January 2014

Evaluation of the Medicare Gainsharing Demonstration

Final Report

Prepared for

William Buczko, PhD
Centers for Medicare & Medicaid Services
Center for Medicare and Medicaid Research and Innovation
Mail Stop WB-06-05
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Leslie Greenwald, PhD
Walter Adamache, PhD
Maggie Cole-Beebe, PhD
Peter Amico, PhD
Elizabeth Hunter, BA
Robert Baker, BA
Nathan West, MS
RTI International
Health, Social, and Economics Research
1440 Main Street, Suite 310
Waltham, MA 02451

RTI Project Number 0209853.003.001.002

RTI INTERNATIONAL

EVALUATION OF THE MEDICARE GAINSHARING DEMONSTRATION:

FINAL REPORT

FINAL REPORT

by Leslie Greenwald, PhD
Walter Adamache, PhD
Maggie Cole-Beebe, PhD
Peter Amico, PhD
Elizabeth Hunter, BA
Robert Baker, BA
Nathan West, MS

Associate Project Director: Leslie Greenwald, PhD

Federal Project Officer: William Buczko, PhD

RTI International

CMS Contract No. HHSM-500-2005-00029I; Task 3

January 2014

This project was funded by the Centers for Medicare & Medicaid Services under contract no. HHSM-500-2005-00029I; Task 3. 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. RTI assumes responsibility for the accuracy and completeness of the information contained in this report.

RTI International is a trade name of Research Triangle Institute.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

This page left intentionally blank.

SECTION 1
LEGISLATIVE MANDATE AND OVERVIEW OF THE GAINSHARING MODEL

1.1 Legislative Mandates for the Medicare Gainsharing Demonstration

The Congress, under Section 5007 of the Deficit Reduction Act (DRA) of 2005, required that the Centers for Medicare & Medicaid Services (CMS) conduct a qualified gainsharing program to test alternative ways that hospitals and physicians can share in efficiency gains. The primary goal of the demonstration was to evaluate gainsharing as a means to align physician and hospital incentives to improve quality and efficiency. The DRA-mandated Gainsharing demonstration mandate was amended by Section 3027 of the Affordable Care Act. As part of the statutory mandate, the Secretary of the Department of Health and Human Services is required to submit two reports to Congress: a quality improvement and savings report and a final report. This report presents the final evaluation results for the two sites participating in the demonstration at the time of implementation:

- Beth Israel Medical Center (BIMC), New York, New York
- Charleston Area Medical Center (CAMC), Charleston, West Virginia

CMS solicited volunteer participating sites for the Gainsharing demonstration in the fall of 2006. Applications were due to CMS on November 17, 2006. The DRA legislation originally mandated participation of a total of six sites (four urban and two rural). CMS initially selected four sites from this solicitation for demonstration participation, although no rural sites were selected from this first round. CMS issued a new announcement to solicit additional rural demonstration applications.

CMS ultimately designated four urban sites for participation in the demonstration: BIMC, CAMC, Saint Joseph's Hospital in Syracuse, New York, and Deaconess Hospital in Evanston, Illinois. The follow-up rural solicitation resulted in the additional designation of Lake Cumberland Regional Hospital in Somerset, Kentucky. Three sites moved from the initial selection phase to sign terms and conditions, including the mandated rural site. Two sites (Saint Joseph's and Deaconess) withdrew from the demonstration at various stages and for various reasons, primarily concerns about their potential exposure to post-acute care (PAC) financial risk. Later, the Lake Cumberland rural site voluntarily determined that its proposed model for gainsharing could not meet the implementation and evaluation requirements of the demonstration. At the time of implementation in October 2008, only two sites (BIMC and CAMC) participated in the demonstration. CAMC operated for one year then elected not to continue participation in the demonstration past December 2009.

Section 2 in this report summarizes each site's gainsharing approach as envisioned in its original demonstration application. Additional details on post implementation experiences of the gainsharing sites are described in the site visit and physician focus group analyses (Section 4). Although the anticipated start date for the demonstration was January 1, 2007, demonstration sites did not begin the implementation process until October 1, 2008; they operated until December 31, 2009. At that point, the original legislative authorization for the demonstration ended, although the sites were allowed to continue all operations except for actually making

gainsharing payments to physicians. The demonstration was officially extended through September 30, 2011, as a result of the Affordable Care Act enacted on March 23, 2010. BIMC elected to continue implementation through that extended end date. CAMC elected to end its participation in the demonstration as of December 31, 2009, and will therefore be evaluated only through that time period.

1.2 Overview and History of the Gainsharing Model

Current trends in health care reform emphasize moving the health care system toward models that hold health care providers more accountable for the costs and quality of the care they provide, thereby encouraging greater efficiency and improved outcomes. The gainsharing model is one variant of these systems emphasized under health care reform. Gainsharing models developed in health care because of the potential for misalignment of incentives between hospitals and physicians.[1] The potential for misalignment arises because, under the Medicare fee-for-service program, hospitals and physicians are paid separately for care provided in hospitals under Part A and Part B, respectively. Under the prospective payment system for inpatient hospitals (IPPS), hospitals are paid a fixed amount, based on the principal diagnosis that covers most of the associated hospital costs, including those primarily under a physician's control. Meanwhile, Medicare generally pays physicians per procedure and, implicitly, for volume; a physician who provides more services to a hospitalized patient will typically receive more in reimbursement. Physicians earn no financial gains for providing more efficient care or for lowering hospital costs; they are independent agents who by their use of hospital facilities can directly or indirectly, knowingly or unknowingly, affect hospital costs. Physicians may unknowingly increase hospital costs through unnecessary or inefficient use of hospital resources such as operating room time or disposable surgical supplies. Physicians may also knowingly increase hospital costs by, for example, ordering duplicative testing. Local practice patterns, not necessarily consistent with evidence-based or best clinical practice guidelines, may also influence physician behavior and lead to less-than-efficient clinical care.

Gainsharing is one potential solution to remedy misalignment of hospital and physician incentives. In a hospital-physician gainsharing program, hospitals offer physicians a share of any cost savings achieved by the hospital as a result of the physicians' behavior or decisions. Gainsharing works by providing physicians with a financial stake in controlling hospital costs.

Fraud and Abuse Laws: Gainsharing programs provide an avenue for improvement in efficiency that should result in savings to both hospitals and third-party payers such as Medicare. However, hospitals and physicians that wish to enter into gainsharing arrangements must meet the requirements of the federal anti-kickback statute, the physician self-referral law, and the Civil Monetary Penalty Law.[2]

Gainsharing in Medicare: CMS' first attempt at hospital-physician gainsharing was in the Medicare Heart Bypass Demonstration, conducted from 1991 to 1996 (Cromwell et al.,

---

1 Gainsharing can also exist between payers and physicians as well as payers and patients.

2 42 C.F.R. Sect 1003.

1998), hospital costs were reduced (Cromwell, Dayhoff, and Thoumaian, 1997). All seven sites had waivers to engage in gainsharing, and groups designed and implemented more or less complicated gainsharing algorithms on their own, subject to CMS' final approval. Surgeons, cardiologists, radiologists, anesthesiologists, and pathologists all received fixed, negotiated payment amounts that were included in the hospital payment (no direct Part B inpatient billing of Medicare). Under this successful demonstration, physicians enjoyed gainsharing bonuses, quality improved, and no negative offsets to Medicare savings occurred as a result of shifts of care to the post-acute setting.

In 2001, the New Jersey Hospital Association (NJHA) submitted an application to CMS to operate an eight-hospital demonstration of gainsharing in its state covering all all-patient refined diagnosis-related groups (APR-DRGs; Marcoux, 2007). The application was approved by CMS and became operational in early 2004 as the Hospital Performance-Based Incentives Demonstration. The New Jersey demonstration plan was to establish maximum pools from generated savings of Part A hospital savings for each APR-DRG in the hospital and to share those savings with the medical staff. These pools were constrained to 25 percent of total Part B outlays. Next, the pools were converted to a per-discharge cost for each APR-DRG, which was based on average costs of the lowest 90 percent of cases (so-called best practice norms). Excluding the most expensive cases from the target baseline cost per discharge was the primary mechanism to achieve reductions in hospital costs. Once responsible physicians were identified, they became eligible for gainsharing depending on how the average cost of their cases related to the mean cost of the 90 percent baseline group of cases. Baseline and demonstration cases were standardized for case severity and inflation. In the early performance years, responsible physicians could participate in gainsharing even if they failed the best practice norms, as long as they showed reductions in their Part A costs per case. Gainsharing pools were carved out for hospital-based and consulting physicians to partially shelter them from lost billings associated with shorter stays and less testing. Process and outcome indicators were to be used to restrict gainsharing to physicians maintaining high quality standards.

The NJHA Hospital Performance-Based Incentives Demonstration differed from its predecessor, the Heart Bypass Demonstration, in that the latter put surgeons at risk for both Part A and Part B billings in a single global payment only for a few cardiac DRGs. The NJHA demonstration maintained separate Part A and Part B billing practices. Also, physicians in the NJHA project were put at risk for excessive post-acute care Medicare outlays from any source (including outpatient physician services). The two demonstrations also differed in that CMS negotiated up-front discounts in its cardiac DRG global Part A and Part B rates for the Medicare Heart Bypass Demonstration, whereas New Jersey hospitals had to reduce baseline Part A and Part B inpatient outlays by 2 percent after adjusting for inflation and case mix changes.

The Hospital Performance-Based Incentives Demonstration did not last long; four New Jersey-area hospitals that were excluded from the demonstration project obtained an injunction in Federal court to stop it. They argued that the NJHA's program was anticompetitive and that it violated the civil monetary penalties and anti-kickback regulations. Closely related to gainsharing projects was the Medicare Physician Group Practice (PGP) demonstration, a shared savings model and one of Medicare's first projects that established incentives for quality improvement and cost efficiency. It shared savings with physicians meeting these targets at the group practice level. A legislative mandate for the PGP demonstration was included in the

Medicare, Medicaid, and SCHIP Benefits Improvement and Protection Act of 2000. It established several goals, including (Kautter et al., 2007) encouraging coordination of health care furnished under Medicare Part A and Part B, encouraging investment in administrative structures and processes for efficient service delivery, and rewarding physicians for improving health care processes and outcomes. The PGP demonstration began on April 1, 2005, and ended March 31, 2010. Ten large multispecialty physician groups participated.

