing Files	2015–2024	The clinical episode–specific Model Year 1–7 preliminary target prices, historical payments, and historical volume for all BPCI Advanced–eligible hospitals.	Used for the hospital efficiency measure, a component of the Model Year 4 preliminary target price, which was used as a matching variable in hospital comparison group construction. Also used to describe the reconciliation amounts of providers with different target prices and to determine provider eligibility for purposes of defining our national comparison group.

Type	Dataset Name	Date Range	Dataset Contents	Use
Provider-Level Secondary Data Sources (Continued)	CMS BPCI Advanced Reconciliation Result Files	2018–2025	Reconciliation amounts and final target prices from the Implementation and Monitoring Contractor (IMC). Contains results for Model Years 1–8 (Performance Periods 1–13).	Used to calculate net Medicare savings due to BPCI Advanced as well as descriptive statistics concerning participant reconciliation.
	CMS BPCI Advanced Reconciliation Result Files (Continued)		Second true-up amounts were used whenever possible; however, for the last two performance periods (12, 13), we used first true-up and initial reconciliation because the second true-up data were not available at the time of data analysis.	
	CMS BPCI Database	2013–2018	Information compiled by CMS on BPCI awardees, including awardee name, CCN, TIN, location, type, clinical episode(s), BPCI Initiative participation start and end dates, and contact information.	Used to identify hospitals and PGPs that participated in the BPCI Initiative.
	CMS Comprehensive Care for Joint Replacement (CJR) Model Database	2016–2024	List of hospitals that ever participated in the CJR Model and their dates of participation.	Used to identify hospitals that participated in the CJR Model and to identify PGP episodes at CJR hospitals.
	CMS Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) Data	2018–2024	HCAHPS survey results for Inpatient Prospective Payment System (IPPS) hospitals, covering a variety of patient-reported quality measures.	Used for weighting and propensity-score matching for the patient survey.
	CMS Hospital Cost Reports	2018–2023	Financial data on hospitals, including revenues, expenditures, assets, and liabilities.	Used to describe BPCI Advanced participating hospitals, assess balance of BPCI Advanced and comparison episodes, and examine how model reconciliation payments relate to hospital finances.
	CMS IPPS Files	2011–2024	Hospital-level file containing provider characteristics such as Medicare days as a percentage of total days, resident-bed ratio, and Disproportionate Share Hospital patient percentage.	Used for hospital characteristics, risk adjustment, and hospital comparison group construction.
	CMS Safety-Net Hospital List	2017–2025	List of facilities identified as safety-net facilities from CMS. Facilities are identified when their patient mix of beneficiaries with dual eligibility or Part D Low-Income Subsidy exceeds the 75th percentile threshold for all congruent facilities that bill Medicare.	Used to analyze reconciliation performance of safety-net hospitals.

Type	Dataset Name	Date Range	Dataset Contents	Use
Provider-Level Secondary Data Sources (Continued)	CMS Provider of Service File	2013–2025	Information on Medicare-approved institutional providers, including CCN, size, ownership, location, and staffing.	Used in hospital comparison group construction, for risk adjustment, and for provider locations.
	CMS Research Identifiable File	2013–2025	Information on Shared Savings Program and Next Generation ACO provider-level participation data.	Used to compile Medicare ACO participation data among BPCI Advanced episode initiators.
	Medicare Data on Provider Practice and Specialty	2017–2020	Mapping physicians billing to practices.	Used to identify physician practice affiliation during the baseline period.
Transaction-Level Secondary Data Sources	IMC Episode-Level Files	Jan. 2016–Apr. 2025	Episode-level data files generated by the IMC as part of model administration. Included target pricing files (covering the episodes used to compute model target prices in Model Years 4 and 7) and monthly files (covering episodes of care during the model). Contains episode-level variables derived from Medicare Part A and B claims and standardized payments.	Used to estimate impacts of the model in Model Years 7–8, including identification of eligible episodes and construction of payment outcomes, claims-based patient outcomes, and risk-adjustment variables. Also used for descriptive analyses and to generate episode counts for Medicare program savings estimation.
	The Master Beneficiary Summary File (MBSF)	2013–2024	Beneficiary and enrollment information, including beneficiary unique identifier, address, date of birth and death, sex, age, and Medicare enrollment status.	Used to identify eligibility for episodes of care, beneficiary demographic characteristics, and beneficiary eligibility for inclusion in the denominator for each of the outcome measures.
	Medicare Fee-for-Service (FFS) Claims	Jan. 2013–Mar. 2024	Medicare Part A and B claims.	Used to create all claims-based patient outcomes, claims-based comparison group matching measures, and various risk-adjusters (including prior health care use). Also used to identify clinicians, clinician specialties, and hospitals associated with BPCI Advanced and BPCI Advanced PGP episode initiators; to determine whether a PGP is considered active or inactive; and to compute various descriptive statistics.
	Medicare Standardized Payments	Jan. 2013–Mar. 2024	Medicare standardized payments for 100% Part A and B claims.	Used to create Medicare standardized payment amounts (Part A and B) and allowed standardized payment outcomes.

Note: See **Appendix A** for definitions and acronyms.

B. Qualitative Methods

This section describes the qualitative methods we used to produce the results included in this report. We discuss the sampling approach and interview protocols for site visits, the sampling approach for key informant interviews (KIIs), qualitative analysis methods used across both site visits and KIIs, and study limitations.

1. Clinical Expert Network

The clinical expert network (CEN) provides clinician insights into and interpretation of quantitative data and corroboration of qualitative findings for this evaluation. Under the leadership of Dr. Christine LaRocca, geriatric medicine physician and medical director at Telligen, five clinical experts were vetted based on their area of expertise and under the condition that their organization not be a model participant; four candidates then participated in the CEN based upon their availability. CMS reviewed and approved this roster.

To provide the CEN members with a strong understanding of the BPCI Advanced model and their role in the evaluation, Dr. LaRocca shared a pre-work slide deck with panelists on Tuesday, November 28, 2023 so all CEN members would have a common understanding of the model and the data they would be reviewing and commenting upon.

On Friday, December 8, 2023, Telligen convened clinical experts for a panel to discuss selected results and unfavorable outcomes from the 2021 and 2022 pooled Beneficiary Survey results. The CEN panel was administered via webinar and Dr. LaRocca facilitated. CMS approved the materials prior to distribution to the CEN. The objectives of the CEN were to:

- Comment upon changes in care patterns and quality outcomes, as identified through data analysis;
- Report on changes in clinical practice that should be considered for the evaluation;
- Present medical or provider community feedback on this model;
- Provide insight into utilization and quality patterns we might expect given the incentives of the model;
- Identify practice changes that may differentially impact subpopulations of Medicare Fee-for-Service (FFS) beneficiaries; and
- Aid in the identification of promising practices and unintended consequences.

2. Site Visits

The BPCI Advanced evaluation team conducted site visits from Model Years 2 through 7. For information on site visits conducted in Model Year 2, see **Appendix C** of the BPCI Advanced First Evaluation Report. For information on site visits conducted in Model Year 3, see **Appendix C** of

the BPCI Advanced Second Evaluation Report. For information on site visits conducted in Model Years 4, 5, and 6, see **Appendix C** of the BPCI Advanced Sixth Evaluation Report.²

In Model Year 7, we interviewed BPCI Advanced participants and carried out six virtual and two in-person site visits with five hospitals and three physician group practices (PGPs). Site visits consisted of approximately six 1-hour interviews with representatives from each hospital or PGP, including executive and financial leaders, care redesign leaders, care coordination staff, front-line nursing and physician staff who cared for patients in BPCI Advanced, and convener representatives (when applicable). During site visits, we asked about BPCI Advanced participation decisions, decisions to rejoin the model, clinical episode service line group (CESLG) selection, care transformation and care redesign activities, how care transformation and care redesign activities may affect non-BPCI Advanced patients, the sustainability of care redesign activities after the model has ended, anticipated participation in future bundled payment models, external provider partnerships, and general successes and challenges in the model. Site visits also provided an opportunity to learn about patient, participant, or market factors that may influence variations in the model's effects.

a. Sampling Approach

The site visit sample was purposive to ensure a diverse sample. We included a mix of hospitals and PGPs that:

- Participated with or without a convener
- Had positive or negative financial performance in BPCI Advanced
- Participated in a variety of CESLGs
- Were located in a variety of geographic markets, including rural areas
- Joined BPCI Advanced at different times
- Did or did not have experience in previous Medicare initiatives

The characteristics of the final interviewee sample for Model Year 7 are shown in Exhibit C.2. When discussing the participant start year, the 2024 *rejoiners* are defined as those that participated in BPCI Advanced previously and rejoined under a new convener or rejoined as a non-convener participant in 2024. The *rejoiners* are not included in the 2018 or 2020 counts.

² The BPCI Advanced First, Second, and Sixth Evaluation Reports and their appendices are available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Exhibit C.2: Model Year 7 Site Visit Participant Characteristics

Site Visit Participant Characteristics		Count (N)
Episode Initiator Type	<i>Hospital</i>	5
	<i>PGP</i>	3
Participant Role	<i>Downstream Episode Initiator</i>	4
	<i>Non-convenor Participant</i>	4
Participant Start Year	<i>2018</i>	2
	<i>2020</i>	1
	<i>2024 (Rejoiner)</i>	3
	<i>2024 (New)</i>	2
Medical CESLG Type	<i>Cardiac Care</i>	2
	<i>Gastrointestinal Care</i>	1
	<i>Medicare & Critical Care</i>	4
	<i>Neurological Care</i>	1
Surgical CESLG Type	<i>Cardiac Procedures</i>	0
	<i>Gastrointestinal Surgery</i>	1
	<i>Orthopedics</i>	2
	<i>Spinal Procedures</i>	1

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

b. Interview Protocols

New topics covered in Model Year 7 site visits included decisions to rejoin BPCI Advanced, anticipated participation in future bundled payment models, interpretation of some self-reported patient care experience findings, and the sustainability of care redesign activities after BPCI Advanced has ended. We tailored interview protocols for each type of respondent. For example, we asked care redesign leaders targeted questions about care redesign activities, while we asked financial and executive leaders about entry decisions and financial performance. Topics covered in each interview are shown in Exhibit C.3.

Exhibit C.3: Model Year 7 Site Visit Interview Topics by Interviewee Type

Topics	Executive and Financial Leadership	Care Redesign Leadership	Care Coordination Staff	Patient Care (Nursing)	Patient Care (Physician)	Convenor
BPCI Advanced Patient Population	X	X	X	X	X	X
BPCI Advanced Successes and Challenges	X	X	X	X	X	X
Care Coordination	X	X	X	X		X
Care Redesign	X	X	X	X	X	X

Topics	Executive and Financial Leadership	Care Redesign Leadership	Care Coordination Staff	Patient Care (Nursing)	Patient Care (Physician)	Convener
CESLG Selection	X					X
Entry Decisions	X					X
Financial Performance	X				X	X
Impact of BPCI Advanced	X	X	X	X	X	X
Impact of COVID-19		X	X	X	X	X
Ongoing Convener Support of Episode Initiator		X				X
Overlap With Other Models	X	X	X	X	X	X
Role of Care Navigators/Care Coordinators			X			
Decision to Rejoin BPCI Advanced	X					X
Anticipated Participation in Future Bundled Payment Models	X					X
Self-Reported Patient Care Experiences Findings	X	X	X	X	X	
Sustainability of Care Redesign Activities	X	X	X		X	X

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

3. Key Informant Interviews

While site visits sought to gather information about how participants approached BPCI Advanced, KIIs provided the opportunity for more in-depth discussion about particular topics of interest to the Centers for Medicare & Medicaid Services (CMS) and the evaluation team. KIIs were 30- to 45-minute virtual interviews conducted with episode initiator and convener staff who were most knowledgeable about the topic of interest. All KIIs followed semi-structured protocols to keep the interview focused while allowing flexibility in topic exploration. The BPCI Advanced evaluation team conducted KIIs from Model Years 2 through 7. For information on KIIs conducted in Model Year 2, see **Appendix C** of the BPCI Advanced First Evaluation Report. For information on KIIs conducted in Model Year 3, see **Appendix C** of the BPCI Advanced Second Evaluation Report.

For information on KIIs conducted in Model Years 4, 5, and 6—and a round of KIIs conducted in Model Year 7 regarding how Medicare Accountable Care Organization (ACO) participation affected participant experience in BPCI Advanced—see **Appendix C** of the BPCI Advanced Sixth Evaluation Report.³

a. Sampling Approach

The second KII topic in Model Year 7 focused on episode initiators' perceptions of the model's target pricing methodology, financial factors influencing continued participation in BPCI Advanced, and perceptions and unintended consequences of the model's target pricing methodology.

The evaluation team identified a sample of 27 episode initiators (both downstream episode initiators and non-convenor participants) that met the following criteria:

- Participated in BPCI Advanced in Model Year 7
- Had at least 20 combined attributed episodes in Performance Periods 8 and 9 (July 1, 2022–June 30, 2023)
- Did not participate in a prior site visit or a recent KII (2023–2024)
- Did not share a convenor with more than two other downstream episode initiators in the sample (that is, the sample did not include more than three downstream episode initiators from the same convenor)

The sample included 10 episode initiators (9 hospitals and 1 PGP) that had participated in BPCI Advanced previously and rejoined under a new convenor or rejoined as a non-convenor participant in Model Year 7. We included these rejoined Model Year 7 participants because they could speak about their prior experience in the model.

In May and June of 2024, the evaluation team conducted phone interviews with 19 episode initiators (10 hospitals and 9 PGPs). In three instances, we interviewed a single convenor or health system on behalf of multiple episode initiators, representing eight episode initiators in total. Of the remaining 11 episode initiators interviewed, 4 shared one of two consultants, and 2 shared a convenor. The characteristics of the final interviewee sample are shown in Exhibit C.4.

³ The BPCI Advanced First, Second, and Sixth Evaluation Reports and their appendices, are available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Exhibit C.4: Participant Characteristics for Model Year 7 KIs on Target Pricing

KI Participant Characteristics		Count (N)
Episode Initiator Type	<i>Hospital</i>	10
	<i>PGP</i>	9
Participant Role	<i>Downstream Episode Initiator</i>	6
	<i>Non-convenor Participant</i>	13
Census Region	<i>Midwest</i>	3
	<i>Northeast</i>	2
	<i>South</i>	11
	<i>West</i>	3
Medical CESLG Type	<i>Cardiac Care</i>	10
	<i>Gastrointestinal Care</i>	1
	<i>Medicare & Critical Care</i>	12
	<i>Neurological Care</i>	4
Surgical CESLG Type	<i>Cardiac Procedures</i>	2
	<i>Gastrointestinal Surgery</i>	2
	<i>Orthopedics</i>	7
	<i>Spinal Procedures</i>	5

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

4. Qualitative Analysis

We used the same qualitative analysis methods for site visits and KIs. We recorded and transcribed site visits and KIs and then analyzed transcriptions using ATLAS.ti (Scientific Software Development GmbH, Berlin, Germany), a qualitative data analysis software.

The evaluation team used qualitative research codebooks, which outline the codes or categories that organize and analyze data in a qualitative study. Codebooks included a list of codes with the code's definition, an example of a response that belonged to that code, inclusion criteria, exclusion criteria, and code search expressions. We created unique codebooks for each round of KIs. While we used the same general site visit codebook each year, we modified the site visit codebook each year to incorporate improvements identified as a result of our data collection activities and new topics explored. Staff involved in coding or analysis were trained on the final codebook to promote a consistent approach and ensure a clear understanding of codes, and coding was reviewed to ensure inter-rater reliability.

We ran queries on the most common codes and analyzed qualitative data by subgroup to identify key themes within those topics. We defined subgroups based on organizational characteristics including episode initiator type, participant role, participation in other Medicare payment initiatives, and use of a convenor.

In addition to analyzing Model Year 7 qualitative data, the evaluation team reviewed past site visit and KI transcripts to inform this report. For example, we conducted thematic analyses and

compiled qualitative findings from site visits and KIs conducted with rural BPCI Advanced participants from Model Year 2 until Model Year 7, using ATLAS.ti to organize and review transcripts. There is potential for inconsistent qualitative coding. However, qualitative coders followed a detailed codebook that clearly defined each code, met weekly to discuss coding, and reviewed and revised coding to ensure inter-rater reliability.

5. Limitations

Qualitative methods have some limitations. One limitation is the potential for inconsistencies or bias in interviewees' responses, which may occur due to their knowledge base or inability to recall certain topics. We minimized this potential by confirming their role before the interviews to ensure that we spoke to people knowledgeable in BPCI Advanced and the specific research topics. When we heard inconsistencies in the responses of interviewees from the same organization, we asked more questions to clarify those inconsistencies. We ensured that the interviewees understood that we were evaluating BPCI Advanced, not their organization. This helped build rapport and put them at ease, which may have helped us obtain more honest answers.

Another consideration with qualitative research is that findings are not generalizable to all participants. By using a purposive sampling strategy, we ensured that we spoke to organizations with a range of characteristics (for instance, varied CESLG participation, size, and geographic location) that could best discuss the topic areas we were researching.

C. Patient Survey Methods

The BPCI Advanced patient survey explored differences in patient care experiences and functional outcomes between Medicare patients cared for by BPCI Advanced hospital and PGPs and a matched comparison group of patients whose hospital or PGP did not participate in BPCI Advanced. The patient survey collected information on a set of patient outcomes related to functional status, health care experience, and satisfaction with care and recovery. This section describes the instrument, sampling, administration, outcomes, and analysis of the patient survey.

1. Patient Survey Instrument

The BPCI Advanced evaluation's patient survey instrument (included in **Appendix D**) is a modified version of the survey instrument used in the BPCI Initiative evaluation,^{4,5} which was based on items adapted from validated survey instruments, such as the Continuity Assessment Record and Evaluation (CARE) Tool,⁶ National Health Interview Survey,⁷ and Short Form 36

⁴ The Lewin Group. (2018). *CMS Bundled Payments for Care Improvement Initiative Models 2-4: Year 5 evaluation and monitoring annual report*. <https://downloads.cms.gov/files/cmimi/bpci-models2-4-yr5evalrpt.pdf>

⁵ Trombley, M. J., McClellan, S. R., Kahvecioglu, D. C., Gu, Q., Hassol, A., Creel, A. H., Joy, S. M., Waldersen, B. W., & Ogbue, C. (2019). Association of Medicare's Bundled Payments for Care Improvement Initiative with patient-reported outcomes. *Health Services Research*, 54(4), 793–804. <https://doi.org/10.1111/1475-6773.13159>

⁶ RTI International. (2012). *The development and testing of the Continuity Assessment Record and Evaluation (CARE) item set*. https://hsr.himmelfarb.gwu.edu/smhs_crl_facpubs/103/

⁷ Centers for Disease Control and Prevention. (2012). National Health Interview Survey.

Health Survey.⁸ Based on input from clinical experts at CMS and the evaluation team, we added four new questions to better measure care experience and satisfaction with care. We adapted the new questions from the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Hospital Survey,⁹ Care Coordination Quality Measure for Primary Care survey,¹⁰ and B-Prepared instrument,¹¹ and they replaced five original questions on discharge timing and level of post-acute care received. The revised instrument underwent cognitive testing with a convenience sample of seven Medicare beneficiaries with recent inpatient and outpatient hospital experience. In Wave 5, we added a new open-ended question adapted from the CAHPS Narrative Item Set for the Clinician & Group Survey.¹²

The Wave 5 patient survey contained 27 closed-ended questions and 1 open-ended question, and it was designed to take an average of 25 minutes to complete. Survey questions covered a range of domains, including functional status, health care experience, and satisfaction with care and recovery (Exhibit C.5). For each of the seven functional areas, the survey asked respondents to recall their functional status before the anchor stay or procedure and to report their current functional status at the time they were completing the survey, which was at least 3 months after the anchor hospital discharge or outpatient procedure.

Exhibit C.5: Domain and Survey Items for Patient Survey

Domain	Description of Items
Functional Status	• Bathing/dressing/toileting/eating • Planning regular tasks • Use of a mobility device • Walking by self without resting • Walking up or down 12 stairs • Physical or emotional problems that interfere with social activities • Pain that interferes with normal activities

⁸ Brazier, J. E., Harper, R., Jones, N. M., O’Cathain, A., Thomas, K. J., Usherwood, T., & Westlake, L. (1992). Validating the SF-36 health survey questionnaire: New outcome measure for primary care. *BMJ (Clinical research ed.)*, 305(6846), 160–164. <https://doi.org/10.1136/bmj.305.6846.160>

⁹ Agency for Healthcare Research and Quality. (2018). *CAHPS Hospital Survey*. <https://www.ahrq.gov/cahps/surveys-guidance/hospital/index.html>.

¹⁰ Agency for Healthcare Research and Quality. (2016). *Care Coordination Quality Measure for Primary Care (CCQM-PC)*. <https://www.ahrq.gov/ncepcr/care/coordination/quality/index.html>.

¹¹ Graumlich, J. F., Novotny, N. L., & Aldag, J. C. (2008). Brief scale measuring patient preparedness for hospital discharge to home: Psychometric properties. *Journal of Hospital Medicine*, 3(6), 446–454. <https://doi.org/10.1002/jhm.316>

¹² Agency for Healthcare Research and Quality. (2021). *CAHPS Clinician & Group Narrative Item Set*. <https://www.ahrq.gov/cahps/surveys-guidance/item-sets/elicitation/cg-narratives.html>

Domain	Description of Items
Health Care Experience	• Respondent felt prepared to leave the hospital • Medical staff took patient preferences into account when arranging for health care services after leaving the hospital • Respondent had a good understanding of how to take care of herself or himself prior to leaving the hospital • Medical staff clearly explained how to take medications • Medical staff clearly explained needed follow-up appointments • Respondent and caregiver able to manage their health care needs • Medical staff discussed whether patient would have the help they needed when they got home • If help needed at home, medical staff arranged for services at home to help patient manage health
Satisfaction	• Overall satisfaction with recovery • Rating of all post-hospital care from 0 to 10
Personal Characteristics	• Highest level of education
Open-Ended Care Experience	• Description of what went well in experiences with all health care received during episode

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

2. Patient Survey Sample

a. Timing of Survey Wave

The Final Evaluation Report presents results from Wave 5 of the BPCI Advanced patient survey, which covered episodes with discharges or outpatient procedures between April 2024 and January 2025. We fielded the survey from July 2024 until May 2025 to patients with episodes initiated by BPCI Advanced and comparison acute care hospitals and PGPs. Patients received surveys about 90 days after leaving the hospital.

b. Sample Frame

The patient survey used a stratified random sampling method to obtain a representative sample of the BPCI Advanced population and a matched comparison group. We created the sampling frame using Medicare fee-for-service (FFS) claims from “rolling” 1-month samples because a single month of claims was not adequate to reach the necessary sample size at the levels of clinical precision used to define the strata. For example, we pulled claims in May 2024 for April 2024 discharges and procedures and mailed the surveys to the monthly sample 2 months after drawing the sample (July 2024). We repeated this process for all 10 months of data collection, drawing the last sample in February 2025 and mailing surveys in April 2025. We used this rapid sampling process to reduce recall bias. This process also had the effect of limiting the sample to patients whose claims were filed relatively quickly, within 1 month of discharge or outpatient procedure.¹³ We stratified the sample by CESLG, separately for hospital and PGP episode initiators. We did

¹³ Although claims submitted within 1 month may not represent the entire Medicare population within a stratum due to provider delays in submitting claims, this issue should affect BPCI Advanced and comparison samples equally and not bias our estimates.

not sample *cardiac procedures, gastrointestinal surgery, or neurological care* episodes for PGP episode initiators because these CESLGs had either no volume or such low volume that we could obtain only a handful of survey responses. This approach yielded 13 unique strata. Within our CESLG-based strata, we also oversampled certain groups of patients to ensure all corresponding analyses were adequately powered.

c. Sample Construction

We constructed the hospital and PGP samples with the goals of (1) selecting a representative sample of BPCI Advanced discharges and outpatient procedures and (2) identifying and selecting an appropriate comparison group of discharges and procedures. We started with the universe of BPCI Advanced hospitals and nonparticipating hospitals. We then excluded hospitals that were not eligible for BPCI Advanced and additionally excluded hospitals in various ways to limit exposure of the comparison group to the BPCI Advanced Model and the Comprehensive Care for Joint Replacement (CJR) Model. We describe these exclusions in more detail below.

Our sampling universe comprised all discharges or outpatient procedures that met BPCI Advanced Model rules (for example, no patients assigned to ACOs participating in ACO REACH Model; no discharges or outpatient procedures from hospitals in the Maryland All-Payer Model). We followed identical steps to construct the hospital and PGP samples from within the sampling universe, because the PGP comparison group was based on all eligible hospital discharges or outpatient procedures, not those only attributed to comparison PGPs (that is, we did not identify comparison PGPs from which to sample patients with episodes). We constructed the hospital and PGP survey samples in four steps.

Step 1 – Excluding hospitals: All BPCI Advanced episode initiators active through the end of a given sampling month were eligible for our sample (that is, if we were sampling episodes with a discharge or procedure in July of Model Year 7, then all episode initiators active through July of Model Year 7 were eligible for our sample). We generated a comparison group pool specific to each clinical episode by applying five types of exclusions by clinical episode and CESLG. We excluded hospitals from the clinical episode comparison group pool if:

- They were BPCI Advanced participants for any clinical episode sharing the same CESLG.
- They were ineligible to participate in BPCI Advanced due to low baseline volume for a given clinical episode. For example, a comparison group hospital with a low baseline volume of major joint replacement of the lower extremity (MJRLE) surgeries was excluded from the MJRLE comparison group but would be eligible for inclusion in the sepsis comparison group if this hospital had sufficient sepsis discharges.
- More than 10% of their baseline episodes within the clinical episode or corresponding CESLG were touched by BPCI Advanced participants (hospital or PGP).
- They were located in a core-based statistical area where more than 50% of baseline episodes in a given clinical episode were touched by BPCI Advanced participants (hospital or PGP).

- Finally, CJR hospitals were excluded from the comparison group pool for all clinical episodes in the *orthopedics* CESLG. This CESLG includes MJRLE, the clinical episode corresponding to the CJR Model.

Step 2 – Excluding individual hospital discharges and outpatient procedures: We excluded individual discharges and outpatient procedures from the comparison group to further reduce the risk of contamination from other bundled payment episodes. Specifically, we excluded discharges and procedures for which:

- The attending or operating physician listed on the claim belonged to a BPCI Advanced PGP, but the discharge or procedure was not attributed to BPCI Advanced.
- The patient was in a BPCI Advanced or CJR episode at the time of the discharge or outpatient procedure (meaning the discharge or outpatient procedure occurred within 90 days after the start of a BPCI Advanced or CJR episode).
- The patient was treated in any hospital by a physician belonging to a BPCI Advanced PGP in the prior 90 days.
- The patient had more than one discharge or procedure in the month of our data. This exclusion ensured that a patient could only be selected into the sample one time and that the survey we mailed to them referenced their most recent hospitalization. In that case, only the most recent discharge or procedure was eligible for selection into our sample.

We applied the last two exclusions to the BPCI Advanced group as well as the comparison group to ensure that the characteristics of the patients would be similar between the two groups.

Step 3 – Selecting BPCI Advanced patients: We created sampling cells of unique combinations of clinical episodes based on age category (<65, 65–74, 75–84, 85+ years) and the presence of a major complication or comorbidity for each stratum. We selected a random sample of BPCI Advanced patients within each sampling cell.

Step 4 – Selecting matched comparison hospital discharges and outpatient procedures: Lastly, we matched each selected BPCI Advanced patient one to one with a comparison patient from the same sampling cell. Within sampling cells, we matched comparison patients to a given BPCI Advanced patient using a nearest-neighbor propensity score approach. We estimated propensity scores of being in BPCI Advanced based on the patient, hospital, and market factors described in Exhibit C.6.

Exhibit C.6: Predictive Risk Factors Used to Match BPCI Advanced and Comparison Patients

Domain	Variables
Service Mix^a	- Clinical episode - Lower body fracture (MJRLE, orthopedics and related procedures) - Knee procedure (MJRLE episode only) - Large vessel ischemic stroke (stroke episodes only) - Intracerebral hemorrhage (stroke episodes only) - Major complication or comorbidity
Patient Demographics and Enrollment	- Age (under 65, 65–74, 75–84, 85+ years) - Sex - Dual eligibility status - Originally qualified for Medicare due to disability - Additional patient characteristics
Prior Utilization Measures	Any inpatient admission in the prior 90 days
Discharging Hospital Characteristics	2017 linear HCAHPS score^b - Academic medical center - Bed size (≤250; 251–500; 501–850; >850) - Safety-net status - Census region (Northeast, Midwest, South, West) - Urban - Ownership (for profit, not for profit, government/other)
Neighborhood and Market Characteristics	- ADI^c - Mean personal income (CBSA level) - Percentage of population older than 65 years (county level) - Rurality^d

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms. We used predictive risk factors to estimate propensity scores of being in BPCI Advanced.

^a Additional variables for MJRLE, orthopedics and related procedures and for stroke episodes control for clinical heterogeneity that is not accounted for by MS-DRGs and that is easily identifiable from ICD-10 codes.

^b The linear HCAHPS score captures patient ratings of their overall experience with a hospital from 0 to 100, adjusted for patient mix and HCAHPS Survey mode. We used 2017 data to avoid possible contamination of HCAHPS responses attributable to BPCI Advanced.

^c The ADI is a measure of socioeconomic status developed by researchers at the University of Wisconsin.¹⁴

^d We added rurality as a matching variable beginning in Wave 3. Rurality indicates that a patient resided in a ZIP Code designated as rural by the Federal Office of Rural Health Policy. We obtained eligible ZIP Codes in July 2022: <https://www.hrsa.gov/rural-health/about-us/what-is-rural/data-files>.

3. Administration of the Patient Survey

We mailed each sampled patient a paper survey, a postcard reminder, and, for patients who did not respond to the initial mailings, a second paper survey using priority mail. We mailed the first survey to patients within about 90 days of when they left the hospital. Beginning in Wave 5, we included a QR code that allowed respondents the option of completing the survey online. We

¹⁴ University of Wisconsin School of Medicine and Public Health. (n.d.). *2015 Area Deprivation Index Version 2.0*. <https://www.neighborhoodatlas.medicine.wisc.edu>

present sample sizes, response rates, and response data (for instance, web versus mail response) in Section C.5.

4. Closed-Ended Outcome Measures

The BPCI Advanced patient survey instrument asked about seven measures of physical function, and for each, it asked respondents to recall their status before the anchor hospitalization or procedure (questions 2 through 8) and to report their current functional status at the time of the survey (questions 9 through 15). The seven functional status measures asked about (1) bathing, dressing, toileting, and eating; (2) planning regular tasks; (3) moving using a mobility device; (4) walking without resting; (5) going up or down stairs; (6) the frequency with which physical or emotional health interferes with regular social activities; and (7) the frequency with which pain interferes with normal activities.

For each functional status measure, we created trinary measures for improvement, maintenance, or decline in initial function. We marked the outcome as *improved* if a patient moved to a better functional status level after the episode (for example, from “complete help needed” before the episode to “no help needed” after the episode) or if the patient recalled having the highest functional status prior to hospitalization and remained in that high status at the time of survey response (for example, “no help needed” both before hospitalization and after the episode). We marked the outcome as *maintained* function if the patient did not recall the highest or lowest function prior to hospitalization and reported that their function was the same before the episode and at the time of the survey. We marked the outcome as *declined* if the patient moved to a worse functional status level after the episode or if the patient recalled having the lowest functional status prior to hospitalization and remained in that low status at the time of the survey.

The BPCI Advanced survey asked eight questions about care experience and two about satisfaction with recovery and care received. We constructed binary outcome measures for all questions, except for the rating of all post-hospital care, from which we created a trinary outcome. More detail on measure specifications for these two domains is shown in Exhibit C.7. All questions and possible responses to each question are available in **Appendix D**.

Exhibit C.7: Definitions for Measures of Care Experience and Satisfaction

Domain	Outcome Measure	Response (Positive Response for Binary Outcomes)
Care Experience	Felt “very” or “somewhat” prepared to leave the hospital	Yes
	Medical staff took your preferences into account in deciding what health care services you should have after you left the hospital	Agree/strongly agree
	Good understanding of how to take care of self before going home	Agree/strongly agree
	Medical staff clearly explained how to take medications before going home	Agree/strongly agree
	Medical staff clearly explained what follow-up appointments or treatments would be needed before going home	Agree/strongly agree

Domain	Outcome Measure	Response (Positive Response for Binary Outcomes)
Care Experience (Continued)	Able to manage your health needs since returning home	Agree/strongly agree
	Medical staff talked with you about whether you would have the help you needed when you got home	Yes
	Medical staff arranged services for you at home to help manage your health, if you needed it	Yes
Satisfaction With Care and Recovery	Overall satisfaction with recovery since leaving hospital	Extremely satisfied/quite a bit satisfied
	Rating of all post-hospital care from 0 to 10 ^a	High (rating 9–10), middle (7–8), low (0–6)

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

^a The rating of post-hospital care was a trinary measure.

5. Analysis of the Patient Survey

a. Closed-Ended Survey Analysis

Analytic Approach

We estimated differences between BPCI Advanced and comparison patients averaged across all CESLGs within each group. We used logistic regression to estimate risk-adjusted differences in binary survey outcomes between the BPCI Advanced and comparison respondents. We used multinomial logistic regression to estimate differences for the trinary survey outcomes and estimated the joint statistical significance of differences across all three categories using standard errors clustered at the hospital level.

Analytic Groups

We analyzed all hospital- and PGP-initiated episodes in aggregate and separately stratified results by medical and surgical CESLGs. We report analytic sample sizes, response rates, and minimum detectable differences for these levels of analysis in Exhibit C.8. Exhibit C.9 shows the proportion of respondents who responded online through the QR code, the proportion who had their survey completed by a proxy, and the average time that elapsed between anchor hospital discharge or outpatient procedure and our receipt of survey response. The standardized mean difference provides a scale of how different the two groups are for a given measure relative to the “natural” amount of variation in that measure. A difference of 0.10 or less is a conservative threshold for identifying potentially problematic differences between two groups.¹⁵

¹⁵ Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399–424.

Exhibit C.8: Wave 5 Response Rates, Sample Size, and Minimum Detectable Difference

Level of Analysis	BPCI Advanced Response Rate (%)	Comparison Response Rate (%)	Difference (pp) [P-Value]	BPCI Advanced Respondents	Minimum Detectable Difference
All Clinical Episodes	27.1	28.0	-0.9 [0.369]	6,775	2.1
Medical	21.9	22.4	-0.5 [0.552]	3,867	2.8
Surgical	39.6	41.6	-1.9 [0.089]	2,908	3.2

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

Exhibit C.9: Wave 5 Response Method by Clinical Episode Type

Level of Analysis	Response Measure	BPCI Advanced	Comparison	Difference	Standardized Mean Difference
All Clinical Episodes	<i>Proxy Response (%)</i>	8.6	8.1	0.4	0.02
	<i>Web Response (%)</i>	8.8	7.7	1.1	0.04
	<i>Days Elapsed (Days)</i>	114.7	114.5	0.1	0.01
Medical Episodes	<i>Proxy Response (%)</i>	12.1	11.9	0.2	0.01
	<i>Web Response (%)</i>	7.4	6.2	1.2	0.05
	<i>Days Elapsed (Days)</i>	115.3	115.1	0.2	0.01
Surgical Episodes	<i>Proxy Response (%)</i>	3.8	3.2	0.6	0.03
	<i>Web Response (%)</i>	10.7	9.7	1.0	0.03
	<i>Days Elapsed (Days)</i>	113.9	113.8	0.1	0.00

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

Weighting and Risk Adjustment

For each of the strata, we calculated entropy-balanced weights representative of the BPCI Advanced respondents to improve the generalizability of results. The weights account for the possibility that BPCI Advanced yields different outcomes for different types of patients. The purpose of weighting the comparison group to reflect the BPCI Advanced group is to obtain “doubly robust” estimates of the difference between BPCI Advanced and comparison respondents.¹⁶ This means if either the regression or weights are correctly specified, our estimates should still be unbiased; they do not both need to be correctly specified.

Within each CESLG, we weighted the analytic data in two stages. First, we calculated entropy-balancing weights¹⁷ that made the BPCI Advanced respondents representative of the BPCI Advanced population (that is, the sampling frame) based on the risk factors described in Exhibit C.10. Second, we calculated entropy-balancing weights that made the comparison

¹⁶ Robins, J. M., Rotnitzky, A., & Zhao, L. F. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association*, 89(427), 846–866.

¹⁷ Hainmuller J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20, 25–46.
Hainmuller, J., & Xu, Y. (2013). ebalance: A Stata package for entropy balancing. *Journal of Statistical Software*, 54(7), 1–18.

respondents representative of the (weighted) BPCI Advanced respondents, such that both groups reflected the BPCI Advanced population after applying the survey weights.

Additionally, we performed regression-based risk adjustment to ensure comparability between the BPCI Advanced and comparison groups, which included the factors listed in Exhibit C.10.

Exhibit C.10: Predictive Risk Factors Used to Risk-Adjust Survey Outcomes

Domain	Variables
Service Mix^a	- Clinical episode type - Major complication or comorbidity - Lower body fracture (MJRLE, orthopedics and related procedures) - Knee procedure (MJRLE episodes only) - Large vessel ischemic stroke (stroke episodes only) - Intracerebral hemorrhage (stroke episodes only)
Patient Demographics and Enrollment	- Age (under 65, 65–74, 75–84, 85+ years) - Sex - Dual eligibility status 4-year degree or higher - Aligned to CMS population-health program (for example, ACO, advanced primary care) - Additional patient characteristics
Prior Health Conditions	- Hierarchical Condition Category (HCC) score: index score based on the patient's diagnosis information during the 6 months prior to the episode start date, using v22 of CMS' 2019 Risk Score software and diagnosis to chronic condition mappings - Squared HCC score - Functional status using three summary measures^b
Prior Utilization Measures	- Any inpatient admission in the prior 6 months - Any other institutional care (SNF, IRF, LTCH, or psychiatric hospital) in prior 6 months - Any nursing home care in the prior 6 months
Discharging Hospital Characteristics	2017 linear HCAHPS score^c
Neighborhood Characteristics	ADI^d
Survey Dimensions	- Proxy status (patient had help from someone else in responding to the survey) - Survey mode (response obtained via mail vs. online) - Days elapsed between leaving the hospital and survey response - Survey month fixed effects

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms. We used predictive risk factors to estimate propensity scores of being in BPCI Advanced.

^a Additional variables for MJRLE, orthopedics and related procedures and for stroke episodes control for clinical heterogeneity that is not accounted for by MS-DRGs and that is easily identifiable from ICD-10 codes.

^b Three of the functional status questions have only three possible responses, two functional status questions have four possible responses, and two questions have five possible responses. For each of the outcomes with less than five possible responses, the best functional status was coded as 1, the middle status (or two statuses) was coded as 2, and the worst status was coded as 3. We created a variable summing the number of functional status measures with 2 and the number with 3 as well as a binary indicator for “missing functional status.” For the two measures with five possible responses, we created binary indicators for “all of the time/most of the time” and created a control variable summing the number of indicators equal to 1, as well as a binary indicator for “missing activity status.” For functional status variables with four possible

responses, we considered alternative cutoffs for coding responses as 1, 2, or 3; however, none of these alternative cutoffs altered the results in any meaningful way.

- ^c The linear HCAHPS score captures patient ratings of their overall experience with a hospital from 0 to 100, adjusted for patient mix and HCAHPS survey mode. We use 2017 data to avoid possible contamination of HCAHPS responses attributable to BPCI Advanced.
- ^d The ADI is a measure of socioeconomic status developed by researchers at the University of Wisconsin.

The risk-adjustment model accounts for certain factors that could not be incorporated into our matching algorithm applied at the time of sampling.¹⁸ We also matched on six attributes of the discharging hospital, five of which were like those used by CMS to define target price peer groups but were not included in our final risk-adjustment model.¹⁹

For our aggregate analyses, the regression model for each outcome can be expressed as:

$$Y_{ijk} = \delta BPCIAdvanced_{ij} + \beta_k X_{ij} + CE_i + \varepsilon_{ijk}$$

Y_{ijk} is the outcome of interest for individual i , treated at provider j , in clinical episode k . X refers to the risk-adjustment variables (listed above), CE indicates individual indicators for each clinical episode, and $BPCIAdvanced$ is an indicator for a patient who was treated by a BPCI Advanced participating hospital or PGP. The relationship between Y and $BPCIAdvanced$ (indicated by δ) represents the difference between BPCI Advanced and comparison respondents across all clinical episodes.

Domain-Level Joint Hypothesis Testing

To better understand survey outcomes at the domain level (that is, collectively across all functional status measures, and all care experience and satisfaction measures) we calculated p-values for each of these two domains. These p-values reflect the test of the hypothesis that all estimated differences between BPCI Advanced and comparison respondents within a domain are jointly equal to zero.

A standard test of joint statistical significance entails calculating the “Wald Statistic,” which can be expressed as:

$$(\hat{\beta} - \beta_0)' [\text{COV}(\hat{\beta})]^{-1} (\hat{\beta} - \beta_0)$$

where:

- $\hat{\beta}$ is the vector of true marginal effects of all measures within a domain.

¹⁸ For example, we did not have information about education and prehospital functional status at the time of sampling; those data come from the survey responses. Likewise, factors such as Hierarchical Condition Category (HCC) index score and recent institutional care could not be reliably identified at the time of survey sampling because additional claims runout time would be required, and waiting for the data could delay the survey and increase recall bias.

¹⁹ Results from the original BPCI Initiative evaluation indicated that these hospital-level factors were not strongly correlated with survey outcomes. Matching on these factors allows us to ensure that the BPCI Advanced and comparison groups are reasonably similar with regard to these factors, without the loss of statistical precision (that is, larger standard errors) that would likely result from directly controlling for such measures that only weakly predict survey outcomes.

- β_0 is a vector of zeros (since our null hypothesis is that the estimates do not differ from zero).
- $\text{COV}(\hat{\beta})$ is the covariance matrix of the marginal effects.

The covariance matrix captures the variance of each marginal effect i (σ_{ii}), and the covariance between each marginal effect i and j (σ_{ij}):

$$\text{COV}(\hat{\beta}) = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1j} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{j1} & \sigma_{j2} & \cdots & \sigma_{jj} \end{pmatrix}$$

While estimated marginal effects ($\hat{\beta}$) and estimated variances of the marginal effects (σ_{ii}) are standard output from regressions in most statistical software, estimating the covariance between marginal effects (σ_{ij}) requires joint estimation of outcomes. This can be infeasible when many outcomes and the regressions are nonlinear, as is the case for our domain-level analysis with binary outcomes.

To overcome this limitation, we adapted a method proposed by Mullahy (2016).²⁰ Under this approach, we ran bivariate regressions on every combination of outcomes within a domain and used these to retrieve the covariance between all estimated marginal effects. We used these covariance terms to construct the covariance matrix for each domain, allowing us to compute the Wald Statistic, from which we could retrieve the corresponding p-value from the χ^2 distribution.

Open-Ended Survey Question

“Describe in your own words, your time at the hospital listed in the cover letter and all health care you received after leaving the hospital. What has gone well in your experiences with all health care you received since you went to the hospital listed in the cover letter? Please explain what happened, how it happened, and how it felt to you.”

b. Open-Ended Survey Analysis

To better understand why BPCI Advanced patients reported slightly less favorable experiences compared with patients who did not receive care under the model, we added an open-ended survey question about patient care experiences to Wave 5 of the patient survey. Our objective was to identify key themes in patient-reported experiences that may be, in part, attributed to BPCI Advanced–related care redesign. There were 10,834 total responses to the open-ended survey question.

Identifying Themes and Sentiments

We used a generative artificial intelligence (AI) application within the secure Optum Serve IQ platform to identify thematic concepts, or codes, and sentiments in deidentified responses to the open-ended survey question.

²⁰ Mullahy, J. (2016). Estimation of multivariate probit models via bivariate probit. *The Stata Journal*, 16(1), 37–51.

To identify relevant codes, we developed a prompt with context as input for the generative AI. The context included the objectives and structure of the BPCI Advanced Model, the purpose of the patient survey, description of the relevant parts of the patient journey (before, during, and after the hospital stay or outpatient procedure, discharge process, and the discharge destination), how we administered the patient survey, and the text of the open-ended survey question. The generative AI used the prompt and context to identify relevant thematic codes about care experiences related to the patient's episode. The output had three components: the codes assigned to the response, the sentiment associated with each code, and the AI's rationale for assigning the code and sentiment to the response. A single survey response could have had multiple codes applied.

Since multiple codes could have been assigned to a single survey response, the generative AI assigned a positive, negative, or neutral sentiment to each code. Therefore, a single survey response could have multiple codes applied with different sentiments associated with each code. For example, a survey response could have a code about communication with a positive sentiment and a code about pain management with a negative sentiment. Because the survey question asked about what went well in patients' experiences with their health care, the generative AI assigned a positive sentiment when the response did not include a clear sentiment. The evaluation team likewise assumed a positive sentiment during the quality review process described below.

The evaluation team used part of the data (n = 6,815 survey responses) to identify codes. We then applied these codes to the remaining 4,019 responses from the second part of the data. We manually reviewed any responses from the second part of the data that were not assigned at least one code to determine whether there were additional codes in the second sample of responses that were not in the preliminary sample. This approach is a standard practice that accounts for the computational cost required to identify codes in large datasets using AI. We did not find any evidence of additional codes in the second dataset.

Validating Themes and Sentiments

To ensure that codes were relevant and accurate to the research question, we reviewed the list of codes, the definitions of the codes, and a sample of example survey responses that were categorized into each code.²¹ All initial codes aligned with the research question.

To determine whether the generative AI correctly assigned codes to each response, the evaluation team reviewed a sample of responses for each code. We reviewed approximately 5% of responses assigned to that code and a comparable number that were not. For each response, the reviewer classified the responses as correctly categorized (true positive [TP]), correctly not categorized (true negative [TN]), incorrectly categorized (false positive [FP]), or incorrectly not categorized (false negative [FN]). Our goal was to achieve 90% accuracy for the codes assigned

²¹ An example of a code that did not align with the research question would describe satisfaction with meals in the hospital.

by the AI platform. We iteratively revised the prompt to improve the accuracy with which codes were applied to responses if we did not achieve 90% accuracy.

We used the following equation to calculate accuracy:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

The evaluation team used a similar approach to review for the sentiment associated with each code within the sample of survey responses. A sentiment was only applied, and thus reviewed, if a response was assigned to the code. For each response, the reviewer classified the sentiment as sentiment correctly categorized (TP) or incorrectly categorized (FP). For each code, we calculated the percentage of responses for which the AI assigned a correct sentiment. For the sample responses reviewed, the generative AI assigned the correct sentiment to at least 90% of responses.

c. Limitations

Closed-Ended Survey Analysis

Our analysis of the patient survey results identified favorable, unfavorable, and negligible differences in responses between BPCI Advanced and comparison respondents. Because we only collected the survey data during the performance periods, we cannot determine whether any differences existed during the baseline period. Additionally, given the patient survey's sample sizes, we may not be able to draw statistically significant conclusions for results that are small in magnitude. Although our aggregate analyses are powered to detect differences in the range of 2.1 to 3.2 percentage points, consistent differences of lesser magnitude may be meaningful and important.

We used the same patient survey instrument for all clinical episodes. Different clinical episodes have unique aspects of recovery and medical needs after discharge. Although a common instrument was important for understanding the model as a whole, we may not be able to observe nuanced changes in patient-reported outcomes associated with model participation because we did not tailor the survey instrument to each clinical episode.

Furthermore, due to the nature of primary data collection, the patient survey may not be representative of all BPCI Advanced patients. BPCI Advanced response rates for the patient survey were about 27.1% overall, and as high as 39.6% for surgical episodes. Although we applied nonresponse and sampling weights to all observations to make our respondents reflect the overall BPCI Advanced population on key characteristics of patients, hospitals, and neighborhoods, we cannot guarantee that our results are generalizable to the remaining BPCI Advanced patients who did not respond to the survey.

Open-Ended Survey Analysis

Responses to the open-ended survey question depended on the respondents' understanding of the survey question and willingness to describe their perspectives about their experiences with their health care. Open-ended survey responses reflect how the respondents perceived their

treatment and interactions with health care providers. Consequently, our results are limited to care delivery that patients had an opportunity to observe. For example, respondents' descriptions of care included coordination between and support from different health care providers (doctors, surgeons/specialists, nurses, physical/occupational therapists), facilities (including the hospital, rehabilitation facility, or home health agency), or other services during the patient care journey. As a result, the generative AI initially construed care coordination broadly, including cases where patients received support from different providers without indication of the extent to which the providers coordinated. For the purposes of BPCI Advanced, care coordination focuses more narrowly on communication and collaboration throughout a clinical episode, some of which is invisible to the patient. The iterative process we used to refine code definitions with the generative AI tool aimed at correcting possible miscategorization.

Although the open-ended question was designed to capture positive care experiences, many respondents described negative experiences and challenges. Some responses included both positive and negative experiences. Our human reviewers concluded that the generative AI tool successfully parsed the nuance of multiple sentiments across multiple codes within a response, but it is possible we misinterpreted sentiments or that respondents could not clearly communicate their sentiment.

The results of the analysis of open-ended question responses cannot be generalized to all BPCI Advanced and comparison group respondents. The experiences reported do not reflect all possible experiences for BPCI Advanced and comparison group respondents.

Despite our comprehensive manual review of the accuracy with which the generative AI assigned codes and sentiments to responses, the generative AI may have inaccurately applied codes to some responses. Although the accuracy rate exceeded 90% in our quality review for all codes, an accuracy rate below 100% indicates there were still some responses inaccurately categorized into or excluded from certain codes.

Lastly, the survey asked respondents about what went well to capture positive experiences but did not include a complementary question about what did not go well. Although 53% of respondents (BPCI Advanced and comparison group) described entirely positive experiences, 26% described experiences with more than one sentiment (positive, negative, and/or neutral), and 8% described entirely negative experiences (the remaining 13% did not have enough information to assign a category or sentiment). We likely did not capture important negative experiences among many BPCI Advanced and comparison group patients.

D. Quantitative Methods for Secondary Data

This section discusses the quantitative methods used for analyzing secondary data to evaluate the model. First, we discuss how we conducted descriptive analyses of our sample of patients and participating providers. In the second and third sections, we discuss our strategy for assessing the impact of the model, including the outcomes, study samples, selection of matched comparison providers and episodes, and difference-in-differences (DiD) approach. We present the methods

for Model Year 4 through 6 impacts and Model Year 7 and 8 impacts separately. Finally, we describe how we calculated Medicare program savings.

1. Descriptive Analysis

a. Participation

We assessed changes in participation in BPCI Advanced in Model Years 1 through 8 for hospitals and PGPs. We used the CMS BPCI Advanced Participants Database to examine the number of participating hospitals and PGPs, convener status, convener types, and clinical episode and CESLG selections. We used the Medicare FFS Claims, the Master Beneficiary Summary File, and the IMC Episode-Level Files to identify PGPs that were active or inactive in the model, to assess reach of BPCI Advanced to eligible discharges and clinicians, to analyze episode volume, and to identify hospitals and PGPs with at least one episode in the CESLGs they selected.

For our analyses of participation, we defined PGPs (identified by Taxpayer Identification Numbers [TINs]) as inactive if they had no BPCI Advanced eligible discharges or procedures from 2013 through 2025. All participation results that we report excluded inactive TINs. For calculating the shares of participating hospitals and PGPs in medical or surgical clinical episodes and CESLGs, we limited the denominator to participants that had at least one discharge or procedure in medical or surgical clinical episodes and CESLGs for the model year to ensure participants had some engagement in their chosen CESLGs.

We identified unique conveners based on the convener's TIN, meaning that two different participation agreements with the same TIN were considered a single convener. We did this because many conveners had separate participation agreements with each of their downstream episode initiators, allowing them to stop participating with one episode initiator while continuing to participate with others. The counts of hospitals and PGPs under and not under a convener excluded conveners that were also a participant.

Reach of the BPCI Advanced Model

We assessed the reach of the model to eligible hospitals, clinicians, and discharges. Eligible hospitals were Inpatient Prospective Payment System (IPPS) hospitals that met the minimum hospital baseline volume criterion to participate in BPCI Advanced (and thus received a target price) for at least one clinical episode. The sample excluded hospitals that met any of the following criteria: PPS-exempt cancer hospital, inpatient psychiatric hospital, critical access hospital, hospital located in Maryland, or hospital participating in the Pennsylvania Rural Health Model or in the Rural Community Health Demonstration.

We defined eligible clinicians as attending and operating National Provider Identifiers (NPIs) who treated Medicare patients meeting the BPCI Advanced inclusion criteria at a BPCI Advanced eligible hospital (the minimum hospital baseline volume criterion was not applied). We defined clinicians who participated in BPCI Advanced as (1) any attending or operating NPI at a BPCI Advanced hospital for a clinical episode in which the hospital was participating or (2) any attending or operating NPI on the hospital claim when the patient had a corresponding Part B

claim during the anchor stay or procedure (including 1 day prior) where the BPCI Advanced PGP TIN was the billing provider and the PGP was participating in the given CESLG.

Eligible discharges and procedures were hospital discharges and outpatient procedures that met the model eligibility requirements, including having a BPCI Advanced Medicare Severity-Diagnosis Related Group (MS-DRG) or Healthcare Common Procedure Coding System (HCPCS) code, being at a hospital that was eligible to be in BPCI Advanced, and meeting other beneficiary inclusion criteria. The minimum hospital baseline volume criterion was not applied to eligible discharges and procedures.

Limitations

The number of conveners, hospitals, and PGPs is based on the model's participation data. We counted unique conveners using TINs to prevent double counting single legal entities that entered into multiple participation agreements. However, it is possible that a single organization had participation agreements using multiple TINs and was counted as multiple conveners. Likewise, providers that participated as non-conveners may have had arrangements that mirrored those between a convener and its episode initiators. For example, most hospitals that participated in the model were part of a health system; even those whose health system was not a convener likely still coordinated participation decisions with health system leadership and had some degree of financial protection from the broader health system. Additionally, many non-convener participants worked with convener-like organizations, often consulting firms, which provided some of the same organizational and analytic support as conveners.

b. Hospital Characteristics

Safety-Net Hospitals

We defined safety-net hospitals as those above the 75th percentile in share of Medicare patients dually eligible for Medicaid or the Part D Low Income Subsidy (LIS) in 2021²²; the 75th percentile for each was roughly 30%. This definition does not match the definition used by the model to assign providers to peer groups for target pricing purposes: The model identifies a hospital as a safety-net hospital if at least 60% of its Medicare patient population is eligible for Medicaid.

Hospital Finances

We obtained data on hospital revenue and expenditures from the CMS Hospital Cost Reports, which are filed each year by hospitals participating in Medicare and provide detailed information about hospital finances and related characteristics. We constructed a calendar-year-by-hospital dataset by taking a weighted average of cost reports (based on various fiscal years) that overlapped with a given calendar year and applied various data cleaning steps.

²² We opted to define safety-net hospitals this way, instead of assessing safety-net status separately for each year (and comparing the hospital's dual eligibility and LIS eligibility percentages for that year with that year's 75th percentiles), to obtain a constant sample from year to year. This way we could (for example) calculate the median combined per-episode net payment reconciliation amounts across Model Years 4 through 6 for safety-net versus non-safety-net hospitals.

We defined a hospital's *operating margin* as total operating revenue (meaning net revenue from patient services) less total operating expenses, divided by total operating revenue.²³

Limitations

There are many ways of defining safety-net hospitals, and concordance between different definitions is often low.²⁴ A different definition of safety-net hospitals might have produced different results.

Our data on hospital finances were limited because cost reports have well-documented data errors and inconsistencies.²⁵ Although we performed extensive data cleaning, we ultimately did not have any external way to validate the data.

c. Participant Reconciliation

We analyzed reconciliation data provided by the IMC to describe financial performance in the model. We used the most recent reconciliation data available. For Model Years 1 through 6 and the first quarter of Model Year 7 (Performance Period 11), we used finalized second true-up data. For the second and third quarter of Model Year 7 (Performance Period 12), we used the first true-up data. For the final quarter of Model Year 7 and the first half of Model Year 8 (Performance Period 13), we used the initial reconciliation data.

We calculated the average per-episode reconciliation payment, our measure of financial performance in the model, by summing the reconciliation amount earned by each hospital and PGP in a given period and dividing by the number of episodes. We then took the average or quantiles of these per-episode amounts to produce the reported distributions and sample statistics.

²³ There are multiple ways to calculate operating margins from hospital cost reports, which can differ considerably; see, for instance: Ma, J., Gurzenda, S., Reither, K., & Pink, G. (2024). *Differences in measurement of operating margin: An update*. N.C. Rural Health Research Program, University of North Carolina. <https://www.shepscenter.unc.edu/product/differences-in-measurement-of-operating-margin-an-update/>. Our method is what the Ma et. al. (2024) refer to as “patient care only.”

²⁴ See, for example: Chatterjee, P., Liao, J., Amagai, K., Zhao, Y., Shirk, T., & Navathe, A. (2025). Variation, overlap, and stability in defining safety net hospitals. *JAMA Network Open*. 2025;8(7):e2523923. <http://doi.org/10.1001/jamanetworkopen.2025.23923>; Popescu, I., Fingar, K., Cutler, E., Guo, J., & Jian, J. (2019). Comparison of 3 safety-net hospital definitions and association with hospital characteristics. *JAMA Network Open*. 2019;2(8):e198577. <http://doi.org/10.1001/jamanetworkopen.2019.8577>

²⁵ For instance, see Rhodes, J., Santos, T., & Young, G. (2023). Hospitals' uneven recovery from the COVID-19 pandemic. *Health Affairs Scholar*, 1(3), 1–6. <https://doi.org/10.1093/haschl/qxad034> (cost reports “have several well-documented limitations, including missing and inaccurate submissions”); Lamboy-Ruiz, M., No, W., & Watanabe, O. (2018). Discrepancies in hospital financial information: Comparison of financial data in state data repositories and the Healthcare Cost Reporting Information System. *Journal of Information Systems*, 33(3), 19–44. <https://doi.org/10.2308/isys-52149> (authors found “widespread,” “material discrepancies” between cost reports and hospital financial statements, along with “many avoidable computational errors” in cost reports); and Locey, K., & Stein, B. (2023). Democratizing insights into hospital cost reports. *Healthcare Analytics*, 4, 100274. <https://doi.org/10.1016/j.health.2023.100274> (“wage index data [often] contained extreme values and erratic year-to-year changes, [suggesting] potential reporting errors,” amounts of charity care “were often suspect” and sometimes flatly impossible). Our own analysis of cost report data suggested similar data quality concerns.

The Composite Quality Score (CQS) adjustment and stop-loss/gain limits are applied at the episode initiator and performance period level, after summing the reconciliation amounts across clinical episodes. To evaluate financial performance for medical and surgical episodes separately, we adjusted the clinical episode–level data by multiplying each episode initiator’s clinical episode–level reconciliation amounts by the ratio of that episode initiator’s total reconciliation after the CQS adjustment and stop-loss/gain limits for that performance period over the episode initiator’s total reconciliation before the CQS adjustment and stop-loss/gain limits. This is the same way that we adjusted reconciliation amounts to compute Medicare program savings.