The Acute Care Episode (ACE) Demonstration (authorized under the Medicare Prescription Drug, Improvement and Modernization Act of 2003) is a 3-year demonstration that primarily tested the use of a global payment covering all Medicare Part A and Part B services for specified cardiovascular and orthopedic procedures. Five hospitals were selected to participate and began participation on November 1, 2010. The demonstration concluded on…..Gainsharing arrangements for participating sites and their physicians are allowed under this demonstration, and four of the five participating sites have implemented gainsharing arrangements. The evaluation of this demonstration is expected to be complete in 2014…

CMS also implemented the Medicare Physician-Hospital Collaboration (PHC) Demonstration, as required under Section 646 of the Medicare Prescription Drug, Improvement and Modernization Act of 2003. The PHC demonstration operated from July 2009, through September, 2012. Similar to the Medicare Gainsharing Demonstration, the primary goal of the PHC demonstration was to evaluate gainsharing strategies aimed at improving the quality of care in a health delivery system. The demonstration sites implemented approaches to align physician and hospital financial incentives and ultimately lead to reductions in internal hospital costs of care and overall Medicare costs of care, in and up to 90 days beyond the acute inpatient stay. The only participant in the demonstration was the NJHA/New Jersey Care Integration Consortium, with 12 participating hospitals. Participating sites were not required to generate Medicare program savings, but they must maintain budget neutrality within the inpatient stay and up to 90 days beyond the acute inpatient stay as well as meet a minimum quality of care.

1.3 Overview of the Evaluation Design

For the evaluation of this demonstration, RTI consulted CMS, the Actuarial Research Corporation (ARC), and the demonstration sites for information on the development and implementation of their demonstration models. RTI also conducted independent analyses using Medicare administrative data. RTI prepared a series of reports for this project. As the Final Report, this report summarizes the hypotheses and research questions, methods, data collection, findings, and policy relevance of the demonstration and presents overall evaluation findings.

The evaluation addresses a variety of research questions and will assess the effects of different gainsharing models on

- hospital efficiency,
- physician referral patterns,
- Medicare expenditures,

- quality of care, and
- beneficiary satisfaction.

A summary of the primary analytic tasks follows.

Comparison Groups: CMS used a trended-baseline methodology to determine whether participating hospitals have achieved budget neutrality and if Medicare payments and quality of care have changed during the demonstration. Comparison groups are necessary because the demonstration applicants otherwise can only compare their own performance year experience to that of a base year (i.e., a simple pre-post analysis). Observing only pre-post differences does not control for changes experienced by similar nonparticipants during the demonstration period. One must observe both types of differences to determine the effects attributable to the Gainsharing demonstration. Therefore, RTI also compared the performance of the demonstration sites with that of independent comparison sites not participating in the Gainsharing demonstration. A complete summary of how this was accomplished is provided in Section 3 of this report.

Site Visits and Physician Focus Groups: Site visits and physician focus groups were required under this evaluation contract. This qualitative data collection process will document and analyze initial implementation and ongoing operations of the different Gainsharing demonstration sites. Site visits were conducted for CAMC and BIMC in the fall of 2010. A follow-up second visit was conducted with BIMC in November 2011; we did not revisit CAMC as it had withdrawn from the demonstration. We discussed the participation decision, details of the demonstration design, and initial and ongoing implementation; methods and evidence for cost reductions and quality impacts attributable to the intervention; and relationships with physicians and other providers. Paralleling and in coordination with the site visits, physician focus group discussions were also conducted. The goal of the physician focus groups was to gather information on physicians' experience and satisfaction with the gainsharing arrangements at their respective sites. In these focus groups, RTI collected in-depth information on physicians' behavioral responses to incentives, the evolution of gainsharing methods at each site, and physician satisfaction with the arrangements, along with patient referral patterns and evidence of biased selection.

Analysis—Implementation and Organizational Response Analysis: ARC, the implementation contractor, had the lead responsibility for monitoring gainsharing arrangements and ensuring that payments adhered to the payment policies set in the demonstration protocols. The evaluation's analysis of organizational responses will be qualitative and based on the site visits and physician focus groups. Issues to be investigated include

- overall perceptions of the Gainsharing demonstration,
- rationale for participation in the Gainsharing demonstration,
- perceptions of methods used to achieve savings and efficiency,
- changes in relationship between physicians and hospitals as a result of gainsharing,

- changes in clinical patterns of care (e.g., clinical pathways, shorter stays, fewer consults), and

- roles of physicians and hospitals in developing and monitoring changes in care delivery.

Analysis—Medicare Expenditures and Savings: The RTI evaluation of Medicare payments and savings overlaps to some degree with the responsibilities of the demonstration implementation support contractor (ARC), which has official responsibility for determining whether the demonstration is budget neutral. RTI and ARC are jointly involved in analyzing financial reconciliation and quality performance. The RTI evaluation will also

- describe and critique gainsharing methods,

- determine financial impacts of gainsharing on providers,

- adjust for patient severity and for substitution of PAC for inpatient care, and

- analyze sources of Medicare savings: inpatient hospital compared with PAC.

Analysis—Quality of Care: A critical aspect of the evaluation is an assessment of whether quality of care has been affected by the gainsharing financial incentives. Quality-of-care analyses in the evaluation will compare changes in quality measures for demonstration hospitals with those from comparison hospitals. Because all of these indicators are constructed from Medicare claims data, RTI will have complete data for all hospitals. Quality measures analyzed include

- inpatient and 30-day post discharge mortality,

- readmissions within 30 days of discharge, and

- inpatient quality indicators and patient safety indicators from the Agency for Healthcare Research and Quality (AHRQ).

These analyses adjust for patient severity using the APR-DRG risk adjustment grouper.

Analysis—Patient Satisfaction: An important aspect of quality of care is patients' perspectives about the care they receive during their hospital stays. The Hospital Consumer Assessment of Health Plans Survey (HCAHPS) provides annual measures on patient satisfaction for participating hospitals. CMS has made participation in HCAHPS a requirement for the demonstration sites. We analyze the difference in beneficiary satisfaction between demonstration and comparison hospitals before and after program implementation.

Analysis—Physician Referral Patterns: The potential for additional incentive payments for physicians under gainsharing may affect the decisions physicians make, including increasing the probability of certain attractive patients' being admitted to a demonstration hospital by participating physicians. Participating physicians may also have an incentive either to transfer

very costly and difficult-to-manage cases to other acute care hospitals (IPPS transfers) or to discharge them to PAC providers. Increased transfers might, in turn, result in a reduction in demonstration hospital outlier cases. To monitor these potential referral patterns and market competition impacts due to gainsharing, RTI presents descriptive analyses that include tabulating and statistically testing differences between the demonstration hospital and its competitor hospitals (before and during the demonstration) using the following indicators:

- shares of more or less complex Medicare Severity (MS)-DRG cases,
- ER admissions,
- overall transfers in and out,
- transfers of more or less complex MS-DRG cases, and
- outliers.

1.4 Outline of This Report

This report will focus on the three performance analyses in the Gainsharing demonstration and comparison sites. Section 2 provides an overview of the gainsharing models implemented by the CAMC and BIMC demonstration sites. Section 3 describes in detail the comparison site selection process that forms the basis of our comparative and difference-in-difference analyses. These comparison sites are being used by both the evaluation and implementation contractors for the quality of care, Medicare expenditures, and savings analyses. Section 4 summarizes the implementation and organizational response analyses, which are based on the evaluation site visits and physician focus groups. Section 5 summarizes Medicare expenditures and savings findings. Section 6 presents the quality-of-care findings for the demonstration and comparison sites. Section 7 presents baseline patient satisfaction indicators. Section 8 summarizes the referral patterns analysis. Section 9 offers a summary discussion.

This page left intentionally blank.

SECTION 2
SUMMARY OF MEDICARE GAINSHARING DEMONSTRATION SITE PARTICIPANTS

2.1 Charleston Area Medical Center, Charleston, West Virginia

Overview: CAMC has 893 beds and is by far the largest of the six acute care hospitals in the Charleston, West Virginia, area. It is almost five times larger than the next largest hospital in the area, which has 189 beds. CAMC is the main tertiary care hospital serving West Virginia, northeastern Kentucky, and southeastern Ohio, including more than 300,000 people in the Charleston metropolitan area. Three data systems were used to support the demonstration project: CathSource, HeartSource, and TSI Cost Accounting System. In detail:

- CathSource and HeartSource are software tools that compiled data during procedures in accordance with guidelines from the American College of Cardiology and the Society of Thoracic Surgeons.
- Data consistency in all CathSource/HeartSource users provided consistent benchmarks (national, volume-based, or best-in-class).
- Data on the following items were collected at the point of care:
  - cost
  - clinical quality
  - productivity
  - laboratory work and radiology tests
- TSI Cost Accounting was used as a supplement to collect data on laboratory work and radiology tests.

Eligible DRGs, Patients, and Physicians: The CAMC gainsharing model focused on cardiac DRGs. Almost 40 percent of CAMC's Medicare revenue is generated from cardiovascular medicine, and cardiovascular DRGs have a direct annual cost of $31 million—more than 55 percent of direct costs. The following DRGs were proposed by CAMC for inclusion in the demonstration: DRGs 104–106, 117, 118, 121, 122, 127, 130, 131, 138, 139, 143, 515, 518, 535, 536, 547, 548, 549, 550, 551, 552, 555, 556, 557, and 558. Patients were identified on admission.