We then added the average reconciliation payment across all episodes and years to the centered reconciliation amounts and took the average centered per-episode reconciliation for each hospital, weighting by episode count. We broke the benchmark prices into deciles and calculated descriptive statistics for the hospitals with average per-episode reconciliation payments within each decile.

Limitations

Although we used the latest reconciliation data available, we could not include data for the last performance period of the model, which included the end of Model Year 8 (episodes ending in June through October 2025, Performance Period 14). We also used preliminary data for the second half of Model Year 7 (episodes ending June through December 2024, Performance Period 12) and the first half of Model Year 8 (episodes ending January through June 2025, Performance Period 13). CQS amounts were not calculated during the first round of reconciliation in a given model year. Instead, all providers were given a CQS value of zero, and their reconciliation payments were adjusted accordingly; in subsequent rounds of reconciliation, these amounts were adjusted using the true CQS value. Consequently, for Model Year 8 episodes in Performance Period 13, our reconciliation amounts reflected these preliminary CQS values of zero.

Our reconciliation data reflect the amounts paid to or owed by providers. However, we did not have access to financial transaction data from CMS and could not confirm that all amounts for repayments owed to CMS were actually paid.

d. Patient Characteristics

We calculated the unadjusted mean values for the patient characteristics in the evaluation baseline period (episodes with anchor stays or procedures beginning on January 1, 2015, or later and ending by September 30, 2018), Model Years 4 through 6, and Model Years 7 and 8 for BPCI Advanced episodes and comparison group episodes. We then calculated the relative change as the difference in the mean value of the characteristics for BPCI Advanced episodes in the model years compared with the baseline period, relative to the change for the comparison group (**Appendix G**).

These means and the relative changes are computed from Medicare claims data for the key patient characteristic variables listed in Exhibit C.11. For Model Years 4 through 6, we included all variables listed. In Model Years 7 and 8, the data did not include prior home health use and

Hierarchical Condition Category (HCC) score, and therefore, the means and relative changes are not available for these two variables in **Appendix G**. The sample was restricted to episodes where patients had (1) complete FFS enrollment history 6 months prior to the anchor stay or procedure and (2) nonmissing age and sex data.

Exhibit C.11: Definition of Characteristics and Patient Mix

Variable	Definition
Age 80+ Years	Percentage of episodes where the patient was 80 years or older as reported in the MBSF.
Male	Percentage of episodes where the patient sex was male as reported in the MBSF.
Count of HCC Indicators	Average number of HCC comorbidity indicators per patient from diagnostic history during the 6 months prior to the episode start date.
Disabled, No End-Stage Renal Disease (ESRD)	Percentage of episodes where the patient had a disability but did not have ESRD as reported in the MBSF.
HCC Score	Average HCC score. The HCC score (or index) was constructed using patient demographics and diagnostic history. Each episode was assigned an HCC score based on the patient's diagnosis information during the 6 months prior to the episode start date, using v22 of CMS' 2019 Risk Score software, and 2016 (ICD-9) and 2019 (ICD-10) diagnosis-to-chronic condition mappings. For example, the HCC score for an episode that started on July 1, 2022, was constructed using diagnoses from January 1, 2022, to June 30, 2022, claims.
Dual Eligibility	Percentage of episodes where the patient was dually eligible for Medicare and Medicaid during the 180 days prior to the beginning of the episode as reported in the MBSF.
Prior Home Health Use	Percentage of episodes where the patient accessed home health services in the 180 days prior to the beginning of the episode.
Prior Institutional Post-Acute Care (PAC) Use	Percentage of episodes where the patient accessed institutional PAC services in the 180 days prior to the beginning of the episode.

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

2. Impact of BPCI Advanced on Claims-Based Outcomes in Model Years 4–6

a. Outcomes

We used data from Medicare claims to create payment, utilization of health care services, and quality outcomes. All outcomes were created for episodes using a “first admission” episode definition following the BPCI Advanced Model specifications for Model Years 4 through 8. Specifically, when a patient had two eligible discharges or procedures within a 90-day period, we retained only the *first* one as an episode.²⁶ Exhibit C.12 provides details about each outcome

²⁶ There were three exceptions to these rules. First, in the case of multiple overlapping MJRLE episodes, the subsequent episode was retained. Second, in cases where two episodes began on the same day, which is only possible when one was inpatient and one was outpatient, the model specified that the inpatient episode was retained as the initial clinical episode. Third, in cases of a percutaneous coronary intervention (PCI) episode overlapping with a following endovascular cardiac valve replacement episode, the subsequent endovascular cardiac valve replacement episode was retained and the preceding PCI episode was excluded.

measure used in our impact analyses. In addition to the restrictions listed under the “Additional Restrictions for Sample Eligibility” column, we restricted all outcomes to beneficiaries who:

- Maintained FFS Part A and B enrollment through the measurement period (or until death, in the case of the mortality outcome)
- Survived the anchor hospitalization stay or procedure (except for all-cause mortality during the anchor stay or 90-day post-discharge period)
- Had an episode with a discharge date between January 1, 2016, through September 30, 2018, or January 1, 2021, through December 31, 2023
- Had a complete Medicare FFS Part A and Part B enrollment history 6 months prior to the anchor stay or procedure
- Had nonmissing age and sex and reported age less than 115 years at the time of the episode

Exhibit C.12: Claims-Based Outcome Definitions for Model Year 4–6 (2021–2023) Impact Analyses

Domain	Outcome Name	Description	Technical Definition	Additional Restrictions for Sample Eligibility
Payment	Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day	Average total Medicare Part A & B standardized allowed amount during the episode.	The sum of nonnegative Medicare Part A & B payments, including patient out-of-pocket amounts for all health care services during the anchor stay or outpatient procedure + 90-day PDP Standardized Medicare dollars.	Patients who had positive payments for the following three categories: anchor hospitalization, total Part A, and total Part B.
	Total Medicare Part A & B Standardized Allowed Payment Amount During the Anchor Stay or Procedure	Average total Medicare Part A & B standardized allowed amount during the anchor stay or procedure.	The sum of Medicare Part A & B payments, including patient out-of-pocket amounts for all health care services during the anchor stay or outpatient procedure. Standardized Medicare dollars.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Total Medicare Part A & B Standardized Paid Amount 90-Day	Average total Part A & B standardized amount paid by Medicare, during the episode.	The sum of Medicare payments for all health care services during the anchor stay or outpatient procedure + 90-day PDP. Does not include patient cost sharing. Standardized Medicare dollars.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Medicare Part A SNF Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for SNF setting.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided by a SNF during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Medicare Part A IRF Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for IRF setting.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided by an IRF during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Medicare Part A HH Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for HH setting.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for HH during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Medicare Part A Readmissions Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for all-cause readmissions.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for readmissions during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
Utilization	Discharged to Institutional Post-Acute Care Setting	The proportion of episodes discharged from the hospital to an institutional PAC setting.	Binary outcome where the first PAC setting was SNF, LTCH, or IRF. Institutional PAC stays are only counted as a first PAC setting if the patient was admitted to SNF, LTCH, or IRF within 5 days of discharge from the hospital.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.

Domain	Outcome Name	Description	Technical Definition	Additional Restrictions for Sample Eligibility
Utilization (Continued)	Number of Days in a SNF	Average number of SNF days.	The total number of days of SNF care (not necessarily consecutive) during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day with the additional requirement that patients had at least 1 SNF day during the 90-day PDP.
Quality	Readmission Rate 90-Day	The proportion of episodes with one or more readmissions for any condition 90 days after the anchor stay or outpatient procedure.	Binary outcome (1 = at least one eligible readmission during measurement period; 0 = no eligible readmission during measurement period). Eligible readmissions are IPPS claims with any MS-DRG, excluding those with a diagnosis of COVID-19 (indicated with ICD-10 code U071).	Patients who were discharged from the anchor stay or outpatient procedure in accordance with medical advice.
	Unplanned Readmission Rate 90-Day ^a	The proportion of episodes with one or more unplanned, all-cause readmissions for any condition 90 days after the anchor stay or outpatient procedure.	Binary outcome (1 = at least one eligible readmission during measurement period; 0 = no eligible readmission during measurement period). Eligible readmissions are IPPS claims not on the list of planned admissions as defined by National Quality Forum measure 1789 and excluding those with a diagnosis of COVID-19 (indicated with ICD-10 code U071).	Same as Readmission Rate 90-Day.
	All-Cause Mortality During the Anchor Stay or 90-Day PDP	The proportion of episodes with death from any cause during the anchor stay or 90-day PDP.	Binary outcome (1 = the patient date of death occurs during the anchor stay or procedure or within 90 days following the anchor stay or procedure end date; 0 if no date of death by the patient during the anchor stay or procedure or within 90 days following the anchor stay or procedure end date).	Patients who had not received hospice care in the 6 months prior to their anchor stay or procedure. The eligible sample includes episodes where the patient died during the anchor stay.
	All-Cause Mortality During the 90-Day PDP	The proportion of episodes with death from any cause during the 90-day PDP.	Binary outcome (1 = the patient date of death is within 90 days of the anchor stay or procedure end date; 0 if no date of death by the patient within 90 days of the anchor stay or procedure end date).	Patients who had not received hospice care in the 6 months prior to admission.
	ED Visit 90-Day PDP	The proportion of episodes with an ED visit during the 90-day PDP.	Binary outcome (1 = the patient has any ED visit during the measurement period for which the patient required medical treatment but was not admitted to the hospital.; 0 otherwise). Eligible ED visits are outpatient claims with a code indicating the patient used the ED but was not admitted to the hospital.	Same as Readmission Rate 90-Day.

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

^a The unplanned readmission rate outcome is based on specifications for the National Quality Forum (NQF) all-cause unplanned readmission measure (NQF measure 1789), whereas planned admissions are based on AHRQ Clinical Classification System procedure and diagnosis codes.

*b. Study Samples***BPCI Advanced Study Population**

We used the same study sample of treated and comparison hospitals and PGPs for Model Year 4 through 6 pooled analyses as we had previously used for Model Year 4 analyses. We analyzed hospitals and PGPs that participated in BPCI Advanced and met certain evaluation exclusions (including sufficient evaluation baseline volume, nonmissing data on matching variables, and, for PGPs, a sufficient share of the PGP's volume at BPCI Advanced–eligible hospitals). We analyzed only clinical episodes with a sufficient count of participating hospitals or PGPs;²⁷ the clinical episodes that we evaluated included 97.3% of the Model Year 4 through 6 volume. For details of our Model Year 4 BPCI Advanced study population, see the BPCI Advanced Evaluation Fifth Evaluation Report.²⁸

*c. Analytic Methods***Comparison Group Selection**

The DiD approach compares the change in episode outcomes for hospitals and PGPs in BPCI Advanced with those of a group of comparable hospitals and PGPs not in the model. This estimation strategy relies on the comparison group serving as a counterfactual of the change in outcomes in the absence of the model. For example, in Exhibits 8 and 9 in the Final Evaluation Report, the dotted arrows express the estimated difference in post-acute care (PAC) use in the absence of the effects of BPCI Advanced, as estimated by our comparison group.

In Model Year 4, we created a comparison group for each clinical episode and separately for hospitals and PGPs, allowing us to infer outcomes for BPCI Advanced participants had the model not existed. We did this by matching each Model Year 4 BPCI Advanced hospital or PGP, for each clinical episode, to a single other hospital or PGP not participating in that clinical episode. We used 1-1 propensity-score matching with calipers and applied eligibility and contamination restrictions for BPCI Advanced and comparison hospitals and PGPs. For more information on how these comparison groups were selected (including which variables we used in matching), see the BPCI Advanced Evaluation Fifth Evaluation Report.²⁸ Because no new participants were allowed to join in Model Years 5 and 6, we reused these Model Year 4 comparison groups for the pooled Model Year 4 through 6 analysis.

To deal with provider attrition during our study period and ensure balance throughout, we dropped episodes initiated by BPCI Advanced participants after they left the model and episodes initiated by comparison providers after their matched BPCI Advanced provider left the model. For example, if hospital A was matched to hospital B, and hospital A left the model at the end of Model Year 4, we would omit all volume at hospitals A and B after Model Year 4.

²⁷ For Model Years 4 through 6, we did not evaluate impacts for the following clinical episodes: back and neck except spinal fusion (inpatient), bariatric surgery, cardiac defibrillator (inpatient), cardiac valve, double joint replacement of the lower extremity, endovascular cardiac valve replacement, and inflammatory bowel disease.

²⁸ The BPCI Advanced Fifth Evaluation Report is available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

To improve comparison group balance given attrition, we re-executed the matching process using a phased approach that prioritized finding an optimal match for participants that stayed to Model Year 6, then for participants that stayed to Model Year 5, and finally for participants that exited in or after Model Year 4. We did this separately for each clinical episode and separately for hospitals and PGPs, using the Model Year 4 selected comparison group and matching characteristics. We applied volume stratification and used Mahalanobis and scaled Euclidean distance matching for hospitals and PGPs, respectively. When this did not result in improved balance, we retained the original match. We also retained the original match for clinical episodes with fewer than five PGP participants remaining in Model Year 6. After finalizing the match, we excluded comparison hospitals and PGPs in model years where their matched BPCI Advanced provider had no episode volume. As a whole, this phased prioritized matching approach resulted in the same comparison group used in Model Year 4 but with different 1-1 matches. For comparison group balance statistics, see **Appendix E**.

Difference-in-Differences Approach

The DiD approach is a statistical technique that quantifies the impact of an intervention or policy by comparing changes in a *treatment* group (BPCI Advanced) with changes in a comparison group between baseline (preintervention) and intervention periods. We used this analytic technique to estimate the impact of BPCI Advanced on the key claims-based patient outcomes while adjusting for differences in episode, patient, and provider characteristics between BPCI Advanced and comparison episodes. This approach mitigates biases from time-invariant differences between BPCI Advanced and comparison episodes and controls for trends that are common between the BPCI Advanced and comparison populations. Using episodes rather than hospitals and PGPs as observations allowed us to directly control for changes in the composition of the patient mix, which may be a response to the model.

We applied the DiD technique at the episode level, including episodes that occurred during the evaluation baseline period (January 1, 2015–September 30, 2018) and the Model Year 4 through 6 intervention period (January 1, 2021–December 31, 2023). We excluded episodes from our DiD sample that occurred on or after the BPCI Advanced hospital or PGP withdrawal date. In the same fashion, we excluded any episodes from a matched comparison provider that occurred on or after the withdrawal date of its matched BPCI Advanced provider.

To illustrate our estimation strategy, consider the stylized equation,

$$(1) \quad Y_{ikt} = \beta_0 + \beta_1 BPCIA_k + \beta_2 Post_t + \theta(BPCIA_k \cdot Post_t) + X_{ikt}' \beta + e_{ikt}$$

where Y_{ikt} is the outcome of interest for episode i from provider k during time t . The variable $BPCIA_k$ is an indicator that takes on the value 1 if provider k participated in BPCI Advanced for the given clinical episode. $Post_t$ is an indicator that takes the value of 1 for every episode in the intervention period, and X_{ikt} is the set of covariates at the episode, patient, provider, and temporal level for episode i with an anchor stay or procedure ending in period t receiving care from provider k . In this linear example, the DiD estimate is the coefficient θ , which determines the differential in outcome Y experienced by patients receiving services from BPCI Advanced providers during the

intervention period relative to patients receiving services from providers in the comparison group. Lastly, the error term is e_{ikt} .

We used multivariable regression models to control for differences in patient demographics, clinical characteristics, and health care use before hospitalization, along with provider characteristics that might be correlated with the outcome. We used ordinary least squares models for continuous and count outcomes (payments, skilled nursing facility (SNF) days) and logistic regression models for the binary outcomes (institutional PAC use, mortality rate, readmission rate, emergency department (ED) use). To calculate the impact of BPCI Advanced on our binary outcomes using logistic regression models, we took the difference between the average marginal effect of the intervention indicator for BPCI Advanced providers and the average marginal effect of the intervention indicator for comparison providers.²⁹ In all specifications, we clustered standard errors at the provider level (hospital or PGP).

Section D.4 of this appendix discusses how we used the combined DiD methodology in the calculation of Medicare program savings.

Risk adjustment. We included the following covariates in regression models to adjust our impact estimates for observable differences:

- MS-DRG or HCPCS code of the anchor stay or procedure (cross-walked to the Fiscal Year 2024 [FY24] definitions)
- Age, sex, dual eligibility, disability (as original Medicare eligibility reason, excluding end-stage renal disease), additional patient characteristics, and HCC score, all interacted with clinical episode of the episode
- Binary indicators for care in a SNF, home health care, IPPS admission, visit to ED, and no stay or admission in the 6 months preceding the start of the episode
- Confirmed COVID-19 diagnosis during the anchor stay or procedure
- Confirmed COVID-19 diagnosis during the 90 days prior to the anchor stay or procedure
- Dementia diagnosis within a 3-year lookback period
- Alignment of episode to BPCI Initiative, interacted with year
- ACO alignment during any point of the episode of care
- Bed count of the hospital (categorical; fewer than 100, 100–250, or more than 250)
- Indicators for quarter of the year

²⁹ Note that this approach implicitly assumes parallel trends on a *probability* scale in the absence of the model. This is why we perform parallel trends tests using a linear probability model (see [Parallel Trends Tests](#)). There has been some debate in the literature concerning on which scale the difference of differences should be taken; see for instance Barkowski, S. (2021). Interpretation of nonlinear difference-in-differences: The role of the parallel trends assumption. Available from the Social Science Research Network: <https://ssrn.com/abstract=3772458> or <http://dx.doi.org/10.2139/ssrn.3772458>.

- Census region of the hospital
- For surgical episodes, indicators for total knee arthroplasties and CJR episodes

We selected additional covariates for each outcome using a Least Absolute Shrinkage and Selection Operator (LASSO) procedure for each clinical episode and combined groupings (Exhibit C.13). For details of this approach, see **Appendix C** of the BPCI Advanced Fifth Evaluation Report.²⁸

Exhibit C.13: Risk-Adjustment Variables, and Required Risk-Adjusters, by Grouping and Outcome

Outcome	Combined	Medical	Surgical
Payments	• Age squared • Hypertension • Prior care in hospice, LTCH, IRF, and inpatient psychiatric settings • For-profit ownership of hospital • Hospital in urban area • Resident-to-bed ratio of hospital • 17 selected HCC flags	• Prior care in hospice, LTCH, IRF, and inpatient psychiatric settings • For-profit ownership of hospital • Hospital in urban area • 11 selected HCC flags	• Age squared • Hypertension • Prior care in an IRF setting • Resident-to-bed ratio of hospital • 7 selected HCC flags
First Discharge to Institutional PAC Setting	• Hypertension • Prior care in hospice, LTCH, or IRF setting • Episode initiated at a safety-net hospital • Resident-to-bed ratio of hospital • 38 selected HCC flags	• Prior care in hospice, LTCH, or IRF setting • Resident-to-bed ratio of hospital • 23 selected HCC flags	• Hypertension • Prior care in IRF setting • Episode initiated at a safety-net hospital • 27 selected HCC flags
SNF Days	• Age squared • Hypertension • Prior care in IRF or inpatient psychiatric setting • Hospital in an urban area • Resident-to-bed ratio of hospital • Safety-net status of hospital • 44 selected HCC flags	• Prior care in IRF or inpatient psychiatric setting • Hospital in an urban area • Resident-to-bed ratio of hospital • Safety-net status of hospital • 25 selected HCC flags	• Age squared • Hypertension • Prior care in IRF setting • Resident-to-bed ratio of hospital • Safety-net status of hospital • 36 selected HCC flags
ED Use	• Hypertension • Prior care in inpatient psychiatric setting • Hospital in an urban area • Six selected HCC flags	• Hypertension • Three selected HCC flags	• Hypertension • Prior care in inpatient psychiatric setting • Hospital in an urban area • Six selected HCC flags

Outcome	Combined	Medical	Surgical
Readmissions	- Hypertension - Prior care in hospice or inpatient psychiatric setting - For-profit ownership of hospital 6 selected HCC flags	- Hypertension - Prior care in hospice or inpatient psychiatric setting - For-profit ownership of hospital 6 selected HCC flags	None
Mortality	- Hypertension - Prior care in inpatient psychiatric setting - Hospital in urban area 16 selected HCC flags	- Hypertension - Prior care in inpatient psychiatric setting - Hospital in urban area 16 selected HCC flags	None

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

Parallel trends tests. A fundamental assumption for the validity of our DiD design is that the BPCI Advanced group would have followed the same trend in outcomes as the comparison group if the model had never existed, often referred to as the parallel trends assumption. Although this assumption is untestable in practice, we can compare the BPCI Advanced and comparison group trends during the evaluation baseline period. Evidence of nonparallel trends during the baseline suggests that any estimated impacts (or lack thereof) could be explained by the preexisting differences in the trends and not an actual impact of the model. We tested the null hypothesis that BPCI Advanced participants and comparison providers had parallel trends during the baseline at the 10% statistical significance level. To do so, we ran a regression of the outcome on a linear time and BPCI Advanced indicator interaction term, in addition to the full set of risk-adjusters that are included in the DiD specification, on baseline data for each DiD model. If there was no difference between the trends of the BPCI Advanced and comparison groups prior to intervention, the interaction coefficient would be near zero and not statistically significant. For binary outcomes (for example, measuring impacts on utilization), we tested parallel trends using a linear probability model framework, allowing us to perform inference on the parallel trend parameter the same way we perform parallel trends inference on continuous outcomes.

We note outcomes for which we rejected the null hypothesis that there were parallel trends in the baseline. Results of the parallel trends tests are reported with results of the relevant DiD analyses in **Appendices H** and **K**.

Sensitivity analyses. To test the robustness of our impact estimates, we performed the three sensitivity analyses described below. Results of these sensitivity analyses are available in **Appendices H** and **K**.

First, we report the results of a sensitivity analysis in which we removed BPCI Initiative episodes. Baseline episodes involving BPCI Advanced hospitals and PGPs were more likely than other episodes to be aligned to the BPCI Initiative, which ended before our intervention period. Because the BPCI Initiative involved incentives similar to the BPCI Advanced Model, we were concerned that including BPCI Initiative episodes in our baseline period might contaminate our impact estimates. To address this concern, we report DiD impact estimates that exclude these episodes.

Second, we report an intention-to-treat (ITT) sensitivity analysis. Participants that were not successful could voluntarily exit the model, leaving a selected group of remaining participants. Our initial comparison group construction for Model Year 4 was done without knowledge of when providers would exit. To address this concern, we report an ITT impact estimate, performing a DiD analysis that included all episodes from Model Year 4 BPCI Advanced and comparison group hospitals and PGPs, even after they left the model.

Third, we report a sensitivity analysis that used a comparison group of all eligible, non-BPCI Advanced Medicare FFS episodes nationwide in place of our selected comparison group. This sensitivity analysis assessed whether our results were robust to our comparison group selection methodology. We used the same DiD framework with the addition of quarter fixed effects in the regressions to account for increasing divergence in episode volume as BPCI Advanced participants left the model and episode costs increased over time.

d. Limitations

We estimated the model's impact using a DiD design with a comparison group constructed via propensity-score matching. A DiD design depends on the comparison group validly reflecting what would have occurred absent the model (the counterfactual). We selected a matched comparison group that was similar during the evaluation baseline period to the BPCI Advanced hospitals and PGPs on key factors expected to influence payment, utilization, and quality outcomes; however, important factors may exist for which we did not account. We extensively assessed the quality of our matched comparison group, including testing our results against a national comparison sample (**Appendix H**) and believe our methodology appropriately addressed voluntary selection into participating in the model. In addition, the model allowed participants to exit at any time, which likely resulted in voluntary selection out of the model. Our main findings are similar to our ITT sensitivity analysis, suggesting voluntary exits are not driving the impacts we find. However, with any voluntary policy, and without a randomized controlled trial, we cannot be certain our comparison group represents the true counterfactual absent the model.

Our estimates may be influenced by the imbalance of certain characteristics resulting from difficulty finding similar hospitals and PGPs from a limited, eligible pool of comparison providers, although the effects from these imbalances are partially mitigated through risk adjustment. See **Appendix E** for the standardized mean differences of variables before and after matching.

We did not evaluate all BPCI Advanced Model Year 4 through 6 episodes for various reasons. First, in our comparison group construction, we had to exclude a small subset of BPCI Advanced episode initiators to achieve balance across important baseline characteristics. Second, we excluded participating PGPs without episodes in the baseline period from our evaluation to

improve comparability of our baseline and intervention samples.³⁰ Third, we did not evaluate some clinical episodes due to their limited sample size and challenges in identifying a suitable matched comparison group. We assessed the impact of BPCI Advanced for a large subset of the 34 clinical episodes: 26 for hospital episode initiators and 18 for PGP episode initiators. The sample of Model Year 4 through 6 episodes that we included in our impact estimates represents the substantial majority of the total BPCI Advanced Model Year 4 through 6 episodes.

3. Impact of BPCI Advanced on Claims-Based Outcomes in Model Years 7–8

a. Outcomes

We used data provided by the BPCI Advanced IMC, which were derived from Medicare claims, to create payment, utilization of health care services, and quality outcomes for Model Years 7 and 8. We created all outcomes for episodes using the BPCI Advanced Model episode definition (first-admission episodes). Exhibit C.14 provides details about each outcome measure used in our impact analyses. Note that this more limited set of outcomes in Model Years 7 and 8 corresponds to the outcomes with the same names in Model Years 4 through 6, but there may be minor differences in definitions due to the different data source.

³⁰ PGP episode initiators could form new TINs specifically to participate in BPCI Advanced. As a result, there were no baseline claims data to use for matching purposes for some PGP episode initiators. To preserve as many PGPs as we could in our sample, we linked new PGP episode initiators to baseline data when possible using data from Medicare Part B claims, Medicare Data on Provider Practice and Specialty, and the BPCI Advanced Database. We performed these linkages where there was sufficient overlap of NPIs and other practice characteristics (74 PGPs across Model Years 1–6). For cases where we could not identify related TINs in the baseline, we removed the PGP episode initiator from the sample (120 TINs across Model Years 1–6).

Exhibit C.14: Claims-Based Outcome Definitions for Model Year 7–8 (2024–2025) Impact Analyses

Domain	Outcome Name	Description	Technical Definition	Eligible Sample
Payment	Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day	Average total Medicare Part A & B standardized allowed amount during the episode.	The sum of Medicare Part A & B payments, including patient out-of-pocket amounts for all health care services during the anchor stay or outpatient procedure and 90-day PDP. Standardized Medicare dollars.	Patients who had nonzero payments for the following three categories: anchor hospitalization, total Part A, and total Part B.
	Medicare Part A SNF Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for SNF setting.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided by a SNF during the 90-day PDP.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
	Medicare Part A HH Standardized Allowed Amount 90-Day	Average Medicare Part A standardized allowed amount, for HH setting.	The sum of Medicare payment and patient out-of-pocket amounts for Part A health care services provided for HH during the 90-day PDP. Differs slightly from the HH payment measure used for Model Years 4–6 due to different treatment of requests for anticipated payment.	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.
Quality	All-Cause Mortality During the 90-Day PDP	The proportion of episodes with death from any cause during the 90-day PDP.	Binary outcome (1 = the patient date of death is within 90 days of the anchor stay or procedure end date; 0 if no date of death by the patient within 90 days of the anchor stay or procedure end date).	Same as Total Medicare Part A & B Standardized Allowed Payment Amount 90-Day.

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

b. Study Samples

In this section, we define the sample used in our evaluation of the BPCI Advanced Model. We made decisions to determine various aspects of our sample, including identifying the BPCI Advanced group and national comparison sample and defining the episode of care.

BPCI Advanced Study Population

We defined the BPCI Advanced hospital or PGP *treatment* group as hospital or PGP episode initiators that participated in at least one clinical episode in Model Year 7. Refer to **Appendix B** for an overview of clinical episodes, their associated service line groups, and designation of clinical episodes used for measuring hospital and PGP impacts.

We then implemented baseline and intervention volume restrictions by removing all volume from participating PGPs and from participating hospitals that lacked volume in the intervention or our evaluation baseline in a given clinical episode. For example, if a PGP had episode volume in Model Year 7 but none in our baseline period, then its Model Year 7 episodes were excluded.

These exclusions were intended to strengthen our DiD identification strategy by improving comparability of our baseline and intervention samples.

National Comparison Sample

The IMC data, generated as part of implementation activities, had a more limited set of variables than the custom analytic files we developed for the evaluation of prior model years but allowed us to more efficiently build the final analytic file for episodes in Model Years 7 and 8. Given data limitations and for efficiency, we relied on a national rather than a matched comparison group for the Model Year 7 and 8 analyses. The national comparison group included all eligible Medicare FFS episodes occurring at nonparticipating hospitals, subject to certain exclusions. We refer to this as a “national comparison sample,” and it is similar to the methods we used in the Model Year 4 through 6 national comparison sample sensitivity analyses described above and in past evaluation reports.

We applied the following exclusions to our comparison sample:

- We restricted the sample to episodes eligible for the model (meaning they involved eligible MS-DRGs or HCPCS codes, occurred at eligible hospitals, and were not subject to other model or beneficiary exclusions).
- We excluded episodes at hospitals that had previously participated in BPCI Advanced in a given clinical episode. For instance, congestive heart failure volume at a hospital that previously participated in BPCI Advanced for congestive heart failure was excluded, but congestive heart failure volume at a hospital that previously participated in MJRLE was included. We also excluded episodes at all hospitals that participated in Model Year 7, regardless of the CESLGs in which they participated. For example, if a hospital participated only in *cardiac care* in Model Year 7, none of its episodes (in *cardiac care* or some other CESLG) would be included in our comparison group. Its episodes in *cardiac care* could be included in our treated group if they satisfied our inclusion criteria.
- We removed all episodes involving all PGPs that participated in Model Year 7 or had previously participated in the model in the same clinical episode. When implementing this restriction, we used historical data to aggregate PGPs that had split up to create new TINs when they started participating in the BPCI Advanced Model so that we could include their episodes in the evaluation.
- We excluded episodes at any hospital that lacked volume in our intervention period or evaluation baseline period, in a given clinical episode. For example, we would exclude a congestive heart failure intervention episode at a hospital that had no congestive heart failure episodes from January 1, 2016, to October 1, 2018.

We used our Model Year 7 and 8 methods to estimate Model Year 4 through 6 impacts and compared the results with those of our main Model Year 4 through 6 impact analyses (Exhibit C.15). This analysis allowed us to assess the magnitude of bias introduced by relying on IMC data and a national comparison group in Model Years 7 and 8.

Exhibit C.15: Assessment of Difference in Model Year 7–8 (2024–2025) Impact Estimates Due to Data Source

Episode Type		DiD Estimates, Total Payments, for Different Model Years, Data Sources, and Comparison				Difference Due to National Comparison Group (%)	Difference Due to National Comparison Group and Different Data Source (%)
		MY7–8 IMC Data National Comparison (\$)	MY4–6 Research File Matched Comparison (\$)	MY4–6 Research File National Comparison (\$)	MY4–6 IMC Data National Comparison (\$)		
All Participants	<i>Combined</i>	-1,600	-1,011	Not estimated	-1,149	Not estimated	13.6
	<i>Medical</i>	-822	-756	-705	-844	-6.7	11.6
	<i>Surgical</i>	-2,814	-1,835	-2,232	-2,169	21.6	18.2
Hospital	<i>Medical</i>	-729	-744	Not estimated	-876	Not estimated	17.7
	<i>Surgical</i>	-2,216	-1,672	Not estimated	-1,608	Not estimated	-3.8

Note: See Exhibit C.1 for data sources and **Appendix A** for definitions and acronyms.

c. Analytic Methods

National Sample Balancing Weights

We applied weights to our Model Year 7 and 8 DiD estimates to improve the balance of the national comparison group. We assigned all BPCI Advanced episodes a weight of one, while we assigned comparison episodes weights based on their CESLG and the quarter of the anchor stay discharge or procedure, which helped balance volume temporally and across service lines.

Absent these balancing weights, imbalances of clinical episode volume could have biased our estimates. For example, clinical episodes in the *orthopedics* CESLG were overrepresented among BPCI Advanced episodes relative to our national comparison sample due to high participation in this CESLG, and *cardiac procedures* clinical episodes were underrepresented due to low participation in this CESLG. This would violate our parallel trends assumption because expenditures for these service lines were on different trajectories during our study period independent of the model, resulting in different counterfactual trends in average episode expenditures at an aggregate model level.

Additionally, the national sample balancing weights provide more consistent units of participation across time. Participants left the model voluntarily during the intervention period, and we exclude their episodes after they exit. The temporal aspects of the balancing weights help mirror the decreases in volume caused by these voluntary exits in the national comparison group.

Difference-in-Differences Approach

To study impacts of the model in Model Years 7 and 8, we used the same DiD approach that we used for Model Years 4 through 6 with modifications due to data limitations. Because we did not have a PGP comparison group, we clustered our standard errors at the hospital level (rather than the provider level, as with our Model Year 4 through 6 analyses). We used a shortened baseline

period (January 1, 2016–September 30, 2018, instead of January 1, 2015–September 30, 2018). We also used balancing weights for the national sample, as described above, and slightly different risk-adjustment covariates, as described below.

Risk adjustment. We used the same risk-adjustment covariates for our Model Year 7 and 8 impacts as we did for our Model Year 4 through 6 impacts, with the following modifications:

- We included only two prior use covariates only: prior care in IPPS hospital settings and prior care in PAC settings.
- We excluded the following covariates because data were unavailable:
 - HCC score
 - Dementia diagnosis
 - COVID-19 diagnosis

Parallel trends tests. With Model Year 4 through 6 impacts, we conducted parallel trends tests to check the parallel trends assumption of DiD estimates. We used the same methods as for the Model Year 4 through 6 analyses, except that regressions were weighted using the CESLG national sample balancing weights from the DiD analyses. Results of these parallel trends tests can be found alongside the corresponding DiD estimates in **Appendices H and K**.

Sensitivity analyses. Because our impact estimation approach for Model Years 7 and 8 differed substantially from the approach used for prior model years, we performed three sensitivity analyses to assess the effect of methodological changes on our estimates.

First, to assess the robustness of our risk adjustment, we ran a sensitivity check that added all individual HCC comorbidity indicators and available covariates in the IMC data as additional risk-adjusters (referred to as *additional risk adjustment*).

Second, because we could not account for provider-level selection into the model as we have in previous years, we ran a sensitivity analysis that used *alternative national sample balancing weights*. The analysis weighted episodes with important covariates in the matching models: bed count of the hospital and presence of an IRF in the market. For details of these weights and balance of the resulting weighted sample, see **Appendix E**.

Finally, we ran a sensitivity analysis with stricter contamination restrictions to remove comparison group providers from the national comparison group that were past BPCI Advanced participants (referred to as *excluding past participants from the comparison group*). Specifically, we also removed all hospitals and PGPs that had ever participated in BPCI Advanced, regardless of when or in which clinical episodes they participated.

d. Limitations

Our Model Year 7 and 8 impacts have most of the same limitations as our Model Year 4 through 6 impacts. First, they depend on an assumption that the comparison episodes represent a valid

counterfactual. This assumption is more likely to be violated for our Model Year 7 and 8 estimates because we could not select a matched comparison group. To estimate the likely bias introduced by using a national comparison group, we estimated the model's impact in Model Years 4 through 6 using the IMC data and our Model Year 7 and 8 methods (including the balancing weights and the national comparison group) (Exhibit C.15). We then compared these results with the results of our main Model Year 4 through 6 analysis using our matched comparison group.

Several additional limitations and operational challenges arose from relying on IMC data for the evaluation. First, exclusion criteria for BPCI Advanced episodes had been applied prior to receipt of the data. As a result, episodes typically included in the evaluation, such as those associated with patients attributed to an ACO, episodes in which the patient died during the anchor stay or procedure, and episodes overlapping with participation in the CJR Model, had already been excluded. Second, the IMC data inputs were from different model years, which had different model design specifications, requiring the application of Model Year 7 and 8 specifications across datasets to ensure consistency in episode construction, a requirement for the evaluation. For example, the lookback period for Medicare enrollment was 90 days in Model Year 4, compared with 180 days for both Model Years 7 and 8. These inconsistencies had to be harmonized as part of our IMC data file construction. Third, MS-DRG definitions change over time, but the evaluation requires consistent definitions. We used the FY24 MS-DRGs to identify triggering episodes, which required mapping episodes in earlier datasets that were coded with FY22 MS-DRGs to FY24 MS-DRGs to maintain uniform classification over time. We had to combine a small number of MS-DRGs in our risk-adjustment variables to achieve this consistent FY24 definition.

Lastly, there is a gap between our baseline and intervention periods, resulting in a long study period: 10 years from the beginning of the baseline (January 1, 2016–September 30, 2018) to the end of the intervention period (January 1, 2024–April 30, 2025). Deviations from the parallel-counterfactual-trend assumption could compound over these 10 years and bias our results. This problem is more pronounced for our Model Year 7 and 8 estimates than it is for our Model Year 4 through 6 estimates, because of the longer gap between the baseline and intervention period.

4. Medicare Program Savings

We followed a four-step process to compute net savings to Medicare from the BPCI Advanced Model. Overall, the approach was consistent for Model Years 4 through 6 and Model Years 7 and 8, with adjustments to address differences in data availability for Model Years 7 and 8.

1. Estimate the per-episode impact of the model in nonstandardized paid dollars using DiD impact estimates converted from standardized paid dollars to nonstandardized paid dollars.
2. Multiply the per-episode impact by the number of episodes impacted by BPCI Advanced.
3. Compute net reconciliation payments made by or to participants.
4. Calculate Medicare program savings as the reduction in nonstandardized paid payments minus net reconciliation payments to participants.

a. Estimate Per-Episode Impact and Convert to Nonstandardized Paid Dollars

We estimated per-episode payment reductions using the DiD impact estimates. Specifically, the per-episode reduction equals the DiD impact estimate multiplied by –1 and then converted from standardized to nonstandardized dollars using an adjustment factor.

For Model Years 4 through 6, as in previous years, we began with the DiD estimate of the model's impact on standardized paid payments, which reflect standardized amounts without patient cost sharing. We converted this estimate to nonstandardized dollars by multiplying it by an adjustment factor, which is the ratio of average nonstandardized paid payments to average standardized paid payments in the BPCI Advanced intervention sample.

For Model Years 7 and 8, this approach was not feasible due to data limitations: The IMC data included allowed amounts but not paid amounts. Accordingly, we multiplied the DiD estimate of the impact on standardized allowed amounts by the ratio of nonstandardized allowed payments to standardized allowed payments for Model Years 7 and 8. We then multiplied this impact by the ratio of nonstandardized paid payments to nonstandardized allowed payments, calculated using 2023 episodes involving Model Year 7 and 8 BPCI Advanced hospitals and PGPs. This conversion relied on supplementing IMC data with our preexisting claims data through 2023.

Because this paid-to-allowed ratio varied over time, we assessed an alternative imputation-based approach. Using historical claims data from 2016 through 2018 and 2023, we estimated a model predicting paid amounts from allowed amounts and applied this model to allowed amounts in the IMC data. We fitted the model on data from 2016 through 2018 and 2023. We then re-estimated DiD impacts using predicted paid payments and used this impact estimate to compute Medicare savings.

b. Multiply by Extrapolation Count

To aggregate total payment reductions, we multiplied the per-episode impact by a count of the BPCI Advanced volume in the model year(s) ("Medicare savings episode count"), representing episodes affected by the model.

For Model Years 4 through 6, the Medicare savings episode count includes all episodes initiated at BPCI Advanced participating hospitals and PGPs during their periods of participation between January 1, 2021, and December 31, 2023. This Medicare savings episode count exceeded the number of episodes in model reconciliation because it included several types of episodes that were excluded from reconciliation but included in the evaluation's DiD analyses.³¹ These include

³¹ For Model Years 4 through 6, the number of episodes reconciled in the model were approximately 67% of the Medicare savings episode count.

episodes with a COVID-19 diagnosis,³² episodes with patients aligned to certain ACO models, episodes initiated by participating PGPs at hospitals with no target price due to lack of volume, and episodes at hospitals subject to major disaster declarations. The Medicare savings episode count was slightly greater than the DiD sample because it included episodes with missing values on covariates, episodes excluded due to applied calipers, and episodes triggered at hospitals that did not meet minimum baseline episode requirements. Some clinical episodes had insufficient volume to evaluate. We excluded these episodes from both the DiD analyses and the Medicare savings episode count.

For Model Years 7 and 8, the Medicare savings episode count was equal to the DiD sample restricted to episodes with anchor end dates on or before April 2, 2025, and closely aligned to the number of episodes in model reconciliation.

c. Compute Total Reconciliation Payments

Net payment reconciliation amounts (NPRA) are payments made by Medicare to BPCI Advanced participants when their episode spending was less than their target prices. Participants whose episode spending exceeded their target prices owed repayment amounts to Medicare. We refer to NPRA and repayment amounts collectively as reconciliation payments.

We used reconciliation data for Model Years 4 through 6 and Model Years 7 and 8 and aggregated across participants and the corresponding clinical episodes we evaluated to compute aggregate net reconciliation amounts. For Model Years 4 through 6 and the first quarter of Model Year 7 (Performance Periods 5–11), we used finalized second true-up amounts. For the second and third quarter of Model Year 7 (Performance Period 12), we used first true-up amounts. For the last quarter of Model Year 7 and the first half of Model Year 8 (Performance Period 13), we used initial reconciliation amounts.

In the model's reconciliation process, reconciliation data were initially calculated at the clinical episode level; they were then summed across clinical episodes for a given provider, after which the CQS adjustment and stop-loss/gain limit were applied. Accordingly, clinical episode-level reconciliation amounts (and thus also medical or surgical reconciliation amounts) do not have these adjustments. To approximate these adjustments at the clinical episode level, we applied an adjustment factor equal to the ratio of total reconciliation after CQS and stop-gain/loss adjustments to total preadjustment reconciliation, calculated at the hospital- or

³² The rules related to excluding or including episodes with a COVID-19 diagnoses changed over this time period. In Model Year 4 (2021), all episodes where the patient was diagnosed with COVID-19 at any time during the clinical episode were excluded from reconciliation. In Model Year 5 (2022), model participants were able to sign an amendment to be held accountable for episodes with a COVID-19 diagnosis, in which case these episodes were included in reconciliation. Beginning in Model Year 6 (2023), all episodes with a COVID-19 diagnosis were included in reconciliation.

PGP-by-performance-period level. We then aggregated adjusted reconciliation payments within each clinical episode and took the sum to compute reconciliation amounts for pooled groupings.³³

d. Compute Medicare Program Savings

We calculated Medicare program savings separately for each grouping as follows:

$$\text{Medicare savings} = \text{reduction in nonstandardized paid payments} - \text{reconciliation payments}$$

Positive estimates represent savings to Medicare and negative amounts represent losses for Medicare. To construct a range for the Medicare program savings estimates, we carried the upper and lower confidence limits of the DiD impact estimates through each calculation step, applying the standardized-to-nonstandardized conversion, multiplying by the extrapolation count, and subtracting reconciliation payments.

Reconciliation payments and Medicare savings episode count sum across groupings (for example, medical and surgical). However, because per-episode payment reductions are estimated from separate regressions with different risk adjustment, combined reductions do not equal the sum of subgroup-specific reductions.

We calculated savings per episode by dividing total savings by the corresponding extrapolation count. To express savings as a percentage of what payments would have been absent the BPCI Advanced Model, we constructed a counterfactual estimate of standardized payments absent BPCI Advanced by adding the change in the risk-adjusted comparison group mean (from baseline to intervention) with the risk-adjusted BPCI Advanced baseline mean. We converted this counterfactual to nonstandardized paid dollars and used it as the denominator.

e. Limitations

Our calculations of BPCI Advanced Medicare program savings are based on the respective DiD estimates, so they inherit the limitations discussed above regarding impacts on total episode payments. In addition, there are a few additional limitations relevant specifically to Medicare savings results.

For Model Years 4 through 6, we extrapolated our estimated reductions in episode payments to BPCI Advanced hospitals and PGPs that were excluded from the evaluation sample because of insufficient baseline volume or because we could not find appropriate matched comparison providers. This extrapolation assumes similar payment reductions for these excluded episodes as we found for episodes included in our evaluation sample. The evaluation sample represents the majority of all BPCI Advanced episodes in the Medicare savings sample (83.6%), and we have no evidence that payment reductions differ between the evaluation sample and the excluded sample to which we extrapolated Medicare savings. Conversely, we estimated net savings for only those clinical episodes for which we had sufficient sample size, including 26 hospital clinical episodes

³³ This adjustment process is relevant for medical and surgical groupings for Model Year 4 through 6 and Model Year 7 and 8. For Model Years 4 through 6, the adjustment process was also relevant for the full-model estimate because we excluded reconciliation payments (and potential cost reductions) for clinical episodes that we did not evaluate.

and 18 PGP clinical episodes, and omitted any cost reductions and reconciliation payments for other clinical episodes. The clinical episodes we included represent the vast majority of BPCI Advanced volume and reconciliation amounts.

For Model Years 7 and 8, we only included episodes in Model Year 8 that ended on or before June 30, 2025 (episodes with anchor end dates on or before April 2, 2025) and hence excluded the last 3 months of Model Year 8. In addition, net savings estimates relied on converting allowed amounts to paid amounts using historical paid-to-allowed ratios. If these ratios changed over time or differed across payment components, estimates could differ from those based directly on paid data.

We report savings estimates in nonstandardized dollars, whereas we estimate impacts on standardized payment data; we convert between the two using the ratio of average nonstandardized to average standardized payments. However, the true ratio of the model impacts on nonstandardized and standardized payments could differ from the ratio of average nonstandardized to average standardized payments. For instance, if the impact of the model on rural hospitals differs from the impact of the model on urban hospitals, this might not hold.

Generally, our estimates capture only a subset of potential savings from the model. They do not account for potential sustained effects after providers exit the model or spillover effects to non-BPCI Advanced episodes.

Appendix D: Patient Survey – Instrument



Health Care Experience Survey

We are interested in the quality of care you received at the hospital listed in the cover letter, and how your recovery has been going. We understand that this was probably a difficult time for you and your family. We appreciate you taking the time to tell us about your health care experiences. Please be assured that all responses are confidential.

There are four sections of this survey. The first section asks about ***how you were feeling just before*** you went to the hospital listed in the cover letter. The second section asks about ***how you are currently feeling***. The third section asks about ***your experience and satisfaction*** with the hospital and any other places where you received care after you left the hospital. The last questions in the survey are about you.

Instructions:

- Please read each question carefully and respond by marking the box next to the response that most closely represents your opinion.
- Please mark only one box for each question, unless it tells you to “Choose all that apply.”
- Many people use a PENCIL in case they want to change their answers. Please erase cleanly or white out any marks you wish to change. Please do NOT use a felt tip pen.
- Please do not make any stray marks on the form.

1. First, please indicate who is completing this survey.

- ☒ Person named in the cover letter
- ☒ Person named in the cover letter, with help from a family member, friend or caregiver
- ☒ A family member, friend, or caregiver of the person named in the cover letter
- ☒ Someone else who is not a family member, friend, or caregiver of the person named in the cover letter
- ☒ **If the person to whom this was mailed cannot complete the survey, and there is no one else who can do it for him or her**, please mark this response and return the blank survey to Abt Associates, P.O. Box 5720, Hopkins, MN 55343 using the postage-paid envelope provided.

Section 1. Before the Hospital

We would like to know how you were doing **before** you went to the hospital listed in the cover letter. Please think about your overall health and all of your medical needs at that time, and not just the reason you went to the hospital listed in your cover letter.

2. Thinking about the week before you went to the hospital, how much help did you need from another person with **bathing, dressing, using the toilet, or eating**?
 - ☐ No help needed from another person
 - ☐ Some help needed from another person
 - ☐ Complete help needed from another person
 - ☐ Don't know/Don't remember
3. Thinking about the week before you went to the hospital, how much help did you need from another person with **planning regular tasks**, such as making a grocery list or remembering to take medication?
 - ☐ No help needed from another person
 - ☐ Some help needed from another person
 - ☐ Complete help needed from another person
 - ☐ Don't know/Don't remember
4. Thinking about the week before you went to the hospital, what best describes your **use of a mobility device** such as a wheelchair, scooter, walker, or cane?
 - ☐ I never used a mobility device
 - ☐ I sometimes used a mobility device
 - ☐ I always used a mobility device
 - ☐ Don't know/Don't remember
5. Thinking about the week before you went to the hospital, what best describes your ability to **walk by yourself** without resting? That is, without the help of another person or the help of a mobility device.
 - ☐ I could walk several blocks by myself without resting or using a mobility device
 - ☐ I could walk one block by myself without resting or using a mobility device
 - ☐ I could walk from one room to another by myself without resting or using a mobility device
 - ☐ I was not able to walk by myself without resting or using a mobility device
 - ☐ Don't know/Don't remember

6. Thinking about the week before you went to the hospital, how much difficulty did you have **walking up or down 12 stairs?**
- ☐ I had no difficulty walking up or down 12 stairs
 - ☐ I had some difficulty walking up or down 12 stairs
 - ☐ I had a lot of difficulty walking up or down 12 stairs
 - ☐ I was not able to walk up or down 12 stairs
 - ☐ Don't know/Don't remember
7. Thinking about the week before you went to the hospital, how often did your **physical health or emotional problems** interfere with your social activities (like visiting friends, relatives, etc.)?
- ☐ All of the time
 - ☐ Most of the time
 - ☐ Some of the time
 - ☐ A little of the time
 - ☐ None of the time
 - ☐ Don't know/Don't remember
8. Thinking about the week before you went to the hospital, how much did pain interfere with your normal activities?
- ☐ All of the time
 - ☐ Most of the time
 - ☐ Some of the time
 - ☐ A little of the time
 - ☐ None of the time
 - ☐ Don't know/Don't remember

Section 2. After the Hospital

It has been a few months since your hospital care and we would like to know how you are doing **today**. Please think about your overall health and all of your medical needs, and not just the reason you went to the hospital listed in your cover letter.

9. How much help do you currently need from another person with **bathing, dressing, using the toilet, or eating?**
- ☐ No help needed from another person
 - ☐ Some help needed from another person
 - ☐ Complete help needed from another person
 - ☐ Don't know/Don't remember

10. How much help do you currently need from another person with **planning regular tasks**, such as making a grocery list or remembering to take medication?
- ☐ No help needed from another person
 - ☐ Some help needed from another person
 - ☐ Complete help needed from another person
 - ☐ Don't know/Don't remember
11. What currently best describes your **use of a mobility device** such as a wheelchair, scooter, walker, or cane?
- ☐ I never use a mobility device
 - ☐ I sometimes use a mobility device
 - ☐ I always use a mobility device
 - ☐ Don't know/Don't remember
12. What best describes your current ability to **walk by yourself** without resting? That is, without the help of another person or the help of a mobility device.
- ☐ I can walk several blocks by myself without resting or using a mobility device
 - ☐ I can walk one block by myself without resting or using a mobility device
 - ☐ I can walk from one room to another by myself without resting or using a mobility device
 - ☐ I am not able to walk by myself without resting or using a mobility device
 - ☐ Don't know/Don't remember
13. Do you currently have difficulty **walking up or down 12 stairs**?
- ☐ I have no difficulty walking up or down 12 stairs
 - ☐ I have some difficulty walking up or down 12 stairs
 - ☐ I have a lot of difficulty walking up or down 12 stairs
 - ☐ I am not able to walk up or down 12 stairs
 - ☐ Don't know/Don't remember
14. How often does your **physical health or emotional problems** currently interfere with your social activities (like visiting friends, relatives, etc.)?
- ☐ All of the time
 - ☐ Most of the time
 - ☐ Some of the time
 - ☐ A little of the time
 - ☐ None of the time
 - ☐ Don't know/Don't remember

15. How much does **pain** currently interfere with your normal activities?
- ☐ All of the time
 - ☐ Most of the time
 - ☐ Some of the time
 - ☐ A little of the time
 - ☐ None of the time
 - ☐ Don't know/Don't remember
16. Overall, since you left the hospital, how satisfied are you with your recovery?
- ☐ Not at all satisfied
 - ☐ Slightly satisfied
 - ☐ Moderately satisfied
 - ☐ Quite a bit satisfied
 - ☐ Extremely satisfied
 - ☐ Don't know/Don't remember

Section 3. Health Care Experiences

Now, we would like to hear about your experiences while you were at the hospital listed in the cover letter and any other place where you received care after the hospital.

In the following questions, the term “medical staff” means doctors, nurses, physical or occupational therapists and any other medical professionals who helped take care of you at the hospital and afterwards, in other facilities or at home. For example, after leaving the hospital you may have received care from medical staff in a nursing home, rehabilitation facility, assisted living facility, a doctor's office, or at home.

We'd like to learn about your experience as you were leaving the hospital in the cover letter.

17. Looking back to the time you left the hospital, overall, ***how prepared did you feel to leave?***
- ☐ Unprepared
 - ☐ Moderately prepared
 - ☐ Very prepared
 - ☐ Don't know/Don't remember

Thinking about when you left the hospital listed in the cover letter, ***how much do you agree or disagree with the following statement?***

18. The medical staff took your preferences and those of your family or your caregiver into account in deciding what health care services you should have after you left the hospital.

☐ Strongly Disagree
☐ Disagree
☐ Agree
☐ Strongly Agree
☐ Don't Know/Don't Remember
☐ Not Applicable

19. Where do you reside now?

☐ At my own home, in someone else's home, or in an assisted living facility (**Continue with the next section by following the arrow**)
☐ In a rehabilitation center, nursing home, or other health care facility (**Skip to Question 26 located on Page 6**)

- 20. Before you prepared to go home (or to someone else's home, or to an assisted living facility), you and your family or caregiver had a good understanding of ***how to take care of yourself***.

☐ Strongly Disagree
☐ Disagree
☐ Agree
☐ Strongly Agree
☐ Don't Know/Don't Remember
☐ Not Applicable

21. Before you prepared to go home (or to someone else's home, or to an assisted living facility), ***medical staff clearly explained how to take your medications***.

☐ Strongly Disagree
☐ Disagree
☐ Agree
☐ Strongly Agree
☐ Don't Know/Don't Remember
☐ Not Applicable, did not receive new medications

22. Before you prepared to go home (or to someone else's home, or to an assisted living facility), **medical staff clearly explained what follow-up appointments or treatments would be needed.**
- ☐ Strongly Disagree
 - ☐ Disagree
 - ☐ Agree
 - ☐ Strongly Agree
 - ☐ Don't Know/Don't Remember
 - ☐ Not Applicable
23. Overall, since you returned home (or to someone else's home, or to an assisted living facility), **you and your caregivers have been able to manage your health needs.**
- ☐ Strongly Disagree
 - ☐ Disagree
 - ☐ Agree
 - ☐ Strongly Agree
 - ☐ Don't Know/Don't Remember
 - ☐ Not applicable
24. Before you prepared to go home (or to someone else's home, or to an assisted living facility), **did doctors, nurses, or other staff talk with you about whether you would have the help you needed when you got home?**
- ☐ Yes
 - ☐ No
 - ☐ Don't Know/Don't Remember
 - ☐ Not applicable
25. Since leaving the hospital, if you needed help at home to manage your health, **did medical staff arrange services for you at home to help manage your health?**
- ☐ Yes
 - ☐ No
 - ☐ Don't Know/Don't Remember
 - ☐ Not Applicable, did not require help at home

26. Now we would like you to think about all of the health care you received **after** leaving the hospital. Using any number from 0 to 10, where 0 is the worst health care possible and 10 is the best health care possible, what number would you use to rate all of your health care after leaving the hospital?

- | | | |
|---|----------------------------|---|
| ☐ 0 (Worst possible) | ☐ 4 | ☐ 8 |
| ☐ 1 | ☐ 5 | ☐ 9 |
| ☐ 2 | ☐ 6 | ☐ 10 (Best possible) |
| ☐ 3 | ☐ 7 | |

The next question asks you to describe your time at the hospital listed in the cover letter and all health care you received after leaving the hospital, in your own words.

27. What has **gone well** in your experiences with all health care you received since you went to the hospital listed in the cover letter? Please explain what happened, how it happened, and how it felt to you.

Section 4. Personal Characteristics

28. What is the highest grade or level of school that you completed?

- ☐ 8th grade or less
- ☐ Some high school, but did not graduate
- ☐ High school graduate or GED
- ☐ Some college or 2-year (associate's) degree
- ☐ 4-year (bachelor's) college degree or equivalent
- ☐ More than 4-year college degree (such as a master's or doctoral degree)

Thank you for completing the survey!
Please mail it back in the enclosed postage-paid envelope
Abt Associates, P.O. Box 5720, Hopkins, MN 55343

Appendix E: Supplemental Statistics on Comparison Groups

This appendix provides statistics on balance between our treated and comparison groups to assess comparability across groups. It also provides descriptive statistics of the national sample balancing weights used for Model Years 7 and 8 (2024–2025) impact results. Exhibits E.1 and E.2 provide summaries of the provider-level balance of the Model Years 4 through 6 (2021–2023) comparison group, for hospitals and physician group practices (PGPs), respectively. They provide the average standardized mean difference (SMD) and the count of outcomes with an average SMD value greater than 0.1 in absolute value, for each clinical episode, between treated and comparison providers, for matching variables. Lists and definitions of matching variables for each clinical episode can be found in **Appendix F** of the BPCI Advanced Fifth Evaluation Report.¹ Exhibit E.3 furnishes descriptive statistics of the balancing weights we used for Model Years 7 and 8 impacts. Finally, Exhibits E.4–E.6 provide episode-level balance, before and after our comparison group procedures, separately for Model Years 4 through 6 and Model Years 7 and 8, and broken down by medical and surgical episodes.

All episode-level balance statistics pertain to impact estimation samples for total allowed payments. Estimation samples for other outcomes were nearly identical but can differ slightly due to different exclusion criteria. “Patient Complexity” refers to the Hierarchical Condition Category (HCC) score for Model Years 4 through 6 estimates; and it refers to the HCC count, computed based on model risk-adjustment HCCs, for Model Years 7 and 8 estimates. “Affiliated IRF” is a binary variable, based on whether the hospital was listed in the Provider of Services (POS) data as having an offsite or onsite inpatient prospective payment system (IPPS)-exempt inpatient rehabilitation facility (IRF). “Swing beds” is a binary variable, based on whether the provider is listed in the POS data as having swing beds. Total and operating margins were winsorized at the 1st and 99th percentiles since cost reports can be prone to data entry and other errors.

For detailed clinical episode-level balance statistics on the full Model Year 4 comparison group, see **Appendix F** of the Fifth Evaluation Report. For detailed balance statistics for the Model Years 1 and 2 and Model Year 3 comparison groups, see **Appendix F** of the Second Evaluation Report (hospital balance in Model Years 1 and 2); **Appendix D** of the Third Evaluation Report (PGP balance in Model Years 1 and 2); and **Appendix D** of the Fourth Evaluation Report (Model Year 3).

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods used to determine the sample, the patient mix measures evaluated, and data sources.