CAMC was motivated to participate in gainsharing by hospital cardiologists and cardiovascular surgeons. The hospital generally experienced few issues in attracting most physicians to the demonstration, with the exception of the major cardiology group. The cardiologists decided to participate in the quality-of-care and efficiency initiatives, but most

declined to participate in the gainsharing payment program because they were concerned about the negative perceptions that might surround such payments.[3]

Gainsharing Strategy: Each included DRG had established savings initiatives. CAMC offered the following example of how these savings initiatives would work. Catheterization laboratory staff at CAMC work less efficiently than ideal and are found to have an average waiting time of 13.5 minutes per case. Assuming 3,000 cases per year, this time translates to 675 hours per catheterization laboratory. Furthermore, assuming three staff per catheterization laboratory, this translates into 2,025 hours per laboratory. At an hourly rate of $30, eliminating this waiting time for catheterization laboratory staff could save $60,000 annually.

Budget Neutrality Strategy: In its original application, CAMC assured budget neutrality for Medicare through internal monitoring. Gainsharing was not to be awarded if no internal savings were generated. CAMC anticipated that internal savings would be generated by the following initiatives:

- examination of physician practice differences
- utilization of laboratory resources as needed
- evaluation of product usage
- increase in patient flow
- negotiation of lower prices for medical devices and supplies

In the end, CAMC reported that coordinating with physicians in negotiations for more efficient purchasing of lower-cost medical devices and surgical supplies became one focal strategy.

Medicare Cost Impact: Medicare payment, internal staff and consultant costs, and Medicare patient volume were expected to remain constant. The site offered the following cost scenario detail in its original proposal (***Table 2-1***).

[3] With the withdrawal of the cardiologists from the Gainsharing demonstration, DRGs 117, 121, 122, 127, 130, 131, 138, 139, 143, 518, 555, 556, 557, and 558 were dropped from the demonstration. On the other hand, discharges with the following ICD-9 procedure codes were included in the demonstration: 00.61 (carotid artery intervention) and 39.90, 00.55, and 39.50 (peripheral vascular intervention).

Table 2-1
CAMC Reported internal savings

Variable	Baseline	Year 1	Year 2	Year 3
Medicare payment	$64,735,318	$64,735,318	$64,735,318	$64,735,318
Direct costs	NA	NA	NA	NA
Patient care	$30,811,844	$29,271,252	$28,346,896	$27,730,660
Internal staff and consultant	$0	$300,000	$300,000	$300,000
Total direct costs	$30,811,844	$29,571,252	$28,646,896	$28,030,660
Cost savings (Baseline − Year X)	NA	$1,240,592	$2,165,948	$2,781,184
Medicare patient volume	4,386	4,386	4,386	4,386

NA = not applicable. SOURCE: Charleston Area Medical Center gainsharing application.

CAMC's decision to participate in the demonstration only through December 31, 2009, means that this site has impacts from only one year. CAMC withdrew from the demonstration for a variety of reasons, including continued concern over financial risk for PAC.

Quality Assurance: Gainsharing's success relies on patients' receiving quality of care that is equal to or better than what they would have received otherwise. CAMC proposed to measure physician care provided on several factors to ensure that quality of patient care remained the same. Worse performance on any of the following standards for an individual physician would make him or her ineligible to receive the gainsharing bonus:

- readmission rates
- repeat procedures
- patient outcomes
- major events during procedures
- antithrombotic usage

2.2 Beth Israel Medical Center, New York, New York

Overview: BIMC is a large, urban, academic hospital with 1,106 beds on two campuses: one in downtown Manhattan (Petrie) and one a community hospital in Brooklyn (Kings Highway). In its application, BIMC argued that it would be able to scale up the demonstration easily because it had implemented a similar gainsharing model for its private insurance patients. BIMC employs the Patient Real Time Information System (PRISM). PRISM includes computerized physician order entry, which maintains information on best prescribing practices and information such as drug interactions and maximum dose checking. A New York State

billing database, the Statewide Planning and Research Cooperative System (SPARCS), provides cost data on best practice norms within the Continuum hospital system (BIMC's parent organizational entity).

Eligible DRGs, Patients, and Physicians: BIMC included most medical and surgical DRGs in its demonstration. Enrollment was voluntary for physicians. At the time of the application, 600 physicians had been employed by the hospital's medical staff for at least 1 year and were thus eligible to enroll in the demonstration; BIMC anticipated that more than 70 percent would enroll. Ultimately, 271 physicians enrolled in the Gainsharing Demonstration.

Gainsharing Strategy: BIMC adopted a gainsharing plan designed by Applied Medical Software, Inc. (AMS). A pool of bonus funds was prospectively estimated from hospital savings on the basis of the following factors:

- Total available incentive is a percentage of the best practice variance for each APR-DRG.
- Best practice variance = (actual spending − best practice cost)
- Best practice cost = spending of the lowest-cost 25th percentile

If no hospital savings were realized, no bonuses were allocated to physicians participating in the demonstration. The total available incentive was defined as

$$\textit{total available incentive} = X\% \times (\textit{actual spending} - 25^{th}\ \textit{percentile spending})$$

where X% = the percentage of spending to allot to the incentive pool.

An incentive pool calculation was made for every APR-DRG and then summed across all APR-DRGs. Put differently, for each DRG, the hospital assigned some percentage of the difference between costs incurred on each Medicare patient minus the costs per case at the 25th percentile. These were summed across all cases.

Purpose of Bonus: In BIMC's proposed strategy, the purpose of the bonus was to counteract the incongruity between the hospital's Medicare payment and physician decision making. Before the gainsharing project, physicians had no direct financial incentive to use hospital resources more efficiently. Gainsharing gives physicians a cumulative incentive to provide only the care that is needed to maintain quality. Physicians earn a share of the total available incentive on the basis of their own efficiency or lower costs.

Gainsharing Distribution to Physicians: In the BIMC model, each patient was assigned to one practitioner who took financial responsibility for the care of the patient. For medical patients, the responsible physician was the attending physician. For surgical patients, the responsible physician was the surgeon. The maximum performance incentive was equal to the APR-DRG–adjusted portion of the total incentive pool allotted to the responsible physician.

The actual bonus paid to physicians, the performance incentive, was calculated as a percentage of the maximum performance incentive, based on performance. Gainsharing payments were capped according to CMS policy at 25 percent of the physician's affiliated Part B reimbursements. The total incentive payment is divided into four categories:

- performance, medical
- performance, surgical
- improvement, medical
- improvement, surgical

Medical and surgical specialists had different gainsharing algorithms: one based on costs relative to their low-cost peers (performance) and another based on their own cost improvement (improvement). Total incentives were weighted toward improvement in the first year and then moved toward performance weighting during later years. By Year 3, the physician incentive depended entirely on cost performance relative to a peer group. This simulation is summarized in ***Table 2-2***, reproduced from BIMC's application.

Once actual implementation began, BIMC decided to maintain the improvement percentage (67 percent) through Year 2 as a way to continue to emphasize improvement. As of Year 3, all annual gainsharing incentives became 100 percent performance based.

Table 2-2
BIMC annual gainsharing incentives

Year	Improvement	Performance
1	67%	33%
2	33%	67%
3+	0%	100%

SOURCE: Beth Israel Medical Center gainsharing application.

Performance Incentives: A physician's peer performance incentive was based on his or her average cost per case relative to the best practice cost per case of a cost-efficient peer group. The total performance incentive (PI) formula was

$$PI = \frac{90th\ percentile\ cost - MD's\ actual\ cost}{90th\ percentile\ cost - best\ practice\ cost} \times maximum\ payment$$

If the physician's actual average cost per case was in the 90th percentile or higher, the performance incentive is equal to 0. If the physician was at the best practice cost, or better, the performance incentive was the maximum payment. The best practice cost established a lower bound on gainsharing to discourage skimping on care.

The performance incentive was calculated by averaging patient costs for each eligible physician, then sorting them from most to least costly. The 90th percentile cost threshold is the average cost cutoff point of the physicians spending in the top 10 percent, on average. Best practice cost is the 25th percentile cost threshold that identifies the least costly 25 percent of physicians' patients. Physicians whose average cost was below the 90th percentile cost were eligible for a bonus, or a fraction of the maximum potential payment. The fraction was determined by scaling each physician's cost savings in the numerator to the maximum allowed savings in the denominator. For example, if the 90th percentile was $15,000, a physician's average cost was $12,000, and the best practice cost was $10,000, then the physician received

$$\frac{\$15{,}000 - \$12{,}000}{\$15{,}000 - \$10{,}000} = \frac{\$3{,}000}{\$5{,}000} = 60\% \text{ of the maximum payment}$$

Improvement Incentives: The improvement incentive was present to compensate physicians because reducing Part A expenditures should result in reduced Part B expenditures (or loss of income). These were defined separately for medicine and surgery. For medical specialists,

$$\textit{(base year ALOS – rate year ALOS)} \times \textit{per diem} \times \textit{rate year admission}$$

where ALOS is average LOS and

$$\textit{per diem rate} = \frac{\textit{medical improvement pool}}{\textit{rate year total days} - \textit{best practice days}}$$

For surgeons,

$$\frac{\textit{base year cost} - \textit{rate year cost}}{\textit{Xth percentile base cost} - \textit{best practice cost}} \times \textit{maximum payment} \times \textit{rate year admissions}$$

As stated in the BIMC application, the percentile of base cost was set to eliminate the outlier effect caused by patients with high utilization rates.

Improvement incentive algorithms differed between medical specialists and surgeons because surgeons control costs directly by ordering services from other doctors and are paid a fixed global fee; however, their fees are seldom affected. Medical specialists exert control over costs by determining the number of inpatient days. Shorter stays reduce hospital costs but also reduce physician fees.

Budget Neutrality Strategy: CMS was concerned that gainsharing may encourage physicians to change their inpatient discharge patterns, resulting potentially in increased overall PAC costs. This is of particular concern when gainsharing models, such as the one proposed by BIMC, focus on reduced lengths of inpatient stays. Should this occur, the demonstration would not be budget neutral. If patients are discharged sooner under the demonstration, Part B and outlier payments may fall, but PAC costs will likely rise. BIMC implemented a communications system with PAC providers to study patterns of post discharge outcomes. BIMC emphasized

strategies to reduce internal facility costs and Part B costs. BIMC's budget neutrality strategy included overall shorter inpatient stays, facilitated by conducting patient rounds on weekends, writing discharge orders early in the morning, and decreasing consultation waiting time. BIMC also planned use of fewer marginal diagnostic tests, a reduction in pharmacy expenses, and more efficient use of operating rooms. BIMC proposed more cost-effective use of critical care, evidence-based selection of medical devices, and avoidance of duplicative care. Finally, BIMC proposed to improve the quality and timeliness of medical records, which should have an overall impact on improved efficiency.

Medicare Cost Impacts: No savings to Medicare were required under this demonstration. BIMC initially proposed a trial year and did not guarantee budget neutrality during the trial year; 1 percent savings were expected in the trial year. BIMC reserved the right to terminate the program after the trial year with no financial penalties. BIMC expected in Year 2 to achieve 3 percent of (base hospital costs – inpatient costs) guaranteed savings (case-mix adjusted difference).

Quality Assurances: BIMC proposed a range of physician quality standards that, if not met by individual physicians, would make the physicians ineligible for the gainsharing bonus. These overall standards are as follows:

- Overall readmission rate within 7 days must not increase.
- Adverse events and malpractice experience must not increase.
- Physicians must comply with available quality measures.

BIMC also proposed to track patient complaints related to premature release, track readmission rates, and implement systematic communications with PAC providers to ensure that post discharge outcomes were not negatively impacted by the demonstration.

This page left intentionally blank.

SECTION 3
COMPARISON SITE SELECTION

3.1 Overview of Comparison Site Selection Methodology

The purpose of the Gainsharing demonstration was to isolate the impacts attributable to gainsharing arrangements in demonstration hospitals. The role of a comparison group was to represent trends in the major impact variables, which are then debited from observed demonstration hospital trends to produce final estimates of gainsharing impacts alone. The following describes the process RTI followed and summarizes the comparison sites selected.

Comparison hospitals were critical to both the budget neutrality analysis and the overall evaluation of the demonstration. The Gainsharing demonstration was required to be budget neutral (i.e., overall Medicare expenditures under the demonstration could not exceed projected costs in the absence of the gainsharing initiative). A trended-baseline methodology was specified by CMS to determine whether participating hospitals achieved budget neutrality.

In a trended-baseline methodology, average actual Medicare payments for the demonstration period are compared with a target. Average actual Medicare payments that are less than the target satisfy the budget neutrality requirement. The target is equal to a participant's actual baseline average payments, trended forward by the participant's expected growth rate. Each participant's expected growth rate is based on the actual growth rates in average Medicare payments for the comparison hospitals. Thus, by comparing spending growth in demonstration and comparison hospitals, demonstration sites are held harmless to growth trends in services such as PAC that are occurring in the absence of the demonstration. Aside from the budget neutrality assessment, the comparison hospitals are used to assess performance.

Basic to the process of the selecting comparison hospitals is that they be representative of or similar to the participants. There are a large number of characteristics from which to choose. We believe that the salient economic ones are

- the growth rates of the comparison hospitals that are used to assess attainment of budget neutrality by the participants,
- factors (e.g., graduate medical education, Medicare case mix, and disproportionate share of low-income patients) that influence both the level of Medicare payments and costs (selected factors should also influence the growth of payments and costs), and
- the competitiveness of the markets in which the participants are located.

These demonstration characteristics for the two sites are elaborated upon in ***Table 3-1***.

Table 3-1
Summary of key gainsharing site features for site selection purposes

Variable	Beth Israel Medical Center New York, New York	Charleston Area Medical Center Charleston, West Virginia
Diagnosis-related groups	All	Cardiac only
Medicare savings	Year 1: 1% Year 2: 3% (Year 2 guaranteed)	No Medicare savings proposed
Post-acute care budget neutrality strategy	Not presented	Reduced readmission rate
Physician payment incentive system	New Jersey system: improvement and performance incentives	Not presented
Internal hospital savings	Not presented	Year 1: 5% Year 2: 8% Year 3: 10%
Number of acute-care hospitals in market owned by parent organization	2 (main: Petrie; other: Kings Highway)	4 (main: General; other: Memorial, Women and Children's, CAMC Teays Valley)
Acute-care beds in main hospital	1,106	710
Other hospitals in local market	Many	3
Number of local hospitals in market that might be important rivals	Many	0

SOURCE: RTI International analysis.

Comparison hospitals should have growth rates that are representative of an attainable growth rate that will influence participant behavior. For instance, a participant's historical growth may be too high, and the participant may wish to have a lower growth rate. What are attainable growth rates? Because there are an infinite number of attainable growth rates, one way to specify attainable growth rates is to define them by the growth rates of comparable or peer hospitals. Comparable hospitals should be subject to cost structures and growth forces that are similar to those of the participants.

For participants located in markets with many other hospitals (e.g., BIMC), peer hospitals can be any of the nonparticipating hospitals in the local area because none of the individual hospitals in the market has much power to influence the collective growth rate of the market hospitals. Each hospital, if judiciously chosen, will have similar cost structures, which, in turn, will be subject to similar growth pressure. By contrast, participants located in markets in which they have near monopoly power (e.g., CAMC) can have great influence on the behavior or

growth rates of the other hospitals in their markets. Such participants can behave with little fear of competitive responses from the other local hospitals. Additionally, the cost structures of monopoly participants may differ greatly from those of other local hospitals. In such cases, the peer hospitals should be selected from external markets. These peer hospitals should be subject to the same growth pressures as the participants. They also should be in markets in which they have the same type of market power as the participants.

In selecting the actual comparison hospitals, a number of problems arise if nonrepresentative hospitals are selected. If the comparison growth rates are too high, participants can more readily attain the required budget neutrality. Conversely, comparison growth rates that are too low can result in bonuses that are not necessarily the result of improved care efficiencies at the hospital level. For this reason, the comparison hospitals should not be limited to just one or two hospitals. Instead, using a larger number of comparison hospitals will help to limit the influence of idiosyncratic factors (e.g., regression to the mean) from each hospital.

One effect of using growth rates in the budget neutrality assessment is to reduce the influence of levels of Medicare payments. Nonetheless, it is important to select comparison hospitals with cost structures similar to those of participants because cost structures can affect growth rates. The influence of cost structures on growth rates is probably more important in markets with many hospitals than in markets with few hospitals. Another reason to select comparison hospitals with cost structures similar to those of participants is that when growth rates are not observable, information on cost structures becomes the best predictor of growth rates.

Factors that can influence hospital cost structures include Medicare volume, the number of short-term acute care beds, the Medicare case mix index, graduate medical education (e.g., indirect medical education [IME]), and the share of low-income patients (i.e., disproportionate share hospitals [DSH]). An additional set of issues is related to growth rates in the assessment of budget neutrality. The Medicare Gainsharing Demonstration protocol indicated that payments for the participants and comparison hospitals should be standardized for Medicare case mix, gender, and age group. Differential changes in the area wage index, IME rates, and DSH rates may also differentially affect the growth rates of the participating and comparison hospitals and may need to be controlled for.

On the basis of this overall approach, RTI set the initial standards for comparison site selections that are shown in ***Table 3-2***.

Table 3-2
Approaches to selecting comparison hospitals

Hospital	Approach
Beth Israel Medical Center	Selected peer hospitals from the greater New York City area.
Charleston Area Medical Center	Selected large, dominant hospitals located in small urban areas. During 2006, the prospective comparison hospital must have performed at least 200 coronary artery bypass grafts or heart valves and at least 400 percutaneous coronary interventions, stents, etc. "Dominant hospital" is defined as one that has a local market share of at least 75% for one of these two sets of cardiac-related procedures.

SOURCE: RTI International analysis.

3.2 Selection of Specific Comparison Site Hospitals

Core Ranking Variables: The following core variables and methodology were used in selecting potential comparison hospitals for both CAMC and BIMC. Data were obtained from the 2008 Impact File.

- For each of the following variables, the absolute value of the potential comparison hospital's value minus the CAMC/BIMC value was calculated: Residents per bed, beds, residents, Medicare discharges, Medicare share of inpatient days, Medicare case mix, and operating DSH adjustment factor.

- For each of the above variables, the hospital that was closest to CAMC/BIMC received a rank of 1, the second closest received a rank of 2, and so forth.

- A weighted mean rank score was calculated for each comparison hospital. The weights used when creating each hospital's mean rank score were as follows:

  – Beds, Medicare share of inpatient days, and residents had a weight of 3 each.

  – All of the other variables had a weight of 1 each.

Hospitals with the lowest mean rank scores were those most similar to CAMC/BIMC. In creating this list of hospitals with the lowest mean rank scores, we were attempting to best reflect the factors associated with the growth of Medicare payments (all, not just IPPS) and the cost structure of the hospitals.

Charleston Area Medical Center—For CAMC's comparison hospitals, we selected those whose market dominance in cardiac surgery is similar to that of CAMC. To be considered a candidate comparison hospital for CAMC, a hospital must annually perform at least 200 major

heart procedures (i.e., coronary artery bypass grafts [CABGs] and heart valve procedures) and at least 400 percutaneous coronary interventions (PCIs). It must also have at least a 75 percent market share of either the major heart procedures or the PCIs.

Candidate comparison hospitals were ranked in terms of similarity to CAMC. Core ranking variables (described above) were used along with four additional variables: CABG/valve volume, percutaneous transluminal coronary angioplasty (PTCA)/stent volume, CABG/valve market share, and PTCA/stent market share. The weights of these four new variables were 6 for each of the cardiac volume measures and 4 for each of the cardiac market share variables. Ten of the 14 hospitals with the lowest mean rank scores were selected as comparison hospitals for CAMC. All but two of CAMC's comparison hospitals are located in the South. The application of the above criteria yielded the 10 comparison sites for CAMC shown in ***Table 3-3***.