¹ All the BPCI Advanced Evaluation Reports and their appendices are available for download at <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Exhibit E.1: Provider-Level Balance for Impact Estimates, BPCI Advanced Hospitals, Before and After Match Prioritization

CESLG	Clinical Episode	Average SMD (Number of Large SMDs)				
		Original Match			After Match Prioritization	
		MY4	MY5	MY6	MY5	MY6
Cardiac Care	<i>Acute Myocardial Infarction</i>	0.04 (0)	0.04 (0)	0.13 (4)	0.07 (0)	0.11 (2)
	<i>Cardiac Arrhythmia</i>	0.03 (0)	0.06 (0)	0.22 (8)	0.06 (0)	0.08 (0)
	<i>Congestive Heart Failure</i>	0.04 (0)	0.06 (0)	0.13 (3)	0.05 (0)	0.10 (0)
Gastrointestinal Care	<i>Disorders of Liver Except Malignancy, Cirrhosis, or Alcoholic Hepatitis</i>	0.07 (0)	0.11 (2)	0.48 (9)	0.12 (0)	0.22 (7)
	<i>Gastrointestinal Hemorrhage*</i>	0.03 (0)	0.06 (0)	0.18 (3)	0.07 (0)	0.16 (5)
	<i>Gastrointestinal Obstruction</i>	0.03 (0)	0.05 (0)	0.22 (7)	0.05 (0)	0.22 (7)
Medical & Critical Care	<i>Cellulitis</i>	0.03 (0)	0.04 (0)	0.15 (2)	0.05 (0)	0.12 (1)
	<i>Chronic Obstructive Pulmonary Disease, Bronchitis, Asthma</i>	0.05 (0)	0.05 (0)	0.19 (6)	0.08 (0)	0.09 (2)
	<i>Renal Failure</i>	0.03 (0)	0.05 (0)	0.18 (4)	0.06 (0)	0.08 (1)
	<i>Sepsis</i>	0.04 (0)	0.05 (0)	0.18 (3)	0.07 (0)	0.07 (0)
	<i>Simple Pneumonia and Respiratory Infections</i>	0.04 (0)	0.06 (0)	0.14 (3)	0.05 (0)	0.08 (2)
	<i>Urinary Tract Infection</i>	0.04 (0)	0.07 (0)	0.26 (6)	0.08 (0)	0.11 (2)
Neurological Care	<i>Seizures</i>	0.06 (0)	0.09 (0)	0.21 (5)	0.09 (0)	0.16 (3)
	<i>Stroke</i>	0.03 (0)	0.07 (0)	0.19 (6)	0.09 (0)	0.08 (2)
Cardiac Procedures	<i>Cardiac Defibrillator (Outpatient)</i>	0.10 (1)	0.09 (1)	0.36 (7)	0.10 (0)	0.19 (2)
	<i>Coronary Artery Bypass Graft</i>	0.08 (1)	0.12 (2)	0.54 (11)	0.10 (1)	0.30 (7)
	<i>Pacemaker</i>	0.08 (0)	0.10 (1)	0.52 (12)	0.15 (2)	0.24 (6)
	<i>Percutaneous Coronary Intervention (Inpatient)*</i>	0.06 (0)	0.08 (0)	0.35 (8)	0.11 (2)	0.26 (6)
	<i>Percutaneous Coronary Intervention (Outpatient)*</i>	0.09 (0)	0.10 (1)	0.33 (7)	0.11 (1)	0.34 (8)
Gastrointestinal Surgery	<i>Major Bowel Procedure*</i>	0.09 (0)	0.09 (1)	0.26 (6)	0.12 (1)	0.29 (8)
Orthopedics	<i>Fractures of the Femur and Hip or Pelvis</i>	0.07 (0)	0.10 (1)	0.27 (7)	0.06 (0)	0.18 (5)
	<i>Hip and Femur Except Major Joint</i>	0.04 (0)	0.09 (1)	0.25 (8)	0.10 (1)	0.13 (2)
	<i>Lower Extremity and Humerus Procedure Except Hip, Foot, Femur</i>	0.09 (1)	0.14 (2)	0.23 (5)	0.12 (1)	0.16 (3)
	<i>Major Joint Replacement of the Lower Extremity</i>	0.08 (0)	0.07 (0)	0.17 (4)	0.10 (0)	0.15 (4)
Spinal Procedures	<i>Back and Neck Except Spinal Fusion (Outpatient)</i>	0.09 (1)	0.14 (2)	0.21 (6)	0.11 (1)	0.08 (1)
	<i>Spinal Fusion</i>	0.05 (0)	0.07 (0)	0.24 (6)	0.09 (0)	0.14 (4)

Note: See the first page of this appendix for data sources and methodology.

Exhibit E.2: Provider-Level Balance for Impact Estimates, BPCI Advanced PGPs, Before and After Match Prioritization

CESLG	Clinical Episode	Average SMD (Number of Large SMDs)				
		Original Match			After Match Prioritization	
		MY4	MY5	MY6	MY5	MY6
Orthopedics	<i>Hip and Femur Procedures Except Major Joint</i>	0.10 (0)	0.13 (3)	0.48 (17)	0.17 (5)	0.42 (18)
	<i>Major Joint Replacement of the Lower Extremity</i>	0.05 (0)	0.06 (0)	0.27 (6)	0.08 (0)	0.19 (3)
	<i>Major Joint Replacement of the Upper Extremity</i>	0.06 (0)	0.11 (0)	0.26 (6)	0.11 (2)	0.17 (6)

Note: See the first page of this appendix for data sources and methodology.

Exhibit E.3: Characteristics of National Sample Balancing Weights for Impact Estimates, Model Years 7–8 (Jan. 2024–Apr. 2025)

Sensitivity	Grouping	Min.	Percentile			Max.
			25 th	50 th	75 th	
CESLG Quarter Weights	<i>Combined Model</i>	0.01	0.02	0.04	0.05	0.06
	<i>Medical</i>	0.02	0.03	0.04	0.04	0.05
	<i>Surgical</i>	0.01	0.01	0.02	0.05	0.06
Sensitivity Analysis: Alternative Weights	<i>Combined Model</i>	0.00	0.01	0.03	0.05	0.10
	<i>Medical</i>	0.00	0.02	0.03	0.05	0.09
	<i>Surgical</i>	0.00	0.01	0.02	0.04	0.12

Note: See the first page of this appendix for data sources and methodology.

Exhibit E.4: Episode-Level Balance of Impact Estimates Samples at Baseline, Standardized Mean Differences, Medical and Surgical Episodes, Model Years 4–6 (Jan. 2021–Dec. 2023) and Model Years 7–8 (Jan. 2024–Apr. 2025)

Category	Variable	MY4–6	MY7–8		
		Matched Sample	No Weights	CESLG Quarter Weights	Alternative Weights
Patient Characteristics	<i>Age</i>	0.02	0.05	0.01	0.00
	<i>Patient Complexity</i>	0.00	0.06	0.05	0.03
	<i>Prior IPPS Hospital Use</i>	0.01	0.07	0.05	0.04
	<i>Prior PAC Use</i>	0.05	0.13	0.09	0.07
CESLG of Episode	<i>Cardiac Care</i>	-0.01	0.18	-0.01	-0.01
	<i>Gastrointestinal Care</i>	0.00	-0.08	0.00	0.00
	<i>Medical & Critical Care</i>	0.00	0.07	0.02	0.02
	<i>Neurological Care</i>	-0.01	-0.10	-0.01	-0.01
	<i>Cardiac Procedures</i>	-0.02	-0.40	0.00	0.00
	<i>Gastrointestinal Procedures</i>	0.00	-0.10	0.00	0.00
	<i>Orthopedics</i>	0.02	0.21	0.00	0.00
	<i>Spinal Procedures</i>	0.01	-0.09	-0.02	-0.02

Category	Variable	MY4–6	MY7–8		
		Matched Sample	No Weights	CESLG Quarter Weights	Alternative Weights
Hospital Ownership	<i>For-Profit</i>	0.25	0.38	0.39	0.42
	<i>Nonprofit</i>	-0.19	-0.20	-0.20	-0.23
	<i>Government</i>	-0.02	-0.20	-0.20	-0.19
Region of Hospital	<i>South</i>	0.09	-0.03	-0.02	-0.03
	<i>Midwest</i>	0.09	0.22	0.22	0.26
	<i>Northeast</i>	-0.22	-0.11	-0.11	-0.16
	<i>West</i>	0.02	-0.13	-0.13	-0.12
Other Hospital Characteristics	<i>IRF in Market</i>	-0.03	0.45	0.48	-0.01
	<i>Bed Count</i>	-0.12	0.08	0.13	0.02
	<i>Affiliated IRF</i>	-0.02	0.22	0.24	0.22
	<i>Swing Beds</i>	-0.09	-0.13	-0.15	-0.07
	<i>Part of Health System</i>	0.01	0.16	0.19	0.08
	<i>Operating Margin (Winsorized)</i>	0.19	0.23	0.24	0.26
	<i>Total Margin (Winsorized)</i>	0.09	0.17	0.19	0.20

Note: See the first page of this appendix for data sources and methodology.

Exhibit E.5: Episode-Level Balance of Impact Estimates Samples at Baseline, Standardized Mean Differences, Medical Episodes, Model Years 4–6 (Jan. 2021–Dec. 2023) and Model Years 7–8 (Jan. 2024–Apr. 2025)

Category	Variable	MY4–6		MY7–8		
		National Sample	Matched Sample	No Weights	CESLG Quarter Weights	Alternative Weights
Patient Characteristics	<i>Age</i>	0.03	0.02	0.03	0.01	0.00
	<i>Patient Complexity</i>	0.04	0.00	0.09	0.07	0.06
	<i>Prior IPPS Hospital Use</i>	0.03	0.01	0.07	0.06	0.05
	<i>Prior PAC Use</i>	0.09	0.05	0.12	0.11	0.09
CESLG of Episode	<i>Cardiac Care</i>	-0.08	-0.01	0.18	-0.01	-0.02
	<i>Gastrointestinal Care</i>	-0.17	0.00	-0.12	0.00	0.00
	<i>Medical & Critical Care</i>	0.20	0.01	-0.01	0.02	0.02
	<i>Neurological Care</i>	-0.07	-0.01	-0.15	-0.01	-0.01
Hospital Ownership	<i>For-Profit</i>	0.27	0.30	0.50	0.50	0.52
	<i>Nonprofit</i>	-0.10	-0.22	-0.26	-0.26	-0.29
	<i>Government</i>	-0.18	-0.04	-0.25	-0.25	-0.23
Region of Hospital	<i>South</i>	0.08	0.08	0.00	0.00	0.00
	<i>Midwest</i>	0.02	0.12	0.11	0.10	0.15
	<i>Northeast</i>	-0.13	-0.20	-0.05	-0.05	-0.10
	<i>West</i>	0.01	-0.02	-0.08	-0.07	-0.07

Category	Variable	MY4–6		MY7–8		
		National Sample	Matched Sample	No Weights	CESLG Quarter Weights	Alternative Weights
Other Hospital Characteristics	<i>IRF in Market</i>	0.31	-0.04	0.51	0.51	0.00
	<i>Bed Count</i>	0.13	-0.14	0.12	0.13	0.01
	<i>Affiliated IRF</i>	0.05	-0.02	0.26	0.27	0.26
	<i>Swing Beds</i>	-0.19	-0.09	-0.13	-0.14	-0.05
	<i>Part of Health System</i>	0.27	-0.01	0.15	0.16	0.05
	<i>Operating Margin (Winsorized)</i>	0.28	0.21	0.27	0.27	0.28
	<i>Total Margin (Winsorized)</i>	0.16	0.10	0.20	0.20	0.20

Note: See the first page of this appendix for data sources and methodology.

Exhibit E.6: Episode-Level Balance of Impact Estimates Samples at Baseline, Standardized Mean Differences, Surgical Episodes, Model Years 4–6 (Jan. 2021–Dec. 2023) and Model Years 7–8 (Jan. 2024–Apr. 2025)

Category	Variable	MY4–6		MY7–8		
		National Sample	Matched Sample	No Weights	CESLG Quarter Weights	Alternative Weights
Patient Characteristics	<i>Age</i>	0.04	0.03	0.05	-0.01	-0.03
	<i>Patient Complexity</i>	-0.08	0.00	-0.18	-0.02	-0.05
	<i>Prior IPPS Hospital Use</i>	-0.03	0.01	-0.03	0.01	-0.01
	<i>Prior PAC Use</i>	0.05	0.04	0.07	0.02	-0.01
CESLG of Episode	<i>Cardiac Procedures</i>	-0.45	-0.06	-0.70	0.01	0.01
	<i>Gastrointestinal Procedures</i>	-0.24	-0.01	-0.12	0.00	0.00
	<i>Orthopedics</i>	0.38	0.04	0.74	0.02	0.02
	<i>Spinal Procedures</i>	0.15	0.01	-0.09	-0.03	-0.03
Hospital Ownership	<i>For-Profit</i>	-0.03	0.00	0.13	0.13	0.18
	<i>Nonprofit</i>	0.00	-0.08	-0.05	-0.05	-0.10
	<i>Government</i>	0.03	0.11	-0.09	-0.09	-0.07
Region of Hospital	<i>South</i>	0.11	0.12	-0.10	-0.07	-0.11
	<i>Midwest</i>	-0.03	-0.06	0.49	0.47	0.53
	<i>Northeast</i>	-0.16	-0.29	-0.32	-0.31	-0.35
	<i>West</i>	0.04	0.21	-0.22	-0.25	-0.23
Other Hospital Characteristics	<i>IRF in Market</i>	0.30	0.03	0.34	0.42	-0.02
	<i>Bed Count</i>	0.15	-0.04	-0.01	0.12	0.05
	<i>Affiliated IRF</i>	-0.05	-0.01	0.12	0.18	0.15
	<i>Swing Beds</i>	-0.11	-0.09	-0.15	-0.19	-0.13
	<i>Part of Health System</i>	0.02	0.08	0.19	0.25	0.13
	<i>Operating Margin (Winsorized)</i>	0.13	0.08	0.17	0.16	0.21
	<i>Total Margin (Winsorized)</i>	0.10	0.04	0.14	0.15	0.20

Note: See the first page of this appendix for data sources and methodology.

Appendix F: Participant Characteristics

This appendix details descriptive analyses of participant characteristics, which we conducted by calculating the counts and percentages for hospital and physician group practice (PGP) participation in Model Years 1 through 8. Because we found that some PGPs that had participation agreements with CMS were not initiating episodes, we removed “inactive” PGPs from our analyses, defined as PGPs for which there were no discharges or procedures associated with the PGP’s tax identification number (TIN) within our claims database in 2013–2025. To calculate the shares of participating hospitals and PGPs in medical or surgical clinical episodes and clinical episode service line groupings (CESLGs), we limited our calculation to participants that had at least one discharge or procedure in medical or surgical clinical episodes or within the CESLGs for the model year to ensure participants had some engagement in their chosen CESLGs.

In our analysis of participation under conveners, a convener is a participant with at least one downstream episode initiator. Conveners facilitate coordination between downstream episode initiators and other providers and bear financial risk. Conversely, a non-convener participant is a hospital or PGP that bears financial risk only for itself. For the counts of unique conveners and counts of hospitals and PGPs under a convener/non-convener we omit conveners with duplicate TINs and exclude episode-initiating conveners; as a result, the total count in Exhibit F.2 differs from that of Exhibit F.1, where we did not apply these restrictions.

In our assessment of the extent of BPCI Advanced participation (the model’s reach), we limited eligible hospitals to those that were eligible for BPCI Advanced in each Model Year and met the baseline volume criterion for at least one clinical episode. For instance, Inpatient Prospective Payment System-exempt cancer hospitals and hospitals located in Maryland were not eligible to participate in BPCI Advanced, so they were excluded. The number of clinicians is a unique count of clinicians who treated Medicare beneficiaries who met the BPCI Advanced beneficiary inclusion criteria at a BPCI Advanced eligible hospital. Clinicians with BPCI Advanced discharges or procedures were identified as either attending or operating National Provider Identifiers (NPIs) at a BPCI Advanced hospital or the attending or operating NPI on the hospital claim when the episode was attributed to a BPCI Advanced PGP. Eligible discharges and procedures include hospital discharges and outpatient procedures that met the model eligibility requirements, including having a BPCI Advanced Medicare Severity-Diagnosis Related Group or Healthcare Common Procedure Coding System code, being at a hospital that was eligible to be in BPCI Advanced, and meeting other beneficiary inclusion criteria. The minimum hospital baseline volume criterion was not applied in our assessment of reach to discharges and clinicians. In counting episode volume for BPCI Advanced Hospitals and PGPs presented in Exhibit F.8, we further restricted the analysis to first admission episodes only. As a result, these episode volume counts differ from the counts of BPCI Advanced discharges and procedures presented in Exhibit F.3.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods.

A. BPCI Advanced Participation

Exhibit F.1: BPCI Advanced Episode Initiators, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Count of BPCI Advanced Hospitals (N)		Count of BPCI Advanced PGPs (N)		
	All Hospitals	Hospitals with Volume	All PGPs	Active PGPs	Active PGPs with Volume
MY1–2	714	713	580	518	504
MY3	1,010	1,002	1,031	728	574
MY4	682	678	523	389	352
MY5	435	431	396	307	254
MY6	106	105	174	87	61
MY7	123	122	124	93	61
MY8	105	104	103	89	39
Unique MY1–8	1,103	1,098	1,198	871	793

Note: See the first page of this appendix for data sources and methodology.

Exhibit F.2: BPCI Advanced Conveners and Participants by Convener Status, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Count of Conveners (N)	Count of Downstream Episode Initiators (N)		Count of Non-Conveners (N)	
		Hospitals	Active PGPs	Hospitals	Active PGPs
MY1–2	75	592	406	116	110
MY3	93	823	540	176	186
MY4	66	580	213	94	174
MY5	42	369	138	60	167
MY6	19	75	25	27	60
MY7	12	51	35	66	58
MY8	12	47	33	55	56
Unique MY1–8	107	892	668	263	209

Note: Results in this exhibit might differ from those presented elsewhere because conveners with duplicate TINs and episode-initiating conveners are excluded. See the first page of this appendix for data sources and methodology.

B. BPCI Advanced Reach

Exhibit F.3: Eligible Discharges in BPCI Advanced, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Discharges	Overall	Medical	Surgical
MY1–2	All Eligible (N)	4,910,164	3,329,712	1,580,452
	BPCI Advanced Discharges (N)	827,940	568,409	259,531
	BPCI Advanced Share (%)	16.9	17.1	16.4

Model Year	Discharges	Overall	Medical	Surgical
MY3	<i>All Eligible (N)</i>	3,446,908	2,304,432	1,142,476
	<i>BPCI Advanced Discharges (N)</i>	587,903	468,711	119,192
	<i>BPCI Advanced Share (%)</i>	17.1	20.3	10.4
MY4	<i>All Eligible (N)</i>	3,257,463	2,174,404	1,083,059
	<i>BPCI Advanced Discharges (N)</i>	615,381	507,675	107,706
	<i>BPCI Advanced Share (%)</i>	18.9	23.3	9.9
MY5	<i>All Eligible (N)</i>	3,244,520	2,091,900	1,152,620
	<i>BPCI Advanced Discharges (N)</i>	383,630	293,316	90,314
	<i>BPCI Advanced Share (%)</i>	11.8	14.0	7.8
MY6	<i>All Eligible (N)</i>	3,224,066	2,025,459	1,198,607
	<i>BPCI Advanced Discharges (N)</i>	93,131	53,429	39,702
	<i>BPCI Advanced Share (%)</i>	2.9	2.6	3.3
MY7	<i>All Eligible (N)</i>	2,022,543	1,245,581	776,962
	<i>BPCI Advanced Discharges (N)</i>	44,945	27,142	17,803
	<i>BPCI Advanced Share (%)</i>	2.2	2.2	2.3
MY8	<i>All Eligible (N)</i>	1,628,620	979,451	649,169
	<i>BPCI Advanced Discharges (N)</i>	24,979	13,698	11,281
	<i>BPCI Advanced Share (%)</i>	1.5	1.4	1.7
Unique MY1–8	<i>All Eligible (N)</i>	21,734,284	14,150,939	7,583,345
	<i>BPCI Advanced Discharges (N)</i>	2,577,909	1,932,380	645,529
	<i>BPCI Advanced Share (%)</i>	11.9	13.7	8.5

Note: See the first page of this appendix for data sources and methodology.

Exhibit F.4: Eligible Hospitals in BPCI Advanced, Hospitals, Model Years 1–8
(Oct. 2018–Dec. 2025)

Model Year	Unique Number of Eligible Hospitals (N)	Unique Number of Participating BPCI Advanced Hospitals	
		Number (N)	Percent (%)
MY1–2	3,269	714	22.0
MY3	3,212	1,010	31.4
MY4	3,148	682	21.7
MY5	3,110	435	14.0
MY6	3,101	106	3.5
MY7	3,068	123	4.0
MY8	3,068	105	3.4
Unique MY1–8	3,410	1,103	32.3

Note: See the first page of this appendix for data sources and methodology.

Exhibit F.5: Eligible Clinicians in BPCI Advanced, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Measurement of Unique Clinicians	Overall	Medical	Surgical
MY1–2	<i>With Eligible Episodes (N)</i>	375,635	234,351	141,284
	<i>With BPCI Advanced Discharges (N)</i>	98,305	65,702	32,603
	<i>With BPCI Advanced Discharges Share (%)</i>	26.2	28.0	23.1
MY3	<i>With Eligible Episodes (N)</i>	343,202	216,380	126,822
	<i>With BPCI Advanced Discharges (N)</i>	97,577	73,018	24,559
	<i>With BPCI Advanced Discharges Share (%)</i>	28.4	33.7	19.4
MY4	<i>With Eligible Episodes (N)</i>	345,178	218,061	127,117
	<i>With BPCI Advanced Discharges (N)</i>	81,546	64,992	16,554
	<i>With BPCI Advanced Discharges Share (%)</i>	23.6	29.8	13.0
MY5	<i>With Eligible Episodes (N)</i>	344,654	218,107	126,547
	<i>With BPCI Advanced Discharges (N)</i>	55,897	43,434	12,463
	<i>With BPCI Advanced Discharges Share (%)</i>	16.2	19.9	9.8
MY6	<i>With Eligible Episodes (N)</i>	344,192	216,015	128,177
	<i>With BPCI Advanced Discharges (N)</i>	12,300	8,231	4,069
	<i>With BPCI Advanced Discharges Share (%)</i>	3.6	3.8	3.2
MY7	<i>With Eligible Episodes (N)</i>	288,116	182,492	105,624
	<i>With BPCI Advanced Discharges (N)</i>	10,133	7,430	2,703
	<i>With BPCI Advanced Discharges Share (%)</i>	3.5	4.1	2.6
MY8	<i>With Eligible Episodes (N)</i>	264,675	165,884	98,791
	<i>With BPCI Advanced Discharges (N)</i>	6,522	4,523	1,999
	<i>With BPCI Advanced Discharges Share (%)</i>	2.5	2.7	2.0
Unique MY1–8	<i>With Eligible Episodes (N)</i>	676,800	425,850	250,950
	<i>With BPCI Advanced Discharges (N)</i>	188,429	134,444	53,985
	<i>With BPCI Advanced Discharges Share (%)</i>	27.8	31.6	21.5

Note: See the first page of this appendix for data sources and methodology.

C. BPCI Advanced CE and CESLG Selection

Exhibit F.6: Medical and Surgical CE/CESLG Selection in BPCI Advanced, Hospitals and PGPs, Combined Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Clinical Episode Type	Hospitals		PGPs	
		Number (N)	Percent (%)	Number (N)	Percent (%)
MY1–8	<i>Medical and Surgical</i>	612	55.7	203	25.6
	<i>Medical Only</i>	422	38.4	132	16.6
	<i>Surgical Only</i>	64	5.8	458	57.8

Note: The denominator is the MY1–8 unique providers with episode count as shown in Exhibit F.1. See the first page of this appendix for data sources and methodology.

Exhibit F.7: Medical and Surgical CE/CESLG Selection in BPCI Advanced, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Clinical Episode Type	Hospitals		PGPs	
		Number (N)	Percent (%)	Number (N)	Percent (%)
MY1–2	<i>Medical</i>	645	90.5	181	35.9
	<i>Surgical</i>	411	57.6	450	89.3
MY3	<i>Medical</i>	930	92.8	258	44.9
	<i>Surgical</i>	492	49.1	438	76.3
MY4	<i>Medical</i>	643	94.8	186	52.8
	<i>Surgical</i>	312	46.0	199	56.5
MY5	<i>Medical</i>	418	96.1	141	55.5
	<i>Surgical</i>	237	55.0	137	53.9
MY6	<i>Medical</i>	96	90.6	7	11.5
	<i>Surgical</i>	49	46.7	60	98.4
MY7	<i>Medical</i>	113	91.9	11	18.0
	<i>Surgical</i>	51	41.8	53	86.9
MY8	<i>Medical</i>	97	92.4	10	25.6
	<i>Surgical</i>	43	41.3	32	82.1

Note: The denominator is the unique providers with episode count for each model year as shown in Exhibit F.1. See the first page of this appendix for data sources and methodology.

Exhibit F.8: First Admission Episode Volume in BPCI Advanced, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Dec. 2025)

Model Year	Hospitals		PGPs	
	Medical (N)	Surgical (N)	Medical (N)	Surgical (N)
MY1–2	295,629	58,705	143,225	178,293
MY3	264,859	46,497	110,499	82,114
MY4	323,832	51,612	85,849	48,481
MY5	198,629	39,797	41,407	46,222
MY6	42,925	13,223	361	25,002
MY7	25,082	6,512	2,060	11,291
MY8	12,383	4,418	1,690	6,918

Note: See the first page of this appendix for data sources and methodology.

Exhibit F.9: CESLG Selection in BPCI Advanced, Hospitals and PGPs, Model Years 1–8
(Oct. 2018–Dec. 2025)

Model Year	Clinical Episode Type	CESLG	Hospitals		PGPs	
			Number (N)	Percent (%)	Number (N)	Percent (%)
MY1–2	<i>Medical</i>	Cardiac Care	500	70.1	159	31.5
		Gastrointestinal Care	193	27.1	115	22.8
		Medical & Critical Care	536	75.2	152	30.2
		Neurological Care	226	31.7	86	17.1
	<i>Surgical</i>	Cardiac Procedures	199	27.9	103	20.4
		Gastrointestinal Surgery	38	5.3	51	10.1
		Orthopedics	292	41.0	369	73.2
		Spinal Procedures	97	13.6	155	30.8
MY3	<i>Medical</i>	Cardiac Care	560	55.9	184	32.1
		Gastrointestinal Care	291	29.0	149	26.0
		Medical & Critical Care	792	79.0	210	36.6
		Neurological Care	342	34.1	123	21.4
	<i>Surgical</i>	Cardiac Procedures	273	27.2	101	17.6
		Gastrointestinal Surgery	75	7.5	39	6.8
		Orthopedics	264	26.3	290	50.5
		Spinal Procedures	171	17.1	114	19.9
MY4	<i>Medical</i>	Cardiac Care	367	54.1	134	38.1
		Gastrointestinal Care	225	33.2	79	22.4
		Medical & Critical Care	497	73.3	148	42.0
		Neurological Care	297	43.8	98	27.8
	<i>Surgical</i>	Cardiac Procedures	82	12.1	26	7.4
		Gastrointestinal Surgery	71	10.5	16	4.5
		Orthopedics	166	24.5	134	38.1
		Spinal Procedures	124	18.3	72	20.5
MY5	<i>Medical</i>	Cardiac Care	265	61.5	102	40.2
		Gastrointestinal Care	184	42.7	57	22.4
		Medical & Critical Care	334	77.5	110	43.3
		Neurological Care	225	52.2	72	28.3
	<i>Surgical</i>	Cardiac Procedures	67	15.5	17	6.7
		Gastrointestinal Surgery	53	12.3	1	0.4
		Orthopedics	123	28.5	106	41.7
		Spinal Procedures	94	21.8	52	20.5

Model Year	Clinical Episode Type	CESLG	Hospitals		PGPs	
			Number (N)	Percent (%)	Number (N)	Percent (%)
MY6	<i>Medical</i>	Cardiac Care	65	61.9	5	8.2
		Gastrointestinal Care	25	23.8	0	0.0
		Medical & Critical Care	73	69.5	4	6.6
		Neurological Care	39	37.1	0	0.0
	<i>Surgical</i>	Cardiac Procedures	13	12.4	2	3.3
		Gastrointestinal Surgery	12	11.4	1	1.6
		Orthopedics	28	26.7	55	90.2
		Spinal Procedures	16	15.2	21	34.4
MY7	<i>Medical</i>	Cardiac Care	70	57.4	6	9.8
		Gastrointestinal Care	32	26.2	2	3.3
		Medical & Critical Care	60	49.2	1	1.6
		Neurological Care	30	24.6	0	0.0
	<i>Surgical</i>	Cardiac Procedures	13	10.7	1	1.6
		Gastrointestinal Surgery	26	21.3	1	1.6
		Orthopedics	12	9.8	47	77.0
		Spinal Procedures	18	14.8	15	24.6
MY8	<i>Medical</i>	Cardiac Care	58	55.8	6	15.4
		Gastrointestinal Care	27	26.0	2	5.1
		Medical & Critical Care	56	53.8	8	20.5
		Neurological Care	25	24.0	0	0.0
	<i>Surgical</i>	Cardiac Procedures	11	10.6	1	2.6
		Gastrointestinal Surgery	21	26.0	2	5.1
		Orthopedics	11	10.6	28	71.8
		Spinal Procedures	17	16.3	4	10.3

Note: The denominator is the count of unique providers with episode count for each model year as shown in Exhibit F.1. See the first page of this appendix for data sources and methodology.