Beth Israel Medical Center: BIMC is a large urban hospital with its main location in lower Manhattan. It is affiliated with an academic medical center and has a large resident program. It proposed to cover all DRGs during the demonstration. Because BIMC is located in a market in which it is but one of many hospitals, it will likely be subject to the same pressures on growth of Medicare payments as the other hospitals. To help select candidate comparison hospitals from the 52 other short-term, acute-care hospitals in the New York City area, we used data compiled from the core ranking variables (described above) to identify a potential list. Of the 52 New York City hospitals, we identified the 16 hospitals with the lowest mean rank scores and selected them as potential comparison hospitals for BIMC. Subsequently, three hospitals were removed from the comparison group because they either closed or merged with other hospital systems. The final 13 hospitals are shown in ***Table 3-4***.

Table 3-3
CAMC comparison hospitals

Medicare provider ID	Hospital name	City	State	MSA code	Mean rank score	Market hospitals	Acute-care beds	Medicare discharges number	Medicare discharges share	CABGs & heart valves, hospital volume	CABGs & heart valves, market volume	CABGs & heart valves, hospital share	PTCA & stents hospital volume	PTCA & stents market volume	PTCA & stents hospital share	DSH adj factor	No. of residents	Medicare inpatient share	Residents per bed	Medicare case mix index
—	—	weight:	—	—	—	0	3	1	0	6	0	4	6	0	4	1	3	3	1	1
510022	Charleston Area Medical Center	Charleston	WV	16620		4	718	13,824	62%	751	751	100%	1,120	1,261	89%	0.12	116	0.52	0.16	1.82
490024	Carilion Medical Center	Roanoke	VA	40220	7.6	3	664	13,381	66%	386	484	80%	1,066	1,308	81%	0.07	83	0.55	0.12	1.76
200009	Maine Medical Center	Portland	ME	38860	8.5	7	581	11,033	47%	424	424	100%	896	949	94%	0.08	171	0.46	0.30	1.95
340002	Memorial Mission Hospital and Asheville Surgery Center	Asheville	NC	11700	8.6	4	646	16,194	65%	571	571	100%	750	750	100%	0.13	39	0.53	0.06	1.79
440002	Jackson-Madison County General Hospital	Jackson	TN	27180	9.6	2	558	12,635	82%	326	326	100%	1,315	1,355	97%	0.16	18	0.54	0.03	1.74
010039	Huntsville Hospital	Huntsville	AL	26620	10.7	3	786	16,256	73%	359	359	100%	684	736	93%	0.07	31	0.48	0.04	1.66
340040	Pitt County Memorial Hospital	Greenville	NC	24780	11.5	1	618	12,619	100%	492	492	100%	749	749	100%	0.24	155	0.48	0.27	1.96
110107	Medical Center of Central Georgia	Macon	GA	31420	12.8	3	534	11,606	68%	493	598	82%	1,323	1,710	77%	0.21	88	0.46	0.16	1.92
440063	Johnson City Medical Center	Johnson City	TN	27740	15.5	4	478	10,734	77%	286	286	100%	755	755	100%	0.16	62	0.45	0.14	1.55
200033	Eastern Maine Medical Center	Bangor	ME	12620	15.6	2	302	8,388	76%	329	329	100%	658	659	100%	0.16	24	0.49	0.08	1.85
340141	New Hanover Regional Medical Center	Wilmington	NC	48900	15.9	3	539	13,331	84%	245	245	100%	563	564	100%	0.12	54	0.54	0.11	1.65

NOTE: CABG = coronary artery bypass graft; DSH = disproportionate share hospital; MSA = major service area; PTCA = percutaneous transluminal coronary angioplasty.

SOURCE: RTI analysis of CMS 2008 Impact File.

Table 3-4
BIMC comparison hospitals

Rank	Provider ID	Hospital name	Borough	Mean rank, score (weighted)	Residents per bed	Beds	Residents	Medicare discharges	Medicare share of IP days	Medicare case mix	Operating DSH adj factor
—	330169	Beth Israel Medical Center	Manhattan	—	0.35	994	349	12,914	0.39	1.39	0.37
1	330194	Maimonides Medical Center	Brooklyn	9.6	0.63	569	356	10,179	0.38	1.75	0.38
2	330236	New York Methodist Hospital	Brooklyn	11.5	0.44	495	217	7,841	0.39	1.57	0.30
3	330119	Lenox Hill Hospital	Manhattan	13.2	0.36	570	203	9,196	0.40	1.78	0.07
4	330055	New York Hospital Medical Center of Queens	Queens	13.8	0.38	439	168	9,295	0.39	1.57	0.24
5	330024	Mount Sinai Hospital	Manhattan	14.3	0.58	901	524	17,350	0.40	1.81	0.24
6	330214	NYU Hospitals Center	Manhattan	15.2	0.55	528	290	8,708	0.42	1.88	0.05
7	330306	Lutheran Medical Center	Brooklyn	15.8	0.63	322	203	5,261	0.38	1.44	0.40
8	330160	Staten Island University Hospital	Staten Island	16.5	0.37	557	204	8,265	0.29	1.53	0.21
9	330056	Brooklyn Hospital Center at Downtown Campus	Brooklyn	16.8	0.46	428	197	4,066	0.33	1.43	0.46
10	330195	Long Island Jewish Medical Center	Queens	18.2	0.79	578	459	9,077	0.32	1.70	0.14
11	330193	Flushing Hospital Medical Center	Queens	20.1	0.42	274	116	3,325	0.41	1.40	0.33
12	330233	Brookdale Hospital Medical Center	Brooklyn	20.2	0.53	455	240	3,343	0.25	1.48	0.52
13	330221	Wyckoff Heights Medical Center	Brooklyn	21.6	0.43	294	127	4,419	0.33	1.36	0.54

NOTE: DSH = disproportionate share hospital; IP = inpatient.

SOURCE: RTI analysis of CMS 2008 Impact File.

This page left intentionally blank.

SECTION 4
IMPLEMENTATION AND ORGANIZATIONAL RESPONSE: SITE VISIT AND FOCUS GROUP ANALYSIS

4.1 Charleston Area Medical Center

4.1.1 Approach and Methods

The RTI Medicare Gainsharing Demonstration team conducted a site visit to CAMC in Charleston, West Virginia, on September 29 and 30, 2010. The interview schedule was coordinated by CAMC staff on the basis of specific requests and discussion guides forwarded to CAMC by the RTI team. We conducted four interview sessions: (1) A discussion with the CAMC chief quality officer, (2) a session with the Medicare Gainsharing Demonstration project leadership, (3) a discussion with the demonstration financial officer, and (4) an interview with the nurse supervisors who coordinated the demonstration initiative on behalf of the CAMC nonphysician staff. Because CAMC withdrew from the demonstration as of December 2010, a second round of site visit interviews and physician focus groups was not conducted.

The interviews were supplemented by physician focus groups conducted during the same two days the RTI team was on site in West Virginia. RTI planned two physician focus groups, one each with participating and nonparticipating physicians. Because the number of demonstration eligible participating and nonparticipating physicians was small, CAMC staff assisted RTI in contacting and inviting all members of both groups to join a focus group discussion. In that recruitment process, invited focus group participants were provided with information on the purpose of the relative discussions. Seven participating physicians agreed to participate in a focus group, and four attended the discussion. Four nonparticipants were recruited to participate in a separate focus group, but none actually attended the session. Combined with the extreme difficulty in recruiting nonparticipants to attend the focus group sessions, we concluded that physicians with no stake in the demonstration, and with unpredictable hospital responsibilities, were not likely to participate.

Discussion guides were used by the leaders of both the interviews and focus groups to structure each session. Designated note takers recorded the feedback gathered in each session. Quoted words and phrases are from those notes.

4.1.2 Overview of Participation

Rationale for Organizational Demonstration Participation: A key topic of discussion throughout the site visit was CAMC's rationale for participating in the Medicare Gainsharing Demonstration. Feedback we received from both the interviews and the physician focus group suggested that CAMC staff saw the Gainsharing demonstration as a natural extension of quality improvement initiatives already under way to improve cardiac care performance relative to national standards.

CAMC staff described a number of ongoing quality improvement and cost savings initiatives. Before the demonstration, CAMC's approach was to identify cost savings and quality improvement opportunities, then share resulting savings with the relevant hospital departments, which could use the money for their priority projects (such as the acquisition of new medical

equipment). Other initiatives included distribution of gas gift cards for hospital staff (for example nursing and technical staff, but not physicians) in departments that achieved cost savings targets. Specific examples of these pre-Gainsharing demonstration cost savings initiatives included reductions in the waste associated with surgical supplies. These prior initiatives saved approximately $500,000 per year, with savings shared with participating hospital departments.

Other pre-demonstration quality initiatives were aimed at meeting or exceeding Society of Thoracic Surgery (STS) national standards for stroke and other co-morbidities. These were viewed as "soft" cost savings. CAMC's logic on these initiatives was that if patients stayed in the hospital longer, costs were added to the hospital. Because CAMC generally operates at nearly 100 percent occupancy, opening beds has a cost implication; the hospital doesn't have to send patients to other facilities. If the patients can be taken care of properly and can be released to the community, then beds can be opened up for new patients.

The Gainsharing demonstration was viewed by CAMC as a way to expand quality improvement initiatives and share savings directly with the physicians rather than just with departments.

Rationale for Physician Participation: One issue we discussed with both interviewees and the physician focus group was the rationale physicians used to elect to participate (or decline to participate) in the Gainsharing demonstration.

CAMC's leadership wanted to use the demonstration to engage a set of their physicians on a common set of goals. Our discussions suggested that relationships between the hospital and physicians may at times be somewhat strained, with physicians exerting considerable independence. For a hospital in a rural setting, competition for and retention of physician staffing is a complex issue. Although CAMC described a long history of quality improvement initiatives, leadership believed that adding a direct gainsharing component to these projects would take these existing initiatives to a higher level. It was also described as a way to engage physicians, particularly thoracic surgeons, toward a greater awareness of costs. A focus on heart procedures was chosen because CAMC was already investing in quality improvements in this area.

CAMC project management described the thoracic surgeon group as sometimes difficult to work with. Working relationships among the surgeons and between surgeons and the hospital have sometimes been strained. One goal of the demonstration was to give this disparate group a common goal.