Exhibit F.10: Characteristics of Hospitals Participating in BPCI Advanced for Years Where Participants Could Join the BPCI Advanced Model

Hospital Characteristic		2019	2020	2024
Number of Hospitals (N)		714	1,010	123
Urban Location (%)		90.2	89.2	92.7
Safety-Net Hospital (%)		22.8	20.7	26.0
Affiliated With a Health System (%)		98.7	98.7	100.0
Teaching Hospital (%)		59.9	57.4	58.5
Ownership (%)	<i>Non-profit</i>	68.6	69.1	47.2
	<i>For-profit</i>	27.0	25.0	47.2
	<i>Government</i>	4.3	5.8	5.7
Bed Size (%)	<i><100</i>	8.7	9.5	5.7
	<i>101–249</i>	32.1	34.6	35.0
	<i>250+</i>	59.2	55.9	59.3
Swing Beds (%)	<i>None</i>	92.6	92.9	91.9
	<i><50</i>	2.7	2.3	2.4
	<i>50–99</i>	1.1	2.1	2.4
	<i>100+</i>	3.6	2.8	3.3
HCAHPS Overall Rating, Mean (%)	<i>Patients Who Gave a Rating of 6 or Lower</i>	8.8	—*	13.0
	<i>Patients Who Gave a Rating of 7 or 8</i>	20.7	—*	22.9
	<i>Patients Who Gave a Rating of 9 or 10</i>	70.5	—*	64.1
Year Joined BPCI Advanced (%)	<i>MY1–2</i>	100.0	63.5	74.0
	<i>MY3</i>	N/A	36.5	9.8
	<i>MY7</i>	N/A	N/A	16.3

Note: * HCAHPS data were not collected in 2020. See the first page of this appendix for data sources and methodology.

Appendix G: Supplemental Sample Characteristics

The following supplemental tables display additional characteristics of the evaluation sample. The sample count tables present the number of BPCI Advanced participants included in the evaluation sample after comparison group construction and the associated intervention episode counts. The counts of BPCI Advanced intervention episodes are based on the sample used to evaluate the impact of the model on total allowed standardized payments. We also present the shares of episodes in which the patient had a confirmed Coronavirus Disease 2019 (COVID-19) diagnosis during either the anchor hospitalization or procedure or the 90-day post-discharge period (PDP) for BPCI Advanced episodes and comparison group episodes in Model Years 4 through 6.

The patient mix tables present the results by episode type for BPCI Advanced episodes and comparison group episodes. The tables present the mean values in the baseline period and in Model Years 4 through 6 or Model Years 7 and 8 and are not risk adjusted. The relative change represents the difference in the means of the characteristics for BPCI Advanced episodes in the model years compared to the baseline period, relative to the change for the comparison group. For example, if the mean of a characteristic decreased more for the BPCI Advanced episodes from the baseline period to model years than the comparison group, then the relative change will be negative, indicating a greater decrease for the BPCI Advanced group. Similarly, if the mean of a characteristic increased more for the BPCI Advanced episodes from the baseline period to model years than the comparison group, then the relative change will be positive, indicating a greater increase for the BPCI Advanced group. The N reported represents the sample population in the model years, before removing any episodes with missing patient characteristics. Prior institutional post-acute care (PAC) setting and prior home health indicate the share of the sample with recent prior use of PAC services. The Hierarchical Condition Category (HCC) count is the average number of HCC indicators per patient in the sample, and the HCC score is the average score (or index). Prior institutional PAC setting, prior home health, HCC count, and HCC score are based on the 6 months prior to the anchor hospitalization or procedure. Results are also presented separately for hospitals and physician group practices (PGPs). An asterisk denotes statistically significant differences at the 10% level. Please note that prior home health and HCC score variables are not available in the Model Years 7 and 8 data and are absent from Exhibit G.7.

All results are based on the BPCI Advanced evaluation team's analysis of Medicare claims and enrollment data for episodes attributed to BPCI Advanced and comparison hospitals and PGPs and matched comparison providers. These results are based on clinical episodes with sufficient sample size for evaluation for Model Years 4 through 6 (January 1, 2021–December 31, 2023) and based on all clinical episodes for Model Years 7 and 8 (January 1, 2024–April 30, 2025).

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods.

A. Sample Counts

1. Model Years 4–6

Exhibit G.1: Count of BPCI Advanced Hospitals and Episodes Included in the BPCI Advanced Evaluation Sample, by Clinical Episode Type, Model Years 4–6
(Jan. 2021–Dec. 2023)

Clinical Episode Type	Count of Hospitals (N)	Count of Episodes (N)
All Clinical Episodes	671	596,116
Medical	602	493,997
Surgical	291	102,119

Note: See the first page of this appendix for data sources and methodology.

Exhibit G.2: Count of BPCI Advanced PGPs and Episodes Included in the BPCI Advanced Evaluation Sample, by Clinical Episode Type, Model Years 4–6
(Jan. 2021–Dec. 2023)

Clinical Episode Type	Count of PGPs (N)	Count of Episodes (N)
All Clinical Episodes	189	183,132
Medical	99	109,133
Surgical	98	73,999

Note: See the first page of this appendix for data sources and methodology.

Exhibit G.3: Share of Episodes With Confirmed COVID-19 Diagnosis, Model Years 4–6
(Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced Episodes (%)	Comparison Group Episodes (%)
All Clinical Episodes	20.1	20.0
Medical	24.5	23.6
Surgical	5.1	5.4

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit G.4: Count of BPCI Advanced Hospitals and Episodes Included in the BPCI Advanced Evaluation Sample, by Clinical Episode Type, Model Year 7–8
(Jan. 2024–Apr. 2025)

Clinical Episode Type	Count of Hospitals (N)	Count of Episodes (N)
Combined	122	42,417
Medical	113	33,757
Surgical	51	8,660

Note: See the first page of this appendix for data sources and methodology.

Exhibit G.5: Count of BPCI Advanced PGPs and Episodes Included in the BPCI Advanced Evaluation Sample, by Clinical Episode Type, Model Year 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	Count of PGPs (N)	Count of Episodes (N)
Combined	62	17,652
Medical	11	2,860
Surgical	54	14,792

Note: The surgical episode count includes a PGP that has episodes only in MY8 but not MY7. See the first page of this appendix for data sources and methodology.

B. Patient Characteristics

1. Model Years 4–6

Exhibit G.6: Patient Mix, Hospitals and PGPs, Baseline (Jan. 2015–Sept. 2018) and Intervention (Jan. 2021–Dec. 2023)

Clinical Episode	Patient Characteristic	BPCI Advanced		Comparison Group		Relative Change	P-Value
		Baseline Mean	MY4–6 Mean	Baseline Mean	MY4–6 Mean		
Medical	<i>Prior Institutional PAC Use</i>	32.1%	29.3%	31.8%	29.1%	-0.18 pp	0.387
	<i>Prior Home Health</i>	27.2%	25.4%	24.9%	23.8%	-0.65 pp*	0.014
	<i>Age: 80+ Years</i>	44.4%	44.0%	43.7%	43.8%	-0.48 pp	0.124
	<i>Male</i>	43.7%	45.1%	43.7%	45.2%	-0.04 pp	0.817
	<i>Disabled, No ESRD</i>	26.8%	23.6%	27.6%	23.6%	0.83 pp*	0.005
	<i>Dual Eligibility</i>	26.8%	24.1%	26.8%	23.3%	0.73 pp*	0.093
	<i>HCC Count</i>	3.0	3.0	2.9	3.0	0.00	0.937
	<i>HCC Score</i>	1.9	2.0	1.9	2.0	0.00	0.692
Surgical	<i>Prior Institutional PAC Use</i>	13.2%	9.3%	13.1%	9.9%	-0.71 pp*	0.027
	<i>Prior Home Health</i>	11.7%	9.9%	10.2%	9.5%	-1.17 pp*	0.001
	<i>Age: 80+ Years</i>	26.5%	25.9%	25.8%	25.9%	-0.80 pp	0.142
	<i>Male</i>	41.3%	41.6%	42.1%	42.5%	-0.09 pp	0.863
	<i>Disabled, No ESRD</i>	18.4%	13.4%	19.7%	15.1%	-0.32 pp	0.469
	<i>Dual Eligibility</i>	12.0%	8.4%	12.6%	9.4%	-0.28 pp	0.549
	<i>HCC Count</i>	1.5	1.5	1.5	1.6	-0.04	0.184
	<i>HCC Score</i>	1.1	1.1	1.1	1.1	-0.02	0.119

Note: The means are unweighted. See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit G.7: Patient Mix, Hospitals and PGPs, Baseline (Jan. 2016–Sept. 2018) and Intervention (Jan. 2024–Apr. 2025)

Clinical Episode	Patient Characteristic	BPCI Advanced		Comparison Group		Relative Change	P-Value
		Baseline Mean	MY7–8 Mean	Baseline Mean	MY7–8 Mean		
Medical	<i>Prior Institutional PAC Use</i>	35.6%	30.2%	30.1%	26.9%	-2.22 pp*	0.000
	<i>Age: 80+ Years</i>	44.9%	47.1%	43.5%	47.1%	-1.36 pp*	0.028
	<i>Male</i>	43.5%	45.6%	43.9%	44.6%	1.53 pp*	0.000
	<i>Disabled, No ESRD</i>	26.1%	20.3%	27.2%	21.3%	0.20 pp	0.753
	<i>Dual Eligibility</i>	29.7%	23.1%	29.6%	23.3%	-0.30 pp	0.767
	<i>HCC Count</i>	3.1	4.3	2.9	4.0	0.06	0.132
Surgical	<i>Prior Institutional PAC Use</i>	14.3%	10.2%	12.0%	10.0%	-2.19 pp*	0.000
	<i>Age: 80+ Years</i>	26.9%	27.6%	26.9%	28.8%	-1.27 pp	0.194
	<i>Male</i>	39.4%	39.7%	47.8%	47.9%	0.20 pp	0.760
	<i>Disabled, No ESRD</i>	15.8%	10.9%	20.2%	13.2%	2.17 pp*	0.005
	<i>Dual Eligibility</i>	11.2%	7.0%	14.7%	8.8%	1.69 pp	0.051
	<i>HCC Count</i>	1.5	2.1	1.8	2.4	-0.02	0.693

Note: The means are unweighted. See the first page of this appendix for data sources and methodology.

Appendix H: Expenditure and Utilization Outcomes – Results

The exhibits in Section A display the risk-adjusted difference-in-differences (DiD) results for all claims-based payment and utilization outcomes assessed in the report. Medicare payments were standardized to remove the effects of geographic differences in wages, extra amounts to account for teaching programs, and other policy factors. Total episode payments include all Medicare Parts A and B payments and beneficiary cost sharing, while total paid payments exclude beneficiary cost sharing. The “DiD as a Percent (%)” reflects the DiD estimate as a percentage of the BPCI Advanced baseline mean.

The exhibits in Section B present parallel trends test results that correspond to the claims-based payment and utilization outcomes assessed in the report. As described in **Appendix C**, we tested the null hypothesis that BPCI Advanced participants and comparison providers had parallel trends during the baseline period at the 10% level of statistical significance. We used the dagger symbol (‡) to denote estimates for which we rejected the null hypothesis of parallel trends, indicating a parallel trends failure for that result. A failure of the parallel trends test indicates that BPCI Advanced participants and matched comparison providers did not have the same baseline trends for that outcome; therefore, the impact estimate may partially reflect a continuation of preexisting trends rather than solely the impact of the BPCI Advanced Model.

The exhibits in Section C present sensitivity analyses that evaluate the robustness of BPCI Advanced impact estimates to alternative episode samples and analytic specifications. For Model Years 4 through 6, these sensitivity analyses include estimating impacts after excluding episodes initiated by providers participating in the prior BPCI initiative from the baseline (“No BPCI-C Episodes”), intent-to-treat (ITT) estimates that include episodes for all matched BPCI Advanced providers and their matched comparisons regardless of subsequent model withdrawal, and a sensitivity analysis using a national comparison group. For Model Years 7 and 8, these include estimates from a regression that included additional risk adjusters, such as a variety of Hierarchical Condition Category flags (“additional risk adjustment”), estimates utilizing alternative national sample balancing weights, and estimates that excluded all participants that participated in *any* clinical episode from the national comparison group (“excluding past participants from comparison group”).

These exhibits display pooled impact estimates based on clinical episodes with sufficient sample size for evaluation for Model Years 1 and 2 (October 1, 2018–December 31, 2019)¹ and Model

¹ We provide the pooled impact estimates for the skilled nursing facility (SNF) payments through the 90-day post-discharge period (PDP) and inpatient rehabilitation facility (IRF) payments through the 90-day PDP outcomes for Model Years 1 and 2 as these were not included in earlier reports. For detailed results of other outcomes from Model Years 1 and 2, refer to prior evaluation reports and appendices available on the BPCI Advanced Model web page: <https://www.cms.gov/priorities/innovation/innovation-models/bpci-advanced>.

Years 4 through 6 (January 1, 2021–December 31, 2023), and results based on all clinical episodes for Model Years 7 and 8 (January 1, 2024–April 30, 2025).

All results are based on the BPCI Advanced evaluation team’s analysis of Medicare claims and enrollment data for episodes attributed to BPCI Advanced and comparison hospitals and physician group practices (PGPs) and matched comparison providers.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods.

A. Pooled Impact Estimates

1. Model Years 1–2

Exhibit H.1: Impact of BPCI Advanced on SNF Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 1–2
(Oct. 2018–Dec. 2019)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY1&2 Episodes (N)	Baseline Mean (\$)	MY1&2 Mean (\$)	MY1&2 Episodes (N)	Baseline Mean (\$)	MY1&2 Mean (\$)							
Overall	500,687	5,551	4,702	Not reported	5,390	5,008	-466	-8.4	<0.001	-539	-392	-553	-378
Medical	386,825	5,269	4,622	Not reported	5,085	4,878	-439	-8.3	<0.001	-521	-358	-536	-343
Surgical	115,416	6,629	5,055	Not reported	6,542	5,534	-567	-8.6	<0.001	-718	-416	-747	-386

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.2: Impact of BPCI Advanced on IRF Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 1–2
(Oct. 2018–Dec. 2019)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY1&2 Episodes (N)	Baseline Mean (\$)	MY1&2 Mean (\$)	MY1&2 Episodes (N)	Baseline Mean (\$)	MY1&2 Mean (\$)							
Overall	500,687	1,214	1,116	Not reported	1,147	1,218	-170	-14.0	<0.001	-216	-124	-225	-115
Medical	386,825	1,066	1,113	Not reported	1,064	1,166	-56	-5.3	0.034	-99	-13	-108	-4
Surgical	115,416	1,759	1,142	Not reported	1,454	1,454	-576	-33.0	<0.001	-699	-453	-722	-429

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 4–6

Exhibit H.3: Impact of BPCI Advanced on Total Episode Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	26,716	28,131	748,841	26,418	28,844	-1,011	-3.8	<0.001	-1,190	-832	-1,224	-798
Medical	587,343	25,744	27,849	611,185	25,489	28,350	-756	-2.9	<0.001	-923	-588	-956	-556
Surgical	158,583	30,967	29,968	137,656	30,508	31,345	-1,835	-5.9	<0.001	-2,199	-1,471	-2,269	-1,401

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.4: Impact of BPCI Advanced on Total Paid Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	23,447	24,581	748,841	23,179	25,205	-892	-3.8	<0.001	-1,063	-720	-1,096	-687
Medical	587,343	22,484	24,328	611,185	22,257	24,733	-632	-2.8	<0.001	-783	-482	-812	-453
Surgical	158,583	27,656	26,308	137,656	27,231	27,607	-1,724	-6.2	<0.001	-2,091	-1,356	-2,162	-1,286

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.5: Impact of BPCI Advanced on SNF Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	4,869	4,370	748,841	4,740	4,693	-452	-9.3	<0.001	-553	-352	-572	-333
Medical	587,343	4,804	4,592	611,185	4,669	4,869	-413	-8.6	<0.001	-511	-314	-530	-296
Surgical	158,583	5,141	3,587	137,656	5,048	4,001	-507	-9.9	<0.001	-732	-283	-775	-240

Note: For BPCI Advanced participants in MY4–6 with SNF stays, the cost per diem for SNF stays was \$667 per day. See the first page of this appendix for data sources and methodology.

Exhibit H.6: Impact of BPCI Advanced on IRF Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	1,024	1,098	748,841	1,048	1,501	-378	-36.9	<0.001	-450	-306	-464	-292
Medical	587,343	876	1,066	611,185	930	1,367	-248	-28.3	<0.001	-316	-179	-329	-166
Surgical	158,583	1,656	1,327	137,656	1,585	2,139	-883	-53.3	<0.001	-1,049	-717	-1,081	-685

Note: For BPCI Advanced participants in MY4–6 with IRF stays, the cost per diem for IRF stays was \$1,937 per day. See the first page of this appendix for data sources and methodology.

Exhibit H.7: Impact of BPCI Advanced on HH Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	1,370	1,424	748,841	1,307	1,389	-28	-2.0	0.146	-59	4	-65	10
Medical	587,343	1,283	1,443	611,185	1,234	1,383	11	0.8	0.362	-8	30	-12	33
Surgical	158,583	1,749	1,437	137,656	1,609	1,457	-160	-9.1	0.038	-287	-33	-311	-9

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.8: Impact of BPCI Advanced on Readmission Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)							
Overall	745,926	3,614	4,085	748,841	3,608	4,045	33	0.9	0.221	-11	77	-20	85
Medical	587,343	4,056	4,648	611,185	4,056	4,594	54	1.3	0.074	4	104	-5	113
Surgical	158,583	1,629	1,707	137,656	1,574	1,671	-19	-1.1	0.552	-71	33	-81	43

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.9: Impact of BPCI Advanced on Proportion of Episodes First Discharged to Institutional PAC Setting, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	727,881	28.8	22.7	730,580	27.9	23.4	-1.45	-5.1	<0.001	-2.01	-0.90	-2.12	-0.79
Medical	587,343	26.3	22.6	611,185	25.6	22.5	-0.60	-2.3	0.004	-0.95	-0.26	-1.02	-0.19
Surgical	140,538	41.5	25.8	119,395	39.2	28.3	-4.81	-11.6	<0.001	-6.51	-3.10	-6.84	-2.77

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.10: Impact of BPCI Advanced on Number of Days in SNF Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (days)	DiD as a Percent (%)	P-Value	90% LCI (days)	90% UCI (days)	95% LCI (days)	95% UCI (days)
	MY4–6 Episodes (N)	Baseline Mean (days)	MY4–6 Mean (days)	MY4–6 Episodes (N)	Baseline Mean (days)	MY4–6 Mean (days)							
Overall	179,759	33.6	29.9	172,387	33.8	32.9	-2.85 ‡	-8.5	<0.001	-3.23	-2.47	-3.30	-2.40
Medical	148,146	34.0	30.7	145,565	34.2	33.6	-2.66 ‡	-7.8	<0.001	-3.06	-2.27	-3.14	-2.19
Surgical	31,613	31.5	26.0	26,822	31.6	29.6	-3.50	-11.1	<0.001	-4.20	-2.81	-4.33	-2.67

Note: See the first page of this appendix for data sources and methodology.

3. Model Years 7–8

Exhibit H.11: Impact of BPCI Advanced on Total Episode Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)							
Overall	55,790	27,725	30,481	55,722	26,835	31,191	-1,600	-5.8	<0.001	-2,064	-1,136	-2,152	-1,047
Medical	35,670	26,305	30,725	35,613	25,190	30,433	-822	-3.1	0.002	-1,250	-395	-1,332	-313
Surgical	20,120	30,747	31,082	20,114	30,275	33,423	-2,814	-9.2	<0.001	-3,459	-2,168	-3,583	-2,044

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.12: Impact of BPCI Advanced on SNF Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)							
Overall	55,790	4,937	4,805	55,722	4,890	5,372	-614 ‡	-12.4	<0.001	-881	-346	-933	-294
Medical	35,670	4,647	5,217	35,613	4,619	5,619	-430	-9.2	0.008	-695	-164	-746	-113
Surgical	20,120	5,495	4,072	20,114	5,562	5,079	-940 ‡	-17.1	<0.001	-1,394	-486	-1,481	-399

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.13: Impact of BPCI Advanced on HH Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (\$)	DiD as a Percent (%)	P-Value	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)							
Overall	55,790	1,376	1,307	55,722	1,239	1,307	-137	-9.9	0.003	-212	-62	-226	-48
Medical	35,670	1,142	1,369	35,613	1,061	1,293	-4	-0.4	0.881	-49	41	-57	49
Surgical	20,120	1,916	1,306	20,114	1,606	1,388	-393	-20.5	<0.001	-563	-222	-596	-189

Note: See the first page of this appendix for data sources and methodology.

B. Pooled Parallel Trends

1. Model Years 4–6

Exhibit H.14: Results of Parallel Trends Tests, Payment Outcomes, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Total Episode Payments		Total Paid Payments		SNF Payments		IRF Payments		HH Payments		Readmission Payments	
	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value
Overall	-9	0.174	-7	0.269	-5	0.138	2	0.521	-1	0.243	1	0.650
Medical	-9	0.183	-7	0.294	-6	0.107	3	0.251	-1	0.351	0	0.925
Surgical	-8	0.587	-8	0.557	-1	0.907	-5	0.483	-3	0.345	4	0.230

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.15: Results of Parallel Trends Tests, Utilization Outcomes, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	First Discharge to Institutional PAC Setting		SNF Days	
	Linear Trend Coefficient (pp)	P-Value	Linear Trend Coefficient (days)	P-Value
Overall	-0.01	0.613	-0.04 ‡	0.019
Medical	0.00	0.894	-0.04 ‡	0.024
Surgical	-0.11	0.131	-0.01	0.675

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit H.16: Results of Parallel Trends Tests, Payment Outcomes, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	Total Episode Payments		SNF Payments		HH Payments	
	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value	Linear Trend Coefficient (\$)	P-Value
Overall	-15	0.515	-24 ‡	0.036	1	0.675
Medical	0	0.995	-14	0.243	1	0.689
Surgical	-54	0.147	-55 ‡	0.007	0	0.938

Note: See the first page of this appendix for data sources and methodology.

C. Pooled Impact Estimate and Sensitivity Analyses

1. Model Years 4–6

Exhibit H.17: BPCI Advanced Impact Estimate and Sensitivity Analysis Results, Total Episode Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Result	BPCI Advanced			Comparison Group			DiD (\$)	90% LCI (\$)	90% UCI (\$)
		MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)			
Overall	<i>BPCI Advanced Impact Estimate</i>	745,926	26,716	28,131	748,841	26,418	28,844	-1,011	-1,190	-832
	<i>No BPCI-C Episodes</i>	745,926	26,755	28,176	748,841	26,451	28,883	-1,011	-1,190	-831
	<i>ITT</i>	1,246,214	26,917	29,050	1,249,573	26,635	29,549	-782	-952	-613
Medical	<i>BPCI Advanced Impact Estimate</i>	587,343	25,744	27,849	611,185	25,489	28,350	-756	-923	-588
	<i>No BPCI-C Episodes</i>	587,343	25,762	27,861	611,185	25,491	28,358	-769	-938	-600
	<i>ITT</i>	998,433	25,953	28,749	1,024,920	25,703	29,090	-592	-750	-434
Surgical	<i>BPCI Advanced Impact Estimate</i>	158,583	30,967	29,968	137,656	30,508	31,345	-1,835	-2,199	-1,471
	<i>No BPCI-C Episodes</i>	158,583	31,303	30,354	137,656	30,900	31,731	-1,780	-2,156	-1,403
	<i>ITT</i>	247,793	31,153	30,760	224,653	30,726	31,820	-1,487	-1,817	-1,157

Note: See the first page of this appendix for data sources and methodology.

Exhibit H.18: BPCI Advanced Impact Estimate and National Comparison Sample Sensitivity Analysis Results, Total Episode Payments Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Result	BPCI Advanced			Comparison Group			DiD (\$)	90% LCI (\$)	90% UCI (\$)
		MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)	MY4–6 Episodes (N)	Baseline Mean (\$)	MY4–6 Mean (\$)			
Medical	<i>BPCI Advanced Impact Estimate</i>	587,343	25,744	27,849	611,185	25,489	28,350	-756	-923	-588
	<i>National Comparison Group</i>	669,974	25,738	27,883	3,329,097	25,063	28,404	-1,196 ‡	-1,338	-1,055
Surgical	<i>BPCI Advanced Impact Estimate</i>	158,583	30,967	29,968	137,656	30,508	31,345	-1,835	-2,199	-1,471
	<i>National Comparison Group</i>	161,930	31,396	30,489	1,885,561	30,491	31,983	-2,399 ‡	-2,642	-2,157

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit H.19: BPCI Advanced Impact Estimate and Sensitivity Analysis Results, Total Episode Payments During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	Result	BPCI Advanced			Comparison Group			DiD (\$)	90% LCI (\$)	90% UCI (\$)
		MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)	MY7–8 Episodes (N)	Baseline Mean (\$)	MY7–8 Mean (\$)			
Overall	<i>BPCI Advanced Impact Estimate</i>	55,790	27,725	30,481	55,722	26,835	31,191	-1,600	-2,064	-1,136
	<i>Additional Risk Adjustment</i>	55,790	27,882	29,806	55,722	27,052	30,671	-1,695	-2,129	-1,262
	<i>Alternative National Sample Balancing Weights</i>	55,790	27,901	30,649	55,747	27,197	31,733	-1,788	-2,255	-1,322
	<i>Excluding Past Participants from Comparison Group</i>	55,790	27,641	30,408	55,688	26,745	31,094	-1,582	-2,055	-1,109
Medical	<i>BPCI Advanced Impact Estimate</i>	35,670	26,305	30,725	35,613	25,190	30,433	-822	-1,250	-395
	<i>Additional Risk Adjustment</i>	35,670	26,473	29,770	35,613	25,489	29,747	-961	-1,350	-571
	<i>Alternative National Sample Balancing Weights</i>	35,670	26,407	30,805	35,635	25,563	30,930	-968	-1,401	-535
	<i>Excluding Past Participants from Comparison Group</i>	35,670	26,299	30,734	35,586	25,128	30,462	-899	-1,336	-463
Surgical	<i>BPCI Advanced Impact Estimate</i>	20,120	30,747	31,082	20,114	30,275	33,423	-2,814	-3,459	-2,168
	<i>Additional Risk Adjustment</i>	20,120	30,935	30,416	20,114	30,550	32,873	-2,843	-3,451	-2,235
	<i>Alternative National Sample Balancing Weights</i>	20,120	31,125	31,466	20,130	30,688	34,189	-3,160	-3,810	-2,510
	<i>Excluding Past Participants from Comparison Group</i>	20,120	30,482	30,835	20,111	30,153	33,100	-2,593	-3,260	-1,926

Note: See the first page of this appendix for data sources and methodology.

Appendix I: Medicare Program Savings – Results

The following tables display Medicare savings results. In each exhibit, we report our main results which adjusted reconciliation amounts to account for the model's quality (Composite Quality Score, or CQS, adjustments) and stop-loss/stop-gain limits, as well as results that did not apply these adjustments to reconciliation amounts. Lower and upper bounds of net savings are calculated from the 90% and 95% confidence intervals of the DiD estimate of episode expenditures. Net savings to CMS are reported such that a positive value indicates savings to Medicare and a negative value indicates losses to Medicare. Net reconciliation payments are reported such that a negative value indicates payments made from CMS to participants and a positive value indicates payments made from participants to CMS.

Model Years 4 through 6 (January 1, 2021–December 31, 2023) results are based on clinical episodes with sufficient sample size for evaluation. Model Years 7 and 8 (January 1, 2024–April 2, 2025) results are based on all clinical episodes. Please refer to **Appendix C** for details on the Medicare claims and enrollment data used for each baseline and intervention period, how the methodology differs between the model years, and how these differences may affect savings estimates.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods used to determine and evaluate the sample and data sources used.

A. Net Medicare Savings

1. Model Years 4–6

Exhibit I.1: Net Medicare Savings, BPCI Advanced Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments (\$M)	Reconciliation Payments (\$M)	Savings to Medicare (\$M)	90% LCI (\$M)	90% UCI (\$M)	95% LCI (\$M)	95% UCI (\$M)
All Evaluated Clinical Episodes (n=892,640)	<i>Applied</i>	886.89	28.93	915.82	745.15	1,086.50	712.43	1,119.21
	<i>Not Applied</i>	886.89	43.86	930.76	760.08	1,101.43	727.37	1,134.15
Medical (n=77%)	<i>Applied</i>	483.87	111.77	595.64	480.41	710.86	458.32	732.95
	<i>Not Applied</i>	483.87	129.01	612.88	497.66	728.10	475.57	750.19
Hospitals (n=63%)	<i>Applied</i>	385.13	72.62	457.76	353.84	561.67	333.92	581.60
	<i>Not Applied</i>	385.13	83.32	468.46	364.54	572.37	344.62	592.30
PGPs (n=14%)	<i>Applied</i>	85.17	39.15	124.31	77.60	171.02	68.63	180.00
	<i>Not Applied</i>	85.17	45.69	130.86	84.15	177.57	75.17	186.54
Surgical (n=23%)	<i>Applied</i>	394.80	-82.84	311.97	227.78	396.15	211.63	412.30
	<i>Not Applied</i>	394.80	-85.15	309.65	225.47	393.84	209.32	409.99
Hospitals (n=11%)	<i>Applied</i>	168.22	-4.81	163.41	114.77	212.05	105.43	221.39
	<i>Not Applied</i>	168.22	-4.21	164.01	115.37	212.65	106.03	221.98
PGPs (n=12%)	<i>Applied</i>	211.01	-78.03	132.98	52.90	213.07	37.46	228.50
	<i>Not Applied</i>	211.01	-80.94	130.07	49.99	210.15	34.55	225.59

Note: See the first page of this appendix for data sources and methodology.