Ultimately, CAMC was able to gain the participation of electrophysiologists, thoracic surgeons, and anesthesiologists, but not of the cardiologists, who described themselves as "uncomfortable with the setup." There was a perception among the cardiologist group that the quality improvement and cost reduction initiatives were "something we should be doing anyway." Cardiologists reported that they didn't want to be viewed by the community as being motivated by money rather than patient care.

Participants in the physician focus group believed that this demonstration could build bridges, and it did do this over time. The environment went from adversarial to more collaborative. However, the project took about 5 or 6 months to fully mobilize once approval was given by CMS. By the time gainsharing was fully implemented, CAMC had reached the end of the performance year. Physicians said that a year was not enough time.

Rationale for Discontinuing the Gainsharing Demonstration: CAMC started implementation of the Medicare Gainsharing Demonstration in December 2008 and officially operated the project for 12 months. Some feedback we received in the interviews suggested that full implementation of the specific gainsharing initiatives didn't occur until spring of 2009, suggesting that the project operated fully for less than 12 months. Although the authorization for the demonstration was eventually extended through Congressional legislative approval, CAMC elected not to continue participation past December 2009.

We were told by both interviewees and participants in the physician focus group that the Medicare Gainsharing Demonstration was always viewed as a 1-year project. CAMC participating physicians did not expect that the project would continue after December 2009 (and some were unaware that the project could have, been extended once the Congressional extension was enacted in the Affordable Care Act). CAMC demonstration leadership did acknowledge that they knew an extension was possible, but they argued that the organization had become increasingly concerned about its exposure to financial risk from possible changes in PAC utilization and spending. CAMC leadership told us that they were always concerned about risk exposure on both sides of the demonstration episode window. This was "something they didn't think they [could] control." Part of the CAMC concern was based on an acknowledgement that its local population in West Virginia tends to have an unhealthy lifestyle that might have affected pre- and post-acute care use. CAMC also did not have any contacts or working relationships with any of the PAC settings in town. CAMC leaders did not see a viable feedback mechanism with the PAC providers, and therefore they felt exposed to potential PAC risk, given the 30-day postdischarge episode period under the demonstration.

We were also told that some of the systems used to track data within the context of the demonstration turned out to be more resource intensive than was expected. There was a lot of manual input. There were substantial interface problems. CAMC staff found that the demonstration systems in the catheterization lab, operating rooms, and other departments all used different systems to capture quality-of-care and outcome data. The demonstration required additional data collection, but the demonstration-specific systems did not interface with the existing systems.

In the CAMC demonstration, all three administrative secretaries were burdened with data input. A lot of these data didn't seem to be useful to the clinical staff. Some required information to be pulled from medical records. If information wasn't entered correctly, then medical records were searched. By the time the process settled down and the modifications ended, the CAMC staff had developed their own plan. A few of the staff perceived that some of the impacts of CAMC gainsharing project were things that were in the works and would have happened regardless of the project's status.

In the end, CAMC leadership decided not to extend the demonstration because, in the words of one senior staff member, "the juice wasn't worth the squeeze." CAMC leaders were initially enthusiastic about the gainsharing model and were able to champion the model through implementation. But the work required to continue to model over time eroded that leadership support.

4.1.3 Cost Impacts

Sources of Internal Savings Generated by Gainsharing: CAMC originally envisioned interventions in which savings would be generated from both reductions in utilization and efficiencies in medical devices, both related to cardiac care. But once implementation began, leaders didn't feel that they could pursue both of these strategies, so they focused savings initiatives on efficiencies in medical devices. After CAMC submitted the application to CMS, for a number of external reasons, 2 years elapsed before approval was received. Some things weren't followed through because of the delay. A number of the peripheral vascular doctors did participate in the project, but once they got into the demonstration, they found that time and resources available to tackle issues were much more limited than they had anticipated. They didn't want to over-promise results. Some of the utilization and LOS issues were simply tabled because of lack of internal hospital resources to pursue them. CAMC did, however, continue to pursue internal quality initiatives.

CAMC hospital leadership told us that savings from devices is important and that having the doctors' cooperation and support is critical in getting more favorable prices from vendors. "Physicians want total choice when it comes to implants, joints, and other devices—they want 10 things on the shelf." Hospitals, by contrast, want to limit that choice and drive better bargains with vendors, the leadership said. The bulk of their effort in the demonstration went toward working with the doctors to come to agreement on more focused choices of devices.

Throughout the demonstration, the CAMC demonstration team worked with electrophysiology, cardiovascular surgery, and vascular intervention. Savings for all patients (not just Medicare) from the medical device efficiencies were estimated by CAMC as

- cardiovascular surgery, $110,000;
- electrophysiology, $282,000; and
- vascular intervention, $19,000 (some of the vendor contracts weren't finalized until late in the project).

CAMC also worked with cardiologists (most of whom did not participate in the Gainsharing demonstration) on some stent utilization and pricing outside the official demonstration. Savings there, about $421,000 as estimated by CAMC, were the result of existing initiatives that shared savings with departments rather than individual physicians. Both methods appear to work; CAMC's experience, according to the CAMC staff and given its particular environment, seems to suggest that there isn't a savings advantage to be gained just by sharing savings with individual physicians.

Overall Experience With Generating Savings: Originally, CAMC was interested in the following three strategies of potential savings:

- Reduced surgical preparation and operating room time
- Reduced use of supplies
- Lowered device prices

As reported by CAMC, the bulk of savings generated by CAMC through the demonstration ultimately came from negotiations with device vendors, with the help of the physicians. There were also some utilization savings, but these were limited. The CAMC staff did not feel pressure to make changes and generate savings if those changes led to poorer quality. Some of the physicians were simply not aware of the costs of some of the devices or of the alternatives and their costs. Physicians were instrumental in helping make these decisions and identifying opportunities for savings. CAMC always had a fairly good system in working with physicians, and this was another way to improve this relationship. A few physicians came to the table more consistently because of the dollars on the table, but others took the view that this was something they should have been doing all along. This was simply "a new adventure" and a new way to get physicians engaged.

At least one physician claimed that it was difficult to generate savings by reducing surgery and prep time—how to achieve savings by physicians didn't seem to have been well addressed during the demonstration. There was also considerable disagreement about the lack of yardsticks for surgery time. Product prices were posted on an internal CAMC Web site, but guidelines were not issued on product use. For instance, the guidelines did *not* urge physicians to use the least expensive devices.

Medical device product teams met monthly throughout the demonstration period. This process was in place before the demonstration and continued after the demonstration ended. Product teams (e.g., cardiac team, surgical team), which include physician representatives) meet monthly. Physicians who want to use new devices need to go through team reviews before they can use a new product.

Over time, CAMC sees keeping medical device prices down as a major challenge because the technology is constantly changing and increasing. In the vascular intervention area, Bard's life stent was a new technology that was adopted in part because some of the physicians were involved in a clinical trial. CAMC found that it got some traction with participating physicians on determining whether newer stents were always good choices, particularly in the third and fourth quarters of the demonstration as experience was gained with some of these newer products.

Process for Making Physician Gainsharing Payments: Our physician focus group suggested that there was some unevenness in the amount of information participating physicians had about the process for making gainsharing payments.

Some participants in the physician focus group were very unclear on the standards for bonus payments. For example, concerns were raised that CAMC participating physicians were never told whose pacemaker was cheaper and where more savings would occur, although hospitals knew this information. Physicians with this concern felt that if they were not part of a committee, then they did not know where they could generate the most savings. Physicians in some cases weren't forced to use any particular brand or type of medical device, but they weren't offered guidance, either.

Some focus group participant physicians were also unclear on how or why they received gainsharing payments. According to physicians interviewed during the site visit, they "received a check and were never told where they saved or where there was room to cut costs." The physicians asserted that they were never given feedback, only vaguely told that savings had to do with pacemakers and defibrillators. There was an overwhelming sense of disagreement between clinical and management staff.

Some physicians wanted to know how the payments were calculated. In general, the process took a little longer to generate payments for some physicians, but in general they didn't hear a lot of comments one way or the other. That said, physicians no concerns about the gainsharing payment algorithm ("it was always solid") and CAMC would potentially use it again.

4.1.4 Quality-of-Care Impacts

Demonstration Impact on Quality of Care: CAMC had a long history of experimenting with quality-of-care initiatives before the implementation of the Gainsharing demonstration. For the demonstration, CAMC set quality-of-care thresholds that had to be met as a condition of receiving gainsharing payments. Specifically, CAMC focused on CMS indicators (clinical conditions in heart product line, including acute myocardial infarction [AMI]indicators) and STS indicators. With regards to the STS indicators, CAMC had set a goal to improve its performance relative to the national benchmarks (defined in number of "stars") both before and after the Gainsharing demonstration, and this focus continues. CAMC describes its quality approach as "all or nothing": providers focus not on individual elements of quality standards but on a full set of interdependent quality indicators that affect patient quality of care.

CAMC leadership wanted to get physicians to work toward a common goal. They created a new quality infrastructure for the demonstration. CAMC feels that it did have some success on this goal; many physicians did work together, at least more than they had before. Under the Gainsharing demonstration, feedback on quality performance was given to physicians in the form of a report card showing performance on the quality metrics. Some physicians seemed to understand the report card information; others did not. CAMC leadership perhaps underestimated the need to work with physicians to help them understand and interpret the report card information. Quality-of-care measures were used under gainsharing as a threshold for payments, not as a basis for those payments to participating physicians. Emphasis on meeting quality metrics is continuing and CAMC will seek to accelerate a quality-of-care improvement program even in the absence of the demonstration. The demonstration "absolutely" helped bring hospital management and physicians together to improve quality. The demonstration also prompted the development of some new data collection systems that have had a positive and

continuing impact. Some physicians were especially motivated by the money; others (such as the cardiologists) were not. Improvements were made, but one physician noted: "it was not all kumbayah."

Physicians who participated in our focus group also generally agreed that quality of care was emphasized in the demonstration. However, there were some disagreements on whether sufficiently clear standards were set for the quality metrics. Participants reported that several committees were formed with different departments. These committees met and decided what they could achieve and identified a list of things. However, it took 5 or 6 months before they laid down what they were going to do, which was a long time. Focus group participants told us that the demonstration did not take postoperative complications into consideration. Some were also concerned that they were not told individually by the hospital how they were doing and how they could improve.

4.1.5 Patient Satisfaction

Our discussions with CAMC staff during the site visit interviews and focus group suggested that patient satisfaction monitoring was not a focus of CAMC's gainsharing project. There were no initiatives to monitor changes in patient satisfaction related to the Gainsharing demonstration at CAMC. There wasn't any sense among the CAMC staff we spoke to that patients were aware of the Gainsharing demonstration, that there was any feedback, or that patients were affected in any way. The CAMC Implementation Plan suggested that notices would be posted informing patients about the demonstration, providing a CAMC contact, but CAMC clinical and management staff disagreed as to this was actually done or whether any patients contacted CAMC in response to this notice.

4.1.6 Summary Comments

Overall, the CAMC staff we talked to were positive about the Gainsharing demonstration. Most of the staff involved in the project felt that the institution learned some important lessons about how to get physicians and hospitals to work together, how to control costs, and how to improve quality. They also learned the importance of having good data to monitor quality and costs, although developers of those data systems need to keep in mind the burden of data collection on staff.

- The positive results of the Gainsharing demonstration included involving more physicians in some meetings and in purchasing and quality initiatives than would have been expected otherwise. Increased awareness of costs was a definite plus from the project. This information is now more commonly disseminated.

- New initiatives reduce waste by tracking what materials and supplies are opened and not used. "The key is sharing the information. The key was not the gainsharing payments," said a demonstration leader. Some of the staff we spoke with went so far as to equate the savings and gainsharing payments to money that could be paid to get more staff or support existing staff; this did not, however, appear to be a majority opinion.

- The Gainsharing demonstration and the gainsharing payment made directly to physicians were not perceived to have been any more successful than CAMC initiatives aimed at improved quality and reduced cost. This perception may have been a factor in CAMC electing not to extend its demonstration participation past December 2009.

- CAMC raised a few operational and process concerns regarding the operation of the demonstration. CAMC did not view its experience working with CMS as a collaborative process, although CMS did not plan any joint efforts as part of this demonstration. Also, CAMC leadership would have liked for CMS to provide more clarification and data regarding projected PAC costs and risks associated with them. However, sharing these data with the sites was neither operationally feasible nor anticipated for the demonstration.

- CAMC suggests that medical device manufacturers place significant pressure on hospitals and physicians to have the newest devices, which are really driving up the diffusion of their products and prices. Medical devices are a major cost driver at CAMC and something that the Medicare Gainsharing Demonstration helped to address in a small way. On the other hand, a number of staff we spoke to believed that a better approach was to share savings with departments for their overall use rather than with individual physicians. The CAMC nursing staff we spoke with were not proponents of the bonuses. When one group gets an incentive and another doesn't, just because of the eligibility of the program, "there is animosity among the staff." On the other hand, some CAMC departments did pool incentives to fund department initiatives.

- Some CAMC staff suggested a single, bundled payment including the physician, hospital, and medical device costs. The rationale for this is that "the physician is the one who controls the device choice anyway. Physicians need to be held accountable for this choice." Holding hospitals accountable for the usage of medical device products can't be completely successful because hospital administration doesn't generally control these decisions. CAMC management perceived that "in some cases, the devices cost more than the total DRG payment"—and something that the Medicare Gainsharing Demonstration tried to address.

4.2 Beth Israel Medical Center

4.2.1 Approach and Methods

The RTI Medicare Gainsharing Demonstration team conducted site visits to BIMC in New York, New York, on October 13, 2010, and just over a year later on November 9, 2011. In both cases, the interview schedule was coordinated by BIMC staff on the basis of specific requests and discussion guides forwarded to BIMC by the RTI team. We conducted a combined interview session that included a range of administrative and medical leadership involved with the Medicare Gainsharing Demonstration, including the president (BIMC), the senior corporate vice president, senior vice president of medical affairs and chief medical officer, chief of general surgery, chief of graduate medical education, the medical directors for pay for performance and

quality management/patient safety, and the vice president for administration (BIMC Kings Highway). A representative from BIMC's support contractor, Applied Medical Software, Inc. (AMS), also participated. BIMC had some senior staff turnover in the time between these two visits. However, most BIMC hospital leadership remained active in the Gainsharing demonstration project and we were able to meet both times with all staff except the corporate senior vice president for managed care. These representatives were able to discuss the progress of the demonstration at both the Petrie and Kings Highway facilities.

These interviews were supplemented by two physician focus groups, conducted on the same day as the site visit. During each visit, RTI planned and conducted two physician focus groups of demonstration-participating physicians: one at the Petrie campus in Manhattan and the other at Kings Highway in Brooklyn. Because we were not able to recruit any physicians who did not participate in the demonstration's last year, RTI decided not to conduct a physician focus group with nonparticipating physicians from BIMC during the second site visit. BIMC staff assisted RTI in contacting and inviting all members of both groups to join the focus group discussion. Physicians were recruited to participate in the focus groups on the basis of a sample of participating physicians chosen by RTI staff. In that recruitment process, invited physician focus group participants were provided with information on the purpose of the discussions. In 2010, we were able to include 11 participating physicians at each campus. In 2011, we spoke with 9 participating physicians at the Petrie campus and 15 at Kings Highway.

Discussion guides were used by the leaders of both the interviews and focus groups to structure each session. Designated note takers recorded the feedback gathered in each session; quoted words and phrases are from those notes. Participants were told that they would not be identified individually in our write-up and that their responses would be kept confidential.

4.2.2 Overview of Participation

Rationale for Organizational Demonstration Participation: BIMC has been operating a gainsharing project with its non-Medicare fee-for-service and managed care insured populations since July 2006. Therefore, participation in this demonstration was an opportunity to extend to the Medicare population what the BIMC leadership perceives to be already a successful project that has LOSs and per-case costs within the hospital system. BIMC leadership got the idea about a gainsharing project from the cancelled NJHA Medicare demonstration from the 1990s. BIMC began gainsharing with commercial and managed care insurers and then submitted an application to CMS for a Medicare expansion of the project. There was some frustration at BIMC that CMS took such a long time to approve and allow implementation of the Medicare demonstration. However, a strong commitment to the model on the part of key BIMC institutional leaders kept the project moving forward despite the delays.

The BIMC gainsharing project is focused on reducing inpatient LOSs while maintaining or improving quality-related metrics. Specifically, BIMC pursues the following cost savings initiatives:

- shorter inpatient stays (primary focus)
- less use of marginal but costly diagnostic tests

- reduced pharmacy expenses (minimal focus)
- more efficient use of operating rooms and reductions in turnaround time
- cost-effective use of critical care and telemetry units
- evidence-based selection, purchase, and use of medical devices and hardware (minimal focus)
- eliminated duplication of services
- improved quality and timeliness of medical record and related documentation (also a quality threshold measure)

During the first and second site visits, BIMC leadership acknowledged that many of these separate measures related to the overall reduction in length of inpatient stays and agreed that this has been the primary goal of the gainsharing project.

Physicians who elected to participate in the project were measured against best practice (top 25 percent) norms for BIMC LOSs and other measures by applicable APR-DRG for a base year. BIMC leadership made a particular point of the importance of comparing physician performance to the top performers *within their own facility* because the methodology makes clear that the measures must be realistic and attainable (and are not simply a theoretical national standard that may not take into account the specifics and challenges of the local facility). Participating physicians did not actually receive their gainsharing incentive payments unless they met threshold quality-related standards. These threshold standards included medical chart completion and performance on a set of core medical process and clinical outcome measures. Participating physicians who met the threshold quality standards were eligible for a gainsharing payment based on the specific performance for their eligible cases.

In the BIMC demonstration, the first year of the demonstration focused on *improvement* relative to the best practice norms. BIMC leadership had originally intended to transition to part improvement and part performance standards by the second year of the demonstration but instead decided to continue making incentive payments for demonstrated improvement. By the third year, participating physicians were all measured against a *performance* standard alone, meaning that they no longer received any payments if they only showed improvement toward the best practice goals.

Rationale for Physician Participation: An issue we discussed with both BIMC leadership interviewees and the physician focus groups was the rationale physicians used to decide whether to participate in the BIMC Gainsharing demonstration.

BIMC leadership reported the following physician participation rates by the conclusion of the demonstration:

- Two-thirds (66 percent) of eligible Petrie campus physicians participated.

- Almost all (95 percent) eligible Kings Highway physicians participated.

In the second year of the demonstration, BIMC leadership reported the following physician participation rates:

- Three-quarters (75 percent) of eligible Petrie campus physicians participated.

- Almost all (90 percent) eligible Kings Highway physicians participated.

To be eligible, a physician needed to have had at least 10 cases in at least one medical specialty product line.

From the perspective of BIMC leadership, physician participation was never a significant challenge. Physicians were generally recruited through presentations at the department level. In 2010, follow-up meetings with individual physicians, particularly those with large caseloads at BIMC, were conducted as necessary. It was generally described to physicians as a "no-risk" project. BIMC found that after the early pre-Medicare gainsharing implementation resulted in actual incentive payments to physicians, participation rates rose.

In 2010, BIMC leadership also perceived that the gainsharing payments, per se, were not the primary motivator for physicians to participate; rather, competitiveness against their peers and toward personal improvement seemed really to motivate physicians. This perception persisted as the demonstration ended. Most incentive payments to physicians were modest; over the course of the project most physicians received a few thousand dollars per period, although a few outliers received large incentive payments as a result of large volumes. BIMC focused performance based on local (e.g., BIMC based) top performers rather than a national standard. Also, BIMC leadership feels that physicians want more explicit recognition for their efforts to improve quality and lower costs, and the gainsharing project is a way to accomplish that. In retrospect, BIMC leadership agreed that more one-on-one time with physicians both at the initial recruitment stage, and as follow-ups for currently and newly recruited participating physicians, improved participation rates as the demonstration progressed.