Exhibit I.2: Net Medicare Savings, Expressed as a Percentage of the BPCI Advanced Counterfactual, BPCI Advanced Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments (%)	Reconciliation Payments (%)	Savings to Medicare (%)	90% LCI (%)	90% UCI (%)	95% LCI (%)	95% UCI (%)
All Evaluated Clinical Episodes (n=892,640)	<i>Applied</i>	3.5	0.1	3.6	2.9	4.3	2.8	4.4
	<i>Not Applied</i>	3.5	0.2	3.7	3.0	4.3	2.9	4.5
Medical (n=77%)	<i>Applied</i>	2.5	0.6	3.1	2.5	3.7	2.4	3.8
	<i>Not Applied</i>	2.5	0.7	3.2	2.6	3.8	2.5	3.9
Hospitals (n=63%)	<i>Applied</i>	2.5	0.5	2.9	2.3	3.6	2.1	3.7
	<i>Not Applied</i>	2.5	0.5	3.0	2.3	3.6	2.2	3.8
PGPs (n=14%)	<i>Applied</i>	2.5	1.2	3.7	2.3	5.1	2.0	5.4
	<i>Not Applied</i>	2.5	1.4	3.9	2.5	5.3	2.2	5.5
Surgical (n=23%)	<i>Applied</i>	6.1	-1.3	4.9	3.5	6.2	3.3	6.4
	<i>Not Applied</i>	6.1	-1.3	4.8	3.5	6.1	3.3	6.4
Hospitals (n=11%)	<i>Applied</i>	5.1	-0.1	4.9	3.5	6.4	3.2	6.7
	<i>Not Applied</i>	5.1	-0.1	5.0	3.5	6.4	3.2	6.7
PGPs (n=12%)	<i>Applied</i>	7.4	-2.8	4.7	1.9	7.5	1.3	8.1
	<i>Not Applied</i>	7.4	-2.9	4.6	1.8	7.4	1.2	8.0

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit I.3: Net Medicare Savings, BPCI Advanced Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2, 2025)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments (\$M)	Reconciliation Payments (\$M)	Savings to Medicare (\$M)	90% LCI (\$M)	90% UCI (\$M)	95% LCI (\$M)	95% UCI (\$M)
All Clinical Episodes (n=55,188)	<i>Applied</i>	85.58	0.18	85.76	60.96	110.56	56.21	115.31
	<i>Not Applied</i>	85.58	1.47	87.05	62.25	111.85	57.50	116.60
Medical (n=60%)	<i>Applied</i>	26.62	14.15	40.77	26.94	54.61	24.29	57.26
	<i>Not Applied</i>	26.62	15.67	42.29	28.46	56.13	25.81	58.78
Surgical (n=40%)	<i>Applied</i>	59.37	-13.97	45.40	31.77	59.03	29.16	61.64
	<i>Not Applied</i>	59.37	-14.20	45.17	31.54	58.80	28.92	61.41

Note: See the first page of this appendix for data sources and methodology.

Exhibit I.4: Net Medicare Savings, Expressed as a Percentage of the BPCI Advanced Counterfactual, BPCI Advanced Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2, 2025)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments (%)	Reconciliation Payments (%)	Savings to Medicare (%)	90% LCI (%)	90% UCI (%)	95% LCI (%)	95% UCI (%)
All Clinical Episodes (n=55,188)	<i>Applied</i>	5.0	0.0	5.0	3.6	6.4	3.3	6.7
	<i>Not Applied</i>	5.0	0.1	5.1	3.6	6.5	3.4	6.8
Medical (n=60%)	<i>Applied</i>	2.6	1.4	4.0	2.6	5.3	2.4	5.6
	<i>Not Applied</i>	2.6	1.5	4.1	2.8	5.5	2.5	5.8
Surgical (n=40%)	<i>Applied</i>	8.3	-2.0	6.3	4.4	8.3	4.1	8.6
	<i>Not Applied</i>	8.3	-2.0	6.3	4.4	8.2	4.0	8.6

Note: See the first page of this appendix for data sources and methodology.

B. Net Medicare Savings per Episode

1. Model Years 4–6

Exhibit I.5: Net Medicare Savings per Episode, BPCI Advanced Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments per Episode (\$)	Reconciliation Payments per Episode (\$)	Savings to Medicare per Episode (\$)	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
All Evaluated Clinical Episodes (n=892,640)	<i>Applied</i>	994	32	1,026	835	1,217	798	1,254
	<i>Not Applied</i>	994	49	1,043	851	1,234	815	1,271
Medical (n=77%)	<i>Applied</i>	707	163	871	702	1,039	670	1,071
	<i>Not Applied</i>	707	189	896	727	1,064	695	1,097
Hospitals (n=63%)	<i>Applied</i>	689	130	819	633	1,004	597	1,040
	<i>Not Applied</i>	689	149	838	652	1,024	616	1,059
PGPs (n=14%)	<i>Applied</i>	682	313	995	621	1,369	549	1,441
	<i>Not Applied</i>	682	366	1,048	674	1,422	602	1,494

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments per Episode (\$)	Reconciliation Payments per Episode (\$)	Savings to Medicare per Episode (\$)	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
Surgical (n=23%)	<i>Applied</i>	1,893	-397	1,496	1,092	1,900	1,015	1,977
	<i>Not Applied</i>	1,893	-408	1,485	1,081	1,889	1,004	1,966
Hospitals (n=11%)	<i>Applied</i>	1,705	-49	1,656	1,163	2,149	1,069	2,244
	<i>Not Applied</i>	1,705	-43	1,662	1,169	2,155	1,075	2,250
PGPs (n=12%)	<i>Applied</i>	1,921	-710	1,210	482	1,939	341	2,080
	<i>Not Applied</i>	1,921	-737	1,184	455	1,913	315	2,053

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit I.6: Net Medicare Savings per Episode, BPCI Advanced Hospitals and PGPs, Model Years 7–8 (Jan. 2024– Apr. 2, 2025)

Clinical Episode Type	Reconciliation CQS and Stop-Loss/Gain	Reduction in Non-Standardized Payments per Episode (\$)	Reconciliation Payments per Episode (\$)	Savings to Medicare per Episode (\$)	90% LCI (\$)	90% UCI (\$)	95% LCI (\$)	95% UCI (\$)
All Clinical Episodes (n=55,188)	<i>Applied</i>	1,551	3	1,554	1,105	2,003	1,018	2,089
	<i>Not Applied</i>	1,551	27	1,577	1,128	2,027	1,042	2,113
Medical (n=60%)	<i>Applied</i>	804	428	1,232	814	1,650	734	1,730
	<i>Not Applied</i>	804	474	1,278	860	1,696	780	1,776
Surgical (n=40%)	<i>Applied</i>	2,687	-632	2,054	1,438	2,671	1,319	2,790
	<i>Not Applied</i>	2,687	-643	2,044	1,427	2,661	1,309	2,779

Note: See the first page of this appendix for data sources and methodology.

Appendix J: Participant Reconciliation – Results

This appendix provides details of participant reconciliation results. Exhibit J.1 provides total reconciliation earned by participants, broken down by time period. Exhibit J.2 provides similar details about per-episode reconciliation. Per-episode reconciliation is broken down by medical and surgical episodes. Note that these statistics are computed from participants with at least one reconciled episode. A participant that had no medical volume would not be counted in the medical statistics, even if the participant had surgical volume, and likewise a participant that had no surgical volume would not be counted in surgical statistics.

Exhibits J.3 and J.4 present percentiles of total net payment reconciliation and repayment amounts, expressed as a share of key hospital financial measures from the CMS Hospital Cost Reports, including total revenue, operating revenue, operating expenses, and net operating income (defined as operating revenue minus operating expenses). Note that while many hospitals are nonprofit or government-owned, as shorthand we refer to net operating income as “operating profit/loss.” All percentages are expressed as magnitudes, meaning that a participant that earned \$30,000 in Net Payment Reconciliation Amount and sustained an operating loss of \$3,000,000 would have reconciliation equal to 1% of operating profit/loss that year.

In Exhibits J.5–J.7, we provide descriptive statistics of Composite Quality Score (CQS). Exhibit J.5 provides descriptive statistics of the CQS by model year. Exhibit J.6 provides the average reconciliation payments that hospitals and physician group practices (PGPs) actually earned; the average reconciliation payments that they would have earned had they received the worst possible CQS of 0, the best possible CQS of 100, a CQS at the 5th percentile, or a CQS at the 95th percentile; and the average reconciliation payments that they would have earned given various hypothetical changes in per-episode spending. For example, Exhibit J.6 provides the average reconciliation payment that hospitals and PGPs would have received had their per-episode spending been 0.5% lower or 0.5% higher. Exhibit J.7 provides descriptive statistics on the difference between actual reconciliation payments and various hypothetical reconciliation payments. For example, the exhibit contains the mean and standard deviation of the difference between what participants earned and what they would have earned given a perfect CQS, or between what they earned and what they would have earned had they reduced spending by 1% relative to their observed spending.

In Exhibit J.8, we report descriptive statistics of per-episode reconciliation payments of safety-net hospitals versus non-safety-net hospitals.

Finally, in Exhibit J.9, we provide information about the stop-loss and stop-gain limits. In particular, for each model year and for the model overall, we provide the share of participants that hit the stop-loss limit in at least one performance period; the total additional amount that participants avoided losing as a result of the stop-loss limit; the share of participants that hit their stop-gain limit in at least one performance period; and the total additional amount that participants did not earn as a result of the stop-gain limit.

In all exhibits, positive dollar values reflect payments made to participants from CMS, while negative dollar values reflect amounts owed to CMS by participants.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods used to determine the sample, the measures evaluated, and data sources.

A. General Descriptive Statistics of Reconciliation Results

Exhibit J.1: Distribution of Total Reconciliation Payments, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Apr. 2, 2025)

Time Period	Percentile (\$)					Mean (\$)	Standard Deviation (\$)
	10 th	25 th	50 th	75 th	90 th		
MY1–3	-166,923	-1,940	197,341	853,396	2,164,476	722,310	1,816,211
MY4–6	-918,975	-336,600	-4,251	329,842	849,249	-33,691	1,040,238
MY7–8	-683,196	-231,874	9,304	185,609	599,026	-5,778	586,797
Overall	-415,731	-64,602	146,434	859,706	2,351,703	665,030	1,950,547

Note: See the first page of this appendix for data sources and methodology.

Exhibit J.2: Distribution of Per-Episode Reconciliation Payments, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Apr. 2, 2025)

Time Period	Grouping	Percentile (\$)					Mean (\$)	Standard Deviation (\$)
		10 th	25 th	50 th	75 th	90 th		
MY1–3	Combined	-1,475	-32	1,232	2,709	4,211	1,354	2,540
	Medical	-1,459	85	1,437	2,909	4,421	1,403	2,899
	Surgical	-2,394	-679	792	2,739	5,335	1,128	3,955
MY4–6	Combined	-2,478	-1,182	-20	1,199	2,764	88	2,241
	Medical	-2,555	-1,359	-328	893	2,411	-152	2,171
	Surgical	-2,838	-1,008	388	2,028	3,641	516	3,601
MY7–8	Combined	-3,065	-1,547	58	1,647	3,025	145	2,663
	Medical	-3,223	-1,946	-300	1,010	2,947	-278	2,704
	Surgical	-3,646	-781	823	2,296	3,904	567	3,082

Note: See the first page of this appendix for data sources and methodology.

Exhibit J.3: Reconciliation Payments as a Percentage of Hospital Finances, Among Hospitals that Earned a Reconciliation Payment, Model Years 1–5 (Oct. 2018–Dec. 2022)

Model Year	Payments as a Percentage of:	Percentiles (%)						Mean (%)
		50 th	75 th	80 th	85 th	90 th	95 th	
MY1&2	Total Revenue	0.1	0.3	0.4	0.5	0.7	0.9	0.3
	Operating Revenue	0.1	0.4	0.4	0.6	0.7	0.9	0.3
	Operating Expense	0.2	0.4	0.5	0.6	0.7	1.0	0.3
	Operating Profit/Loss	2.0	6.3	8.5	12.1	20.3	50.8	11.3

Model Year	Payments as a Percentage of:	Percentiles (%)						Mean (%)
		50 th	75 th	80 th	85 th	90 th	95 th	
MY3	Total Revenue	0.1	0.3	0.3	0.4	0.6	0.8	0.2
	Operating Revenue	0.1	0.3	0.4	0.4	0.6	0.8	0.2
	Operating Expense	0.1	0.3	0.4	0.5	0.6	0.9	0.2
	Operating Profit/Loss	1.4	4.9	6.3	9.6	14.1	27.5	12.3
MY4	Total Revenue	0.1	0.2	0.3	0.3	0.4	0.6	0.2
	Operating Revenue	0.1	0.3	0.3	0.4	0.5	0.6	0.2
	Operating Expense	0.1	0.3	0.3	0.4	0.5	0.6	0.2
	Operating Profit/Loss	1.3	3.8	4.7	6.2	10.6	20.7	4.4
MY5	Total Revenue	0.1	0.2	0.2	0.3	0.4	0.5	0.2
	Operating Revenue	0.1	0.2	0.2	0.3	0.4	0.6	0.2
	Operating Expense	0.1	0.2	0.2	0.3	0.4	0.5	0.2
	Operating Profit/Loss	1.0	2.7	3.7	5.0	8.0	14.9	3.6
Overall	Total Revenue	0.1	0.3	0.3	0.4	0.6	0.7	0.2
	Operating Revenue	0.1	0.3	0.4	0.4	0.6	0.8	0.2
	Operating Expense	0.1	0.3	0.4	0.5	0.6	0.8	0.2
	Operating Profit/Loss	1.5	4.6	6.1	9.0	14.1	27.5	9.6

Note: See the first page of this appendix for data sources and methodology.

Exhibit J.4: Reconciliation Payments as a Percentage of Hospital Finances, Among Hospitals that Owed a Repayment, Model Years 1–5 (Oct. 2018–Dec. 2022)

Model Year	Payments Relative to:	Percentiles (%)						Mean (%)
		50 th	75 th	80 th	85 th	90 th	95 th	
MY1&2	Total Revenue	0.0	0.1	0.1	0.2	0.2	0.3	0.1
	Operating Revenue	0.1	0.1	0.2	0.2	0.2	0.3	0.1
	Operating Expense	0.1	0.1	0.2	0.2	0.2	0.4	0.1
	Operating Profit/Loss	0.8	2.1	2.6	3.7	5.4	10.3	2.5
MY3	Total Revenue	0.0	0.1	0.1	0.1	0.1	0.2	0.0
	Operating Revenue	0.0	0.1	0.1	0.1	0.1	0.2	0.0
	Operating Expense	0.0	0.1	0.1	0.1	0.1	0.2	0.0
	Operating Profit/Loss	0.3	1.0	1.5	2.5	3.9	6.6	1.4
MY4	Total Revenue	0.1	0.1	0.2	0.2	0.3	0.4	0.1
	Operating Revenue	0.1	0.1	0.2	0.2	0.3	0.4	0.1
	Operating Expense	0.1	0.2	0.2	0.2	0.3	0.4	0.1
	Operating Profit/Loss	1.0	3.4	4.5	6.2	8.5	21.2	8.3
MY5	Total Revenue	0.1	0.1	0.2	0.2	0.3	0.4	0.1
	Operating Revenue	0.1	0.2	0.2	0.2	0.3	0.4	0.1
	Operating Expense	0.1	0.1	0.2	0.3	0.3	0.4	0.1
	Operating Profit/Loss	1.1	2.2	2.9	4.2	5.8	15.4	6.9
Overall	Total Revenue	0.1	0.1	0.1	0.2	0.2	0.3	0.1
	Operating Revenue	0.1	0.1	0.1	0.2	0.2	0.4	0.1
	Operating Expense	0.1	0.1	0.2	0.2	0.3	0.4	0.1
	Operating Profit/Loss	0.7	2.3	3.0	4.3	6.2	12.4	5.2

Note: See the first page of this appendix for data sources and methodology.

B. Descriptive Statistics and Effects of the CQS

Exhibit J.5: CQS of BPCI Advanced Hospitals and PGPs, Model Years 1–7
(Oct. 2018–Dec. 2024)

Model Year	Percentile					Mean	Standard Deviation
	5 th	25 th	50 th	75 th	95 th		
MY1&2	29.5	48.0	60.2	69.7	83.3	58.5	16.2
MY3	20.5	37.2	50.0	61.1	76.8	49.2	16.9
MY4	26.8	43.9	55.0	66.3	83.5	55.1	16.9
MY5	34.0	50.9	60.4	69.5	83.9	60.1	14.4
MY6	22.7	41.3	55.2	65.0	79.7	53.1	16.9
MY7	20.0	34.0	47.8	57.5	77.7	46.9	16.9

Note: See the first page of this appendix for data sources and methodology.

Exhibit J.6: Reconciliation Payments, Assuming Hypothetical CQS and Spending Reductions, Hospitals and PGPs, Model Years 1–7 (Oct. 2018–Dec. 2024)

Reconciliation		Total (\$)		Per-Episode (\$)	
		Mean	Standard Deviation	Mean	Standard Deviation
Actual	Actual Reconciliation	662,217	1,943,656	827	2,283
Hypothetical, Given CQS of...	0 (Worst Possible)	578,919	1,840,495	718	2,263
	100 (Best Possible)	730,224	2,020,029	919	2,295
	5th Percentile	619,132	1,886,709	771	2,273
	95th Percentile	702,046	1,985,424	881	2,290
Hypothetical, Given Payment Reduction of...	1% of Actual Payments	910,449	2,247,947	1,086	2,277
	0.5% of Actual Payments	786,322	2,088,921	956	2,281
	0.25% of Actual Payments	724,265	2,014,394	891	2,282
Hypothetical, Given Payment Increase of...	0.25% of Actual Payments	600,172	1,877,193	762	2,284
	0.5% of Actual Payments	538,138	1,815,522	697	2,284
	1% of Actual Payments	414,123	1,708,482	567	2,285

Note: See the first page of this appendix for data sources and methodology.

Exhibit J.7: Differences of Reconciliation Payments, Assuming Hypothetical CQS and Spending Reductions, Hospitals and PGPs, Model Years 1–7
(Oct. 2018–Dec. 2024)

Reconciliation difference		Total (\$)		Per-Episode (\$)	
		Mean	Standard Deviation	Mean	Standard Deviation
CQS	0 vs. Actual	-83,298	138,579	-109	73
	100 vs. Actual	68,007	110,820	93	71
	100 vs. 0	151,305	238,585	201	122
	95th vs. 5th Percentile	82,914	130,610	111	69
Payment Reduction	1% Lower vs. Actual	248,232	457,808	260	91
	0.5% Lower vs. Actual	124,105	228,928	130	46
	0.25% Lower vs. Actual	62,048	114,480	65	23

Reconciliation difference		Total (\$)		Per-Episode (\$)	
		Mean	Standard Deviation	Mean	Standard Deviation
Payment Increase	0.25% Higher vs. Actual	-62,045	114,473	-65	23
	0.5% Higher vs. Actual	-124,079	228,931	-130	46
	1% Higher vs. Actual	-248,093	457,724	-259	91

Note: See the first page of this appendix for data sources and methodology.

C. Reconciliation Results of Safety-Net Hospitals

Exhibit J.8: Per-Episode Reconciliation Payments of Safety-Net and Non-Safety-Net Hospitals, Model Years 1–8 (Oct. 2018–Apr. 2, 2025)

Model Year	Safety-Net Status	Hospitals (N)	25 th Percentile (\$)	50 th Percentile (\$)	75 th Percentile (\$)
MY1&2	Non-Safety-Net	569	-165	965	2,353
	Safety-Net	138	-76	1,534	3,313
MY3	Non-Safety-Net	797	-85	1,415	2,991
	Safety-Net	169	216	2,431	4,673
MY4–6	Non-Safety-Net	546	-1,186	-215	769
	Safety-Net	131	-1,754	-167	1,530
MY7–8	Non-Safety-Net	95	-1,657	-470	736
	Safety-Net	27	-2,174	-365	1,730

Note: See the first page of this appendix for data sources and methodology.

D. Stop-loss and Stop-gain Limits

Exhibit J.9: Effects of the Stop-Loss and Stop-Gain Limits, Hospitals and PGPs, Model Years 1–8 (Oct. 2018–Apr. 2, 2025)

Model Year	Stop-Loss Limit		Stop-Gain Limit	
	Share of Providers Affected (%)	Total Impact on Payments (\$)	Share of Providers Affected (%)	Total Impact on Payments (\$)
MY1&2	8	5,938,521	13	-12,188,583
MY3	12	7,070,356	25	-16,274,615
MY4	12	6,907,098	11	-3,109,532
MY5	15	5,285,427	12	-2,715,784
MY6	6	677,628	8	-211,406
MY7	20	2,602,634	14	-967,851
MY8	6	510,113	9	-689,089
Overall	25	28,991,777	32	-36,156,861

Note: See the first page of this appendix for data sources and methodology.

Appendix K: Claims-Based Quality Outcomes – Results

The exhibits in Section A present the risk-adjusted difference-in-differences (DiD) results for the claims-based quality outcomes assessed in this report. For Model Years 4 through 6, we present the readmission rate through the 90-day post-discharge period (PDP), mortality rate during the anchor stay and 90-day PDP, and proportion of episodes with one or more emergency department (ED) visits during the 90-day PDP for both hospitals and physician group practices (PGPs). For Model Years 7 and 8, we present the mortality rate during the 90-day PDP. The “DiD as a Percent (%)” reflects the DiD estimate as a percentage of the BPCI Advanced baseline mean.

The exhibits in Section B present parallel trends test results that correspond to the impact estimates for the quality and related measures assessed in the report. As described in **Appendix C**, we tested the null hypothesis that BPCI Advanced participants and comparison providers had parallel trends during the baseline period at the 10% level of statistical significance. We use the dagger symbol (‡) to denote estimates for which we rejected the null hypothesis of parallel trends, indicating a parallel trends failure for that result. A failure of the parallel trends test indicates that BPCI Advanced participants and matched comparison providers did not have the same baseline trends for that outcome; therefore, the impact estimate may partially reflect a continuation of preexisting trends rather than solely the impact of the BPCI Advanced Model.

The exhibits in Section C present sensitivity analyses that evaluate the robustness of BPCI Advanced impact estimates to alternative episode samples and analytic specifications. For Model Years 4 through 6, these sensitivity tests include estimating impacts after excluding episodes initiated by providers participating in the prior BPCI initiative (“No BPCI-C Episodes”), as well as intent-to-treat (ITT) estimates that include episodes for all matched BPCI Advanced providers and their matched comparisons regardless of subsequent model withdrawal. For Model Years 7 and 8, these include estimates from a regression that included additional risk adjusters, such as a variety of Hierarchical Condition Category flags (“additional risk adjustment”), estimates utilizing alternative national sample balancing weights, and estimates that excluded all participants that participated in any clinical episodes from the national comparison group (“excluding past participants from comparison group”).

These exhibits display pooled impact estimate results based on clinical episodes with sufficient sample size for evaluation for Model Years 4 through 6 (January 1, 2021–December 31, 2023), and results are based on all clinical episodes for Model Years 7 and 8 (January 1, 2024–April 30, 2025). All results are based on the BPCI Advanced evaluation team’s analysis of Medicare claims and enrollment data for episodes attributed to BPCI Advanced and comparison hospitals and PGPs and matched comparison providers. Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods.

A. Pooled Impact Estimates

1. Model Years 4–6

Exhibit K.1: Impact of BPCI Advanced on Mortality Rate During the Anchor Stay and 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	758,759	15.5	16.0	762,674	15.2	15.9	-0.13	-0.9	0.347	-0.37	0.10	-0.41	0.15
Medical	612,336	17.8	18.5	636,514	17.5	18.3	-0.13	-0.7	0.431	-0.41	0.14	-0.46	0.20
Surgical	146,423	3.6	3.5	126,160	3.6	3.7	-0.13	-3.7	0.150	-0.29	0.02	-0.32	0.05

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.2: Impact of BPCI Advanced on Mortality Rate During the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	719,530	11.9	12.6	723,770	11.8	12.5	0.03	0.3	0.767	-0.14	0.21	-0.18	0.24
Medical	573,980	13.8	14.6	598,441	13.7	14.4	0.07	0.5	0.594	-0.14	0.27	-0.18	0.31
Surgical	145,550	2.9	3.0	125,329	2.9	3.1	-0.12	-4.2	0.146	-0.26	0.02	-0.29	0.04

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.3: Impact of BPCI Advanced on Readmission Rate Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	738,983	27.7	27.3	742,701	27.4	26.9	0.06	0.2	0.612	-0.14	0.27	-0.18	0.31
Medical	580,504	30.9	30.7	605,128	30.7	30.3	0.13	0.4	0.341	-0.10	0.36	-0.14	0.41
Surgical	158,479	13.4	12.4	137,573	12.8	11.9	-0.19	-1.4	0.324	-0.51	0.13	-0.57	0.19

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.4: Impact of BPCI Advanced on Unplanned Readmission Rate Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	738,983	26.4	25.9	742,701	26.1	25.6	0.07	0.3	0.552	-0.13	0.28	-0.17	0.32
Medical	580,504	29.6	29.4	605,128	29.4	29.0	0.15	0.5	0.296	-0.08	0.38	-0.13	0.42
Surgical	158,479	12.2	11.0	137,573	11.6	10.6	-0.19	-1.6	0.247	-0.47	0.08	-0.52	0.13

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.5: Impact of BPCI Advanced on Proportion of Episodes With One or More ED Visits During the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)							
Overall	738,983	22.0	21.9	742,701	22.6	22.3	0.23 ‡	1.0	0.097	0.00	0.46	-0.04	0.50
Medical	580,504	23.2	23.2	605,128	23.9	23.6	0.35 ‡	1.5	0.023	0.10	0.61	0.05	0.66
Surgical	158,479	16.7	15.9	137,573	16.9	16.3	-0.11	-0.6	0.620	-0.46	0.25	-0.53	0.31

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit K.6: Impact of BPCI Advanced on Mortality Rate During the 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	BPCI Advanced			Comparison			DiD (pp)	DiD as a Percent (%)	P-Value	90% LCI (pp)	90% UCI (pp)	95% LCI (pp)	95% UCI (pp)
	MY7–8 Episodes (N)	Baseline Mean (pp)	MY7–8 Mean (pp)	MY7–8 Episodes (N)	Baseline Mean (pp)	MY7–8 Mean (pp)							
Overall	53,599	10.9	9.5	53,523	11.4	9.9	0.14	1.3	0.541	-0.24	0.52	-0.31	0.59
Medical	35,670	14.0	12.2	35,613	14.6	12.6	0.30	2.2	0.329	-0.21	0.82	-0.31	0.91
Surgical	17,929	3.3	3.0	17,916	3.5	3.4	-0.19	-5.8	0.273	-0.48	0.10	-0.53	0.15

Note: See the first page of this appendix for data sources and methodology.

B. Pooled Parallel Trends

1. Model Years 4–6

Exhibit K.7: Results of Parallel Trends Tests, Mortality and ED Visit Outcomes, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Mortality Rate During Anchor Stay and 90-Day PDP		Mortality Rate During 90-Day PDP		ED Visit	
	Linear Trend Coefficient (pp)	P-Value	Linear Trend Coefficient (pp)	P-Value	Linear Trend Coefficient (pp)	P-Value
Overall	0.00	0.998	0.00	0.870	0.05 ‡	<0.001
Medical	0.00	0.733	0.00	0.827	0.05 ‡	<0.001
Surgical	-0.01	0.196	-0.01	0.241	0.02	0.256

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.8: Results of Parallel Trends Tests, Readmission Outcomes, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Readmission Rate		Unplanned Readmission Rate	
	Linear Trend Coefficient (pp)	P-Value	Linear Trend Coefficient (pp)	P-Value
Overall	0.01	0.401	0.01	0.249
Medical	0.01	0.468	0.01	0.279
Surgical	0.01	0.485	0.01	0.551

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit K.9: Results of Parallel Trends Tests, Mortality Outcome, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	Mortality Rate During 90-Day PDP	
	Linear Trend Coefficient (pp)	P-Value
Overall	-0.03	0.289
Medical	-0.04	0.197
Surgical	0.00	0.942

Note: See the first page of this appendix for data sources and methodology.

C. Pooled Impact Estimate and Sensitivity Analyses

1. Model Years 4–6

Exhibit K.10: BPCI Advanced Impact Estimate and Sensitivity Test Results, Mortality Rate During the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Result	BPCI Advanced			Comparison			DiD (pp)	90% LCI (pp)	90% UCI (pp)
		MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)			
Overall	<i>BPCI Advanced Impact Estimate</i>	719,530	11.9	12.6	723,770	11.8	12.5	0.03	-0.14	0.21
	<i>No BPCI-C Episodes</i>	719,530	12.0	12.7	723,770	11.9	12.6	0.02	-0.16	0.20
	<i>ITT</i>	1,202,923	12.1	12.3	1,208,531	12.0	12.2	0.05	-0.12	0.21
Medical	<i>BPCI Advanced Impact Estimate</i>	573,980	13.8	14.6	598,441	13.7	14.4	0.07	-0.14	0.27
	<i>No BPCI-C Episodes</i>	573,980	13.7	14.6	598,441	13.6	14.4	0.05	-0.16	0.25
	<i>ITT</i>	975,840	14.0	14.3	1,003,941	13.9	14.1	0.07	-0.13	0.26
Surgical	<i>BPCI Advanced Impact Estimate</i>	145,550	2.9	3.0	125,329	2.9	3.1	-0.12	-0.26	0.02
	<i>No BPCI-C Episodes</i>	145,550	3.0	3.1	125,329	3.1	3.2	-0.11	-0.26	0.03
	<i>ITT</i>	227,083	2.9	3.0	204,590	2.9	3.0	-0.04	-0.16	0.07

Note: See the first page of this appendix for data sources and methodology.

Exhibit K.11: BPCI Advanced Impact Estimate and Sensitivity Test Results, Readmission Rate Through the 90-Day PDP, Hospitals and PGPs, Model Years 4–6 (Jan. 2021–Dec. 2023)

Clinical Episode Type	Result	BPCI Advanced			Comparison			DiD (pp)	90% LCI (pp)	90% UCI (pp)
		MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)	MY4–6 Episodes (N)	Baseline Mean (pp)	MY4–6 Mean (pp)			
Overall	<i>BPCI Advanced Impact Estimate</i>	738,983	27.7	27.3	742,701	27.4	26.9	0.06	-0.14	0.27
	<i>No BPCI-C Episodes</i>	738,983	27.8	27.4	742,701	27.6	27.1	0.08	-0.13	0.28
	<i>ITT</i>	1,234,812	27.7	27.3	1,239,732	27.4	26.9	0.12	-0.07	0.31
Medical	<i>BPCI Advanced Impact Estimate</i>	580,504	30.9	30.7	605,128	30.7	30.3	0.13	-0.10	0.36
	<i>No BPCI-C Episodes</i>	580,504	30.9	30.7	605,128	30.7	30.3	0.13	-0.10	0.37
	<i>ITT</i>	987,188	30.9	30.7	1,015,206	30.7	30.3	0.23	0.01	0.44
Surgical	<i>BPCI Advanced Impact Estimate</i>	158,479	13.4	12.4	137,573	12.8	11.9	-0.19	-0.51	0.13
	<i>No BPCI-C Episodes</i>	158,479	13.7	12.7	137,573	13.1	12.2	-0.15	-0.48	0.18
	<i>ITT</i>	247,624	13.4	12.4	224,526	12.8	12.1	-0.29	-0.57	-0.01

Note: See the first page of this appendix for data sources and methodology.

2. Model Years 7–8

Exhibit K.12: BPCI Advanced Impact Estimate and Sensitivity Test Results, Mortality Rate During the 90-Day PDP, Hospitals and PGPs, Model Years 7–8 (Jan. 2024–Apr. 2025)

Clinical Episode Type	Result	BPCI Advanced			Comparison			DiD (pp)	90% LCI (pp)	90% UCI (pp)
		MY7–8 Episodes (N)	Baseline Mean (pp)	MY7–8 Mean (pp)	MY7–8 Episodes (N)	Baseline Mean (pp)	MY7–8 Mean (pp)			
Overall	<i>BPCI Advanced Impact Estimate</i>	53,599	10.9	9.5	53,523	11.4	9.9	0.14	-0.24	0.52
	<i>Additional Risk Adjustment</i>	53,599	10.9	9.3	53,523	11.5	9.7	0.15	-0.23	0.54
	<i>Alternative National Sample Balancing Weights</i>	53,599	10.9	9.6	53,542	11.4	9.8	0.21	-0.17	0.60
	<i>Excluding Past Participants from Comparison Group</i>	53,598	10.9	9.5	53,487	11.4	9.9	0.10	-0.29	0.48
Medical	<i>BPCI Advanced Impact Estimate</i>	35,670	14.0	12.2	35,613	14.6	12.6	0.30	-0.21	0.82
	<i>Additional Risk Adjustment</i>	35,670	14.0	12.0	35,613	14.7	12.3	0.32	-0.19	0.84
	<i>Alternative National Sample Balancing Weights</i>	35,670	14.0	12.2	35,635	14.6	12.4	0.40	-0.12	0.92
	<i>Excluding Past Participants from Comparison Group</i>	35,670	14.0	12.3	35,586	14.6	12.6	0.27	-0.26	0.79
Surgical	<i>BPCI Advanced Impact Estimate</i>	17,929	3.3	3.0	17,916	3.5	3.4	-0.19	-0.48	0.10
	<i>Additional Risk Adjustment</i>	17,929	3.4	2.9	17,916	3.6	3.2	-0.17	-0.45	0.11
	<i>Alternative National Sample Balancing Weights</i>	17,929	3.5	3.2	17,922	3.7	3.5	-0.16	-0.46	0.14
	<i>Excluding Past Participants from Comparison Group</i>	17,928	3.3	3.0	17,911	3.4	3.3	-0.23	-0.52	0.06

Note: See the first page of this appendix for data sources and methodology.

Appendix L: Patient Survey – Results

The exhibits in Section A present results for the closed-ended items of the patient survey, including self-reported changes in functional status, care experiences, and satisfaction for BPCI Advanced and comparison patients with episodes that began between April 1, 2024, and January 31, 2025. The closed-ended questions cover two domains: functional status and patient experience and satisfaction.

Exhibit L.1 presents results from domain-level joint hypothesis tests assessing whether, taken together, responses within each domain differed significantly between BPCI Advanced and comparison patients. These joint hypothesis tests can result in statistical significance even if individual measures are in different directions: that is, BPCI Advanced patients responded more favorably than comparison patients on some measures, but more negatively on others. The evaluation designated the domain as having “room for improvement” if BPCI Advanced patients responded less favorably than comparison patients, “improvement” if BPCI Advanced patients responded more favorably than comparison patients, and “varied” if the direction of association varied. We included these results in our interpretation of findings.

Exhibits L.2–L.4 compare BPCI Advanced and comparison patients on individual survey measures within these domains. Estimates are based on cross-sectional logistic regression models for binary or trinary indicators, controlling for patient, hospital, and neighborhood characteristics. Responses were weighted for non-response and sampling design, and results are reported in percentage points. Findings are shown for all 34 clinical episodes combined and separately for medical and surgical episodes, pooled across hospitals and physician group practices (PGPs). Lower confidence interval (LCI) and upper confidence interval (UCI) are displayed for the 5% and 10% significance levels, and p-values are reported; for trinary functional status measures, p-values reflect joint significance across multiple outcomes of the measure.

Section B presents findings from open-ended survey responses. We used generative artificial intelligence (AI) to conduct a thematic analysis and identify 34 thematic concepts related to patient experience. Exhibit L.5 shows open-ended response rates by group and episode type, and Exhibit L.6 describes the identified thematic concepts. Each response could be assigned one or more concepts and coded as having negative, neutral, or positive sentiment. Exhibit L.7 shows the number of thematic concepts assigned per response. Exhibit L.8 reports the number and percent of responses with each thematic concept, along with sentiment distributions by group. Cells marked with an asterisk (*) are suppressed because the sample size is fewer than 11. Exhibit L.9 shows the 15 thematic concept pairs with the greatest overlap in responses, based on correlation coefficients.

Please refer to **Appendix A** for the definitions of all acronyms used within this appendix. See **Appendix C** for additional information on the methods

A. Closed-Ended Survey Response Results

Exhibit L.1: Domain-Level Joint Statistical Significance Tests for the Patient Survey Outcomes, All Clinical Episodes, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Episode Type	Functional Status Domain		Care Experience and Satisfaction Domain	
	Summary	P-Value	Summary	P-Value
Overall	Varied	0.038	Room for improvement	0.026
Medical	Room for improvement	0.093	Room for improvement	0.011
Surgical	Varied	0.486	Room for improvement	0.361

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.2: Patient Survey Outcomes, All Clinical Episodes, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status	<i>Bathing, dressing, using the toilet or eating</i>	Improvement	63.2	65.6	-2.4	-4.1	-0.7	-3.8	-1.0	0.006
		Maintained	13.8	14.0	-0.2	-1.5	1.1	-1.3	0.9	
		Declined	23.0	20.4	2.6	0.9	4.3	1.2	4.0	
	<i>Planning regular tasks</i>	Improvement	65.8	67.3	-1.5	-3.1	0.1	-2.8	-0.2	0.124
		Maintained	11.3	11.4	-0.1	-1.3	1.1	-1.1	0.9	
		Declined	22.9	21.3	1.6	0.0	3.3	0.2	3.0	
	<i>Use of mobility device</i>	Improvement	44.7	44.7	0.0	-1.5	1.5	-1.3	1.3	0.491
		Maintained	12.9	12.2	0.7	-0.5	1.8	-0.3	1.6	
		Declined	42.5	43.1	-0.6	-2.2	0.9	-1.9	0.7	
	<i>Walking without rest</i>	Improvement	39.0	40.1	-1.2	-2.9	0.6	-2.6	0.3	0.252
		Maintained	21.7	21.7	0.0	-1.5	1.4	-1.2	1.2	
		Declined	39.3	38.1	1.2	-0.3	2.7	0.0	2.4	
	<i>Going up or down stairs</i>	Improvement	38.3	37.5	0.8	-0.8	2.5	-0.6	2.2	0.483
		Maintained	24.0	24.9	-0.9	-2.5	0.7	-2.3	0.4	
		Declined	37.7	37.6	0.1	-1.4	1.6	-1.2	1.4	

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status (Continued)	<i>Physical/emotional problems limiting social activities</i>	Improvement	51.8	52.0	-0.2	-2.2	1.8	-1.9	1.5	0.658
		Maintained	21.8	22.4	-0.6	-2.3	1.2	-2.0	0.9	
		Declined	26.4	25.7	0.8	-1.0	2.5	-0.7	2.2	
	<i>Pain limiting regular activities</i>	Improvement	51.1	50.5	0.6	-1.3	2.6	-1.0	2.2	0.329
		Maintained	26.4	27.8	-1.4	-3.2	0.5	-2.9	0.2	
		Declined	22.5	21.7	0.8	-1.1	2.6	-0.8	2.3	
Care Experience	<i>Felt prepared to leave the hospital</i>	Very or Somewhat	91.3	92.1	-0.8	-2.1	0.6	-1.9	0.3	0.251
	<i>Medical staff took your preferences into account in deciding what health care services you should have after you left the hospital</i>	Strongly Agree or Agree	86.9	89.8	-2.9	-4.4	-1.3	-4.2	-1.6	0.000
	<i>Good understanding of how to take care of self before going home</i>	Strongly Agree or Agree	91.6	92.5	-1.0	-2.1	0.2	-2.0	0.0	0.112
	<i>Medical staff clearly explained how to take medications before going home</i>	Strongly Agree or Agree	90.3	91.4	-1.1	-2.4	0.3	-2.2	0.0	0.113
	<i>Medical staff clearly explained what follow-up appointments or treatments would be needed before going home</i>	Strongly Agree or Agree	91.2	92.5	-1.3	-2.6	0.0	-2.4	-0.2	0.045
	<i>Able to manage your health needs since returning home</i>	Strongly Agree or Agree	94.1	94.7	-0.6	-1.6	0.4	-1.4	0.2	0.248
	<i>Medical staff talked with you about whether you would have the help you needed when you got home</i>	Yes	87.5	88.5	-0.9	-2.5	0.6	-2.2	0.3	0.226

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Care Experience (Continued)	<i>If you needed help at home to manage your health, medical staff arranged services for you at home to help manage your health</i>	Yes	71.5	71.3	0.3	-2.4	2.9	-2.0	2.5	0.853
	<i>Overall satisfaction with recovery</i>	Quite a Bit or Extreme	61.8	62.6	-0.8	-2.8	1.2	-2.5	0.9	0.432
Satisfaction With Care	<i>Rating of all care received after leaving the hospital</i>	9-10	61.0	61.0	0.1	-2.1	2.2	-1.7	1.9	0.410
		7-8	23.4	24.4	-1.0	-2.8	0.9	-2.5	0.6	
		0-6	15.6	14.7	0.9	-0.8	2.6	-0.5	2.3	

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.3: Patient Survey Outcomes, Medical Episodes, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status	<i>Bathing, dressing, using the toilet or eating</i>	Improvement	55.8	59.2	-3.4	-5.6	-1.2	-5.3	-1.6	0.003
		Maintained	17.0	17.3	-0.3	-2.1	1.5	-1.8	1.2	
		Declined	27.2	23.4	3.7	1.4	6.0	1.8	5.7	
	<i>Planning regular tasks</i>	Improvement	57.5	59.5	-2.0	-4.3	0.2	-3.9	-0.1	0.153
		Maintained	13.7	13.9	-0.1	-1.8	1.6	-1.6	1.3	
		Declined	28.8	26.7	2.1	-0.2	4.4	0.2	4.1	
	<i>Use of mobility device</i>	Improvement	37.6	38.3	-0.7	-2.6	1.2	-2.3	0.9	0.597
		Maintained	13.9	13.2	0.7	-0.8	2.1	-0.6	1.9	
		Declined	48.5	48.5	0.0	-2.0	2.0	-1.7	1.7	
	<i>Walking without rest</i>	Improvement	29.6	30.7	-1.2	-3.2	0.9	-2.9	0.6	0.315
		Maintained	24.5	24.8	-0.3	-2.2	1.6	-1.9	1.3	
		Declined	45.9	44.4	1.5	-0.5	3.5	-0.2	3.1	

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status (Continued)	Going up or down stairs	Improvement	30.4	29.7	0.7	-1.4	2.8	-1.0	2.5	0.402
		Maintained	24.7	26.1	-1.4	-3.5	0.6	-3.1	0.3	
		Declined	44.9	44.2	0.7	-1.3	2.7	-1.0	2.4	
	Physical/emotional problems limiting social activities	Improvement	44.2	45.1	-0.8	-3.6	1.9	-3.1	1.5	0.515
		Maintained	25.0	25.6	-0.6	-3.0	1.9	-2.6	1.5	
		Declined	30.8	29.4	1.4	-1.0	3.8	-0.6	3.4	
	Pain limiting regular activities	Improvement	42.9	42.0	0.9	-1.7	3.5	-1.3	3.1	0.424
		Maintained	31.4	33.1	-1.7	-4.1	0.8	-3.7	0.4	
		Declined	25.6	24.9	0.8	-1.8	3.3	-1.4	2.9	
Care Experience	Felt prepared to leave the hospital	Very or Somewhat	89.6	91.1	-1.6	-3.5	0.3	-3.2	0.0	0.100
	Medical staff took your preferences into account in deciding what health care services you should have after you left the hospital	Strongly Agree or Agree	84.6	89.0	-4.4	-6.5	-2.3	-6.2	-2.6	0.000
	Good understanding of how to take care of self before going home	Strongly Agree or Agree	90.5	91.8	-1.3	-3.0	0.3	-2.7	0.0	0.112
	Medical staff clearly explained how to take medications before going home	Strongly Agree or Agree	88.7	90.0	-1.3	-3.0	0.5	-2.8	0.2	0.162
	Medical staff clearly explained what follow-up appointments or treatments would be needed before going home	Strongly Agree or Agree	89.3	91.1	-1.8	-3.7	0.1	-3.4	-0.2	0.060
	Able to manage your health needs since returning home	Strongly Agree or Agree	93.2	93.5	-0.3	-1.7	1.1	-1.5	0.9	0.698
	Medical staff talked with you about whether you would have the help you needed when you got home	Yes	84.5	85.5	-1.0	-3.2	1.3	-2.9	0.9	0.400

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Care Experience (Continued)	<i>If you needed help at home to manage your health, medical staff arranged services for you at home to help manage your health</i>	Yes	69.1	69.1	0.0	-3.5	3.4	-2.9	2.8	0.983
Satisfaction With Care	<i>Overall satisfaction with recovery</i>	Quite a Bit or Extreme	55.9	56.5	-0.6	-3.3	2.1	-2.9	1.7	0.660
	<i>Rating of all care received after leaving the hospital</i>	9-10	55.8	55.6	0.2	-2.7	3.1	-2.3	2.6	0.654
		7-8	25.9	26.9	-1.0	-3.5	1.5	-3.1	1.1	
		0-6	18.3	17.5	0.8	-1.5	3.1	-1.2	2.8	

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.4: Patient Survey Outcomes, Surgical Episodes, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status	<i>Bathing, dressing, using the toilet or eating</i>	Improvement	78.2	78.4	-0.2	-2.4	1.9	-2.0	1.6	0.892
		Maintained	7.2	7.4	-0.2	-1.5	1.1	-1.3	0.9	
		Declined	14.6	14.1	0.4	-1.6	2.5	-1.3	2.2	
	<i>Planning regular tasks</i>	Improvement	82.6	83.1	-0.5	-2.6	1.6	-2.2	1.3	0.623
		Maintained	6.2	6.5	-0.3	-1.6	1.0	-1.4	0.8	
		Declined	11.2	10.4	0.8	-1.0	2.5	-0.7	2.3	
	<i>Use of mobility device</i>	Improvement	58.8	57.6	1.2	-1.1	3.6	-0.7	3.2	0.283
		Maintained	10.8	10.3	0.5	-1.2	2.2	-0.9	1.9	
		Declined	30.4	32.1	-1.7	-3.8	0.4	-3.5	0.1	
	<i>Walking without rest</i>	Improvement	57.7	59.2	-1.5	-4.3	1.3	-3.8	0.9	0.568
		Maintained	16.1	15.5	0.6	-1.4	2.5	-1.1	2.2	
		Declined	26.2	25.3	0.9	-1.1	2.9	-0.8	2.6	
	<i>Going up or down stairs</i>	Improvement	53.8	53.1	0.7	-1.9	3.3	-1.5	2.9	0.746
		Maintained	22.5	22.4	0.1	-2.1	2.3	-1.8	2.0	
		Declined	23.6	24.4	-0.8	-2.9	1.3	-2.6	1.0	

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Change in Functional Status (Continued)	<i>Physical/emotional problems limiting social activities</i>	Improvement	66.9	66.0	0.9	-1.7	3.4	-1.2	3.0	0.781
		Maintained	15.4	15.9	-0.5	-2.5	1.5	-2.2	1.2	
		Declined	17.8	18.1	-0.4	-2.5	1.8	-2.2	1.5	
	<i>Pain limiting regular activities</i>	Improvement	67.4	67.4	0.0	-2.5	2.4	-2.1	2.0	0.675
		Maintained	16.4	17.2	-0.8	-3.1	1.4	-2.7	1.1	
		Declined	16.2	15.3	0.9	-1.4	3.1	-1.0	2.7	
Care Experience	<i>Felt prepared to leave the hospital</i>	Very or Somewhat	94.8	93.9	0.9	-0.5	2.2	-0.3	2.0	0.206
	<i>Medical staff took your preferences into account in deciding what health care services you should have after you left the hospital</i>	Strongly Agree or Agree	91.5	91.5	0.0	-1.9	1.9	-1.6	1.6	0.972
	<i>Good understanding of how to take care of self before going home</i>	Strongly Agree or Agree	93.6	93.9	-0.4	-1.8	1.1	-1.6	0.9	0.629
	<i>Medical staff clearly explained how to take medications before going home</i>	Strongly Agree or Agree	93.3	93.8	-0.5	-2.0	1.0	-1.8	0.8	0.505
	<i>Medical staff clearly explained what follow-up appointments or treatments would be needed before going home</i>	Strongly Agree or Agree	94.8	95.3	-0.4	-1.8	0.9	-1.6	0.7	0.534
	<i>Able to manage your health needs since returning home</i>	Strongly Agree or Agree	95.8	97.0	-1.2	-2.2	-0.1	-2.0	-0.3	0.027
	<i>Medical staff talked with you about whether you would have the help you needed when you got home</i>	Yes	92.9	93.7	-0.8	-2.3	0.8	-2.1	0.5	0.342

Outcome		Response Category	BPCI Advanced Rate (%)	Comparison Rate (%)	Difference in Rate (pp)	95% LCI (pp)	95% UCI (pp)	90% LCI (pp)	90% UCI (pp)	P-Value
Care Experience (Continued)	<i>If you needed help at home to manage your health, medical staff arranged services for you at home to help manage your health</i>	Yes	76.5	75.4	1.2	-3.1	5.4	-2.4	4.7	0.588
Satisfaction With Care	<i>Overall satisfaction with recovery</i>	Quite a Bit or Extreme	73.7	74.9	-1.2	-3.9	1.5	-3.5	1.0	0.368
	<i>Rating of all care received after leaving the hospital</i>	9-10	71.5	71.8	-0.3	-2.9	2.4	-2.5	1.9	0.304
		7-8	18.4	19.3	-1.0	-3.3	1.3	-2.9	0.9	
		0-6	10.1	8.9	1.2	-0.5	3.0	-0.2	2.7	

Note: See the first page of this appendix for data sources and methodology.

B. Open-Ended Survey Response Results

Exhibit L.5: Response Rates for Open-Ended Survey Question, BPCI Advanced and Comparison Group, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Episode Type	BPCI Advanced		Comparison		Total	
	Count (N)	Response Rate (%)	Count (N)	Response Rate (%)	Count (N)	Response Rate (%)
Overall	5,349	78.2	5,484	76.8	10,834	77.5
Medical	2,923	74.6	2,543	62.5	5,467	68.5
Surgical	2,426	83.2	2,943	95.8	5,369	89.6

Note: The denominator for the response rates is the number of respondents for the closed-ended survey. See the first page of this appendix for data sources and methodology.

Exhibit L.6: Thematic Concepts Identified by Generative AI, BPCI Advanced and Comparison Group, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Thematic Concept	Description
Cardiac Rehabilitation	Structured recovery processes designed to improve heart function following a cardiac event. Includes both successful and unsuccessful experiences, as well as need for additional rehabilitation.
Cardiovascular Care	Cardiovascular procedures (such as stent, valve replacement, and bypass surgery) and treatments (such as medication).
Care for Complex Conditions	Experiences with multiple, complicated, or severe medical conditions requiring care from multiple specialists and/or provider types.
Communication About Medication	Information a provider shares with patient or caregiver about medications, including side effects and proper usage.
Communication Between Providers	Information sharing between providers about the patient's medical condition, treatment, and needs.
Compassionate Communication	Provider communicated to the patient with empathy.
Consistency of Information	Providers shared consistent information with the patient or caregiver. Includes consistency across providers and over time.
Coordination of Care	The overall coordination between and support from different health care providers (such as doctors, surgeons, specialists, nurses, physical/occupational therapists), facilities (such as hospitals, rehabilitation facilities, home health services), or other services during the patient care journey. Such coordination may include scheduling physical therapy, scheduling appointments with specialists, and facilitating transitions within hospitals. It also includes transitions and coordination between pre-operative and post-operative care.
Discharge Instructions	Providers shared clear and comprehensive self-care instructions at discharge (excludes discharge timing and medication).
Discharge Medications	Provider communication about medication changes or other medication-related issues during the discharge process. Such experiences include descriptions of improved health with medication changes at discharge, clear explanations of changes at discharge, and reduced use of medications (such as pain medication).
Discharge Process	The logistics and experience of the actual discharge from the hospital, focusing on general discharge experience and patient readiness.
Discharge Timing	Patient experiences with the timing or scheduling of discharge that created positive or negative experiences or challenges.

Thematic Concept	Description
Educational Resources	Materials and information provided to patients to help them understand and manage their health conditions.
Family Support in Recovery	The role of family members in providing care and support during hospital stay and in the recovery process after discharge.
Family Support Integration	The inclusion of family members in discharge planning and post-discharge care.
Follow-Up Appointment Scheduling	Patient experiences with scheduling and attending follow-up appointments with providers and therapists. Includes support from providers in scheduling appointments and difficulties scheduling appointments on their own.
Follow-Up Communication	Post-discharge communication initiated by providers to check on patient progress.
Heart Failure Management	Treatment and management of heart failure and associated conditions. Includes medication, lifestyle modifications, treatment adherence, symptom monitoring, use of medical devices, coordination with health care providers, and navigating challenges such as hospital readmissions or disease progression.
Home Health Care	Nursing visits and other medical support services (such as physical and occupational therapy) provided at home after discharge. Includes access to or lack of home health care as well as positive and negative experiences.
Hospital Staff Interactions	Patient perceptions about their interactions with doctors, nurses, and other health care professionals during hospitalization, including attentiveness, professionalism, communication style, and overall care quality.
Medical Explanations	Descriptions about how providers explained medical conditions, treatments, and procedures to patients in understandable terms.
Multidisciplinary Team Approach	Care provided by teams of health care professionals from different disciplines.
Pain Management	Experiences with how pain was addressed during hospitalization and after discharge, including medication effectiveness and alternative pain relief strategies.
Patient Concerns	How health care providers responded to or failed to respond to concerns or symptoms reported by patients, responses to patient requests for assistance; does not include general care.
Physical Therapy	Therapy services that affected recovery and mobility, either at home, outpatient, or inpatient. Includes quantity of visits and effectiveness of physical therapy.
Physician Interaction	The quality and quantity of direct interactions between patients and physicians during the hospital stay.
Readmission Experiences	Patient's experiences with hospital readmissions after initial discharge.
Rehabilitation Experience	Experiences with care at dedicated rehabilitation facilities after hospital discharge.
Responsive Staff Communication	How promptly providers addressed questions and concerns raised by patients.
Staffing Issues	Patient perceptions about hospital staffing levels and how staffing levels affected care quality, including response times and availability of assistance.
Support for Solo Living	Support provided to patients to address challenges of living alone, such as home health care, transportation to appointments, and home physical therapy. Includes patient descriptions of the health care challenges of living alone.
Surgery Outcomes	Surgical procedures and associated outcomes, including pain reduction, improved mobility, and overall satisfaction with interventions. A procedure is considered surgery based on additional information indicated; for example, responses that mention both "procedure" and "surgeon," responses that mention broken bones or muscles, or responses that mention atrial fibrillation. A procedure is not considered a surgery if the response merely says "procedure," "minor procedure," or "planned procedure" with no additional information that could indicate surgery.

Thematic Concept	Description
Transition Between Care Settings	Coordination, information transfer, and treatment or support across different health care settings during patient transitions.
Transition to Rehabilitation	The process of moving from hospital to rehabilitation facilities, including coordination and information transfer.

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.7: Number of Thematic Concepts per Response, BPCI Advanced and Comparison Group, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Number of Thematic Concepts	BPCI Advanced Responses		Comparison Group Responses	
	Count (N)	Percent (%)	Count (N)	Percent (%)
One	922	17.0	908	16.6
Two	838	15.4	887	16.2
Three	663	12.6	678	12.4
Four or More	1,555	28.6	1,622	29.6

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.8: Open-Ended Survey Sentiments by Thematic Concept, BPCI Advanced and Comparison Group, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Thematic Concept	Number of Responses (N)	BPCI Advanced Responses			Comparison Group Responses			P-Value
		Sentiment			Sentiment			
		Positive (%)	Neutral (%)	Negative (%)	Positive (%)	Neutral (%)	Negative (%)	
Coordination of Care	6,318	69	11	20	70	12	19	0.158
Hospital Staff Interactions	5,478	85	5	10	86	4	10	0.414
Physical Therapy	2,105	78	12	10	81	11	8	0.071
Surgery Outcomes	2,094	67	19	14	67	20	14	0.959
Physician Interaction	1,914	73	11	16	75	10	14	0.380
Home Health Care	1,578	79	13	9	80	14	7	0.447
Compassionate Communication	1,567	92	*	8	92	*	7	0.574
Care for Complex Conditions	1,385	40	32	28	41	33	25	0.652
Transition Between Care Settings	1,288	51	30	19	55	28	18	0.419
Rehabilitation Experience	1,178	62	17	21	69	16	16	0.034
Patient Concerns	1,150	39	6	55	39	6	56	0.982
Responsive Staff Communication	1,018	74	*	25	75	*	24	0.662
Medical Explanations	966	74	7	19	76	11	14	0.038
Communication About Medication	910	49	15	36	49	21	30	0.034
Pain Management	879	49	13	38	52	11	37	0.524
Follow-Up Appointment Scheduling	860	67	20	13	71	17	12	0.313
Cardiovascular Care	857	56	26	18	64	21	15	0.036
Discharge Process	846	63	9	28	65	8	27	0.726
Multidisciplinary Team Approach	700	91	8	*	92	5	*	0.353
Transition to Rehabilitation	665	49	33	18	57	28	15	0.121
Discharge Medications	654	60	1	28	60	18	21	0.044
Family Support Integration	641	73	15	12	76	13	11	0.553

Thematic Concept	Number of Responses (N)	BPCI Advanced Responses			Comparison Group Responses			P-Value
		Sentiment			Sentiment			
		Positive (%)	Neutral (%)	Negative (%)	Positive (%)	Neutral (%)	Negative (%)	
Family Support in Recovery	569	79	15	7	81	11	8	0.287
Follow-Up Communication	525	83	4	14	84	6	10	0.289
Discharge Timing	453	34	22	44	34	23	43	0.910
Support for Solo Living	442	60	15	26	67	10	23	0.206
Staffing Issues	439	9	8	83	11	*	84	0.317
Discharge Instructions	419	80	*	18	83	*	13	0.196
Readmission Experiences	363	6	30	64	6	25	70	0.492
Communication Between Providers	278	45	*	50	53	*	39	0.160
Educational Resources	258	84	*	15	87	*	11	0.733
Heart Failure Management	223	52	18	30	54	25	21	0.252
Cardiac Rehabilitation	169	58	23	19	76	13	12	0.049
Consistency of Information	159	20	*	80	28	*	72	0.273

Note: See the first page of this appendix for data sources and methodology.

Exhibit L.9: Open-Ended Survey, Thematic Concept Pairs With Greatest Overlap, BPCI Advanced and Comparison Group, Hospitals and PGPs, Model Years 7–8 (Apr. 2024–Jan. 2025)

Thematic Concept 1	Thematic Concept 2	Responses Attributed With Concept 1 Also Attributed with Concept 2 (%)	Responses Attributed With Concept 2 Also Attributed With Concept 1 (%)	Correlation Coefficient
Family Support Integration	<i>Family Support in Recovery</i>	78	69	0.73
Discharge Medications	<i>Communication About Medication</i>	61	85	0.72
Transition to Rehabilitation	<i>Rehabilitation Experience</i>	53	93	0.65
Responsive Staff Communication	<i>Patient Concerns</i>	41	46	0.43
Discharge Instructions	<i>Educational Resources</i>	54	33	0.42
Transition to Rehabilitation	<i>Transition Between Care Settings</i>	33	64	0.40
Hospital Staff Interactions	<i>Compassionate Communication</i>	92	26	0.36
Physical Therapy	<i>Home Health Care</i>	51	38	0.35
Discharge Process	<i>Discharge Timing</i>	56	30	0.33
Physical Therapy	<i>Coordination of Care</i>	29	87	0.32
Cardiac Rehabilitation	<i>Cardiovascular Care</i>	14	73	0.30
Pain Management	<i>Surgery Outcomes</i>	23	56	0.30
Compassionate Communication	<i>Responsive Staff Communication</i>	44	29	0.28
Rehabilitation Experience	<i>Transition Between Care Settings</i>	35	39	0.27
Home Health Care	<i>Coordination of Care</i>	22	86	0.27

Note: See the first page of this appendix for data sources and methodology.