Our focus group discussions with physicians generally confirmed the views of BIMC leadership. Most participants saw the demonstration as a generally positive initiative, and none reported widespread reluctance to participate in gainsharing. Participants in the Petrie (Manhattan) campus felt that they had more exposure to hospital administration as a result of the demonstration (a positive outcome). A number of participants in this group also felt that the project had the positive effect of providing them with more feedback on their performance on LOS and costs relative to others at BIMC. Some physician participants in the demonstration thought that better aligning the physician and hospital incentives to lower costs was a good thing in the era of health care reform.

By 2011, BIMC leadership followed through with conducting more one-on-one meetings with physicians. BIMC found these meetings to be a critical strategy and opportunity to review the physicians' performance. The leadership used a "dashboard" document as the foundation for their conversations. The BIMC dashboard, prepared by the BIMC support contractor AMS, provided quick statistics about LOS for individual physicians compared with local best practice

norms. It also showed the incentive earned, the unearned incentive, and the maximum incentive that could have been earned. Along with these physician-specific statistics, the dashboards contained bar graphs with a LOS summary, cost summary, and a top cost summary for the physician compared with the best practice norm. The physicians at Petrie maintained that the dashboards were confusing and hard to follow. One physician noted the reports "are not very real time," noting that looking back at data from 6 months to 1 year ago is not useful. At the Kings Highway campus, however, physicians we met with told us that the dashboards were "pretty easy to follow."

Although BIMC leadership increased the number of one-on-one meetings they held with individual physicians between our first and second site visits, they found that meeting regularly with all participating physicians was not practical. As a result, hospital leadership told us that they tried to meet with high-volume admitters twice a year. These personal sessions seemed to improve the relationships the physicians had with the hospital. BIMC leadership seemed to agree that an incentive is necessary for physicians to change their behavior, but there was some disagreement regarding whether the size of the check mattered. BIMC leadership noted that the importance of the amount of the incentive earned may correspond to the average salary of the individual physicians; they believed that the financial incentives may be more powerful for hospitalists relative to surgeons (for example) because the former earn substantially less money on average. Also, some of the BIMC hospital leadership continued to believe that the competitive nature of physicians was sometimes sufficient to prompt behavioral changes. BIMC leadership cited a number of instances when individual physicians argued over relatively small incentives and were clearly distressed to learn that their missed incentives were greater than those of their peers.

In both rounds of focus groups, physician perspectives generally confirmed the views of BIMC leadership. Physicians participating in gainsharing clearly believed that the Gainsharing demonstration helped contribute to a team approach and improved relationships with the hospital. They continued to see the demonstration as a generally positive initiative without any real downsides. A number of the physicians at Kings Highway spoke about receiving increased support from the hospital, such as adding more physician assistants and social workers. One physician at Kings Highway said, "I had nothing to lose. I was told to sign the paperwork, and I did." Across the two campuses, physicians seemed to agree that "more feedback" on their performance is a good thing.

4.2.3 Cost Impacts

Sources of Internal Savings Generated by Gainsharing: BIMC's gainsharing model focused on reduced lengths of hospital inpatient stays and overall reductions in costs for inpatient stays. Participating physicians received at first quarterly—and later in the demonstration semi-annual—comparisons of their costs relative to the top 25th percentile of other BIMC physicians within the same APR-DRG. Additional initiatives also sought to reduce costs on the basis of reductions in medical device expenditures and other sources, but feedback from both rounds of the site visits suggested that reduced lengths of inpatient stays were clearly the primary focus of the demonstration.

Overall Experience With Generating Savings: After the first year of the demonstration, BIMC leadership felt strongly that the overall gainsharing model had significantly reduced LOSs, and therefore costs, at both the Petrie and Kings Highway campuses. Because many of the Kings Highway physicians also have privileges at other hospitals, the BIMC leadership had to do more "hand holding" to explain the demonstration to physicians and recruit them, but that effort has paid off.

According to BIMC leadership, some aspects of BIMC's original savings plan did not prove as successful as others. Both the BIMC leadership and participating physicians told us that savings on medical devices and reduced pharmacy services have not been nearly as much a focus as reduced LOSs. The implication was that these strategies were simply not vigorously pursued rather than that they were unsuccessful per se.

BIMC leadership did not cite specific savings during the first site visit, but they strongly believed that significant savings had been achieved. In an article published in 2010 by the BIMC leadership, savings of $25.1 million over 3 years (only the last of these included the performance year) and payouts to physicians of more than $2 million were reported (Leitman et al., 2010).

The direct connection between the information provided in the performance reports and changed physician behavior leading to lower costs was somewhat unclear in our discussions. Most physicians we spoke with in the focus groups did not understand, or did not take the time to focus on, the performance measures reported. One physician stressed multiple times that he participated in the project because "no one was telling me I had to change how I practice." Still, upon probing, the message of an overall reduction in costs and LOSs was clearly communicated.

As of our second site visit, monitoring reports provided by ARC cited that BIMC had achieved $31 million in internal savings, according to BIMC's methodology for calculating savings. BIMC leadership still felt strongly that the overall gainsharing model had significantly reduced LOSs, driving these internal cost reductions at both the Petrie and Kings Highway campuses. In addition to reducing LOS, BIMC leadership also reported that they achieved internal cost savings in more efficient operating room utilization and more efficient use and purchasing of implant devices. BIMC hospital leadership reported an improved understanding among the physicians regarding their performance and impact on costs over the course of the demonstration. Although some physicians expressed discomfort with the idea of changing their behaviors because of financial incentives, physicians we spoke with at both the Petrie and Kings Highway campuses were nonetheless able to point out changes they have made in their practice behaviors since the beginning of the demonstration (for instance, completing medical charts and reducing patient LOS).

Process for Making Physician Gainsharing Payments: BIMC leadership reported no particular problems in the process of making gainsharing payments at our site visit in either 2010 or 2011. For BIMC, the actual calculation of the physician incentives is done with the support of a contractor, AMS. BIMC administration staff commented that they didn't think they could make these calculations on their own. Because of the administrative burden involved in calculating and distributing the gainsharing payments, BIMC moved to a semiannual rather than quarterly basis for gainsharing payments. That change was not completely well received among physicians, who felt that this change created too much of a delay between the actual behavior

and the feedback. By 2011, we did not hear concerns regarding the timing or process for incentive payments, other than the general comment that the time lag between the patient data and reporting and the incentive payments was too long. At Petrie, the feedback from physicians regarding the payment of physicians incentives varied somewhat by specialty. Some departments at the Petrie campus pooled incentive payments because physicians take care of each other's patients. Other departments reported to us that they unofficially pooled incentive payments to hire additional support staff or "finds other ways to make physicians' lives easier."

In 2010, our physician focus groups suggested that there was some unevenness in the amount of information the participating physician had or understood about the process for making gainsharing payments. Many physicians seemed to be at least somewhat confused about the basis by which they were judged. BIMC leadership conceded that they probably needed to do more ongoing training and hold more discussions (perhaps one on one) than they had done since the initial recruiting period. By 2011, we did not hear this concern in our physician focus groups.

Across the focus groups at the Petrie and Kings Highway campuses, a number of physician participants told us that, although the feedback on their performance was a positive aspect of the project, the actual reports received were "incomprehensible." One participant commented, "Just reading the document is not possible. It isn't clear that it was designed to be read, but just a dump of information that couldn't be interpreted." Others felt that, whereas gainsharing was supposed to be used as a "carrot" to improve quality and lower costs, the information they received tended to come too late to change behavior within a gainsharing incentive payment period: "They tell you only after the fact whether you qualify or not. They keep all the money and it becomes more of a penalty." Some additional comments that reflected viewpoints on understanding among physicians about the gainsharing metrics included these:

- "Some of the LOS targets seemed unrealistic. We were told it was the top 25th percentile of our peers, but we were skeptical about it."

- "You can lose an entire bonus payment by missing the mark on one indicator, when you hit the mark with all other indicators. One physician described loss of incentive payments because he did not meet minimum quality and threshold standards as: "It felt like one of those 'gotcha' penalties."

- "I get this mysterious check in the mail on occasion, and it comes with these reports with nice charts, but I don't understand it. It's confusing and time-consuming to take the time to understand the reports or validity of them."

The difference in these responses seemed to be related to medical specialty. Participating physicians in the medicine specialties seemed to have a greater understanding of the reports, the threshold quality measurement, and the methodology by which gainsharing payments were (or were not) made. For example, one physician in the Medicine department commented, "We pretty much understand the metrics in our area. It was explained well to us as a group in various meetings when the program got started. In medicine, they were corralled as a group." These perceptions tended to persist in 2011.

During our second site visit, we noted a distinct increase in the overall understanding of gainsharing incentive payment model among the physicians. We attribute this change to the strategy of one-on-one meetings between BIMC leadership and high-volume participating physicians. Although we still heard some confusion about the incentive payment process from participating physicians in our focus groups, the level of information had clearly increased as the demonstration progressed. In our 2011 physician focus groups, there was a greater level of enthusiasm about the gainsharing model as a way to appropriately recognize physicians who were performing well. This contrasted with our 2010 physician focus groups at BIMC, where we were more likely to hear from physicians that the gainsharing initiative was viewed as punitive. At Kings Highway, the physicians said that they were making changes that they had never been told to make before. One physician remarked that he was denied some of his incentive in an earlier round, so he asked the leadership why this was the case. As a result, he corrected those issues and received a larger check in the next round. Another physician told us that he had reduced his patients' LOS and had identified a number of instances where it was clinically appropriate to change medications from intravenous to outpatient delivery. Yet another physician told us that he had not received the maximum incentive payment as a result of missed documentation of a patient pneumonia shot. After that experience, this physician routinely documents pneumonia shots, something he had not troubled to pay attention to in the past.

4.2.4 Quality-of-Care Impacts

During the first site visit, BIMC leadership described their gainsharing initiative as primarily focused on cost savings. The quality-of-care component is certainly important, and BIMC leadership coordinated the Gainsharing demonstration with their ongoing quality-of-care initiatives. Quality-of-care measures were used under gainsharing as an eligibility threshold for the receipt of potential incentive payments, but they were not used to compute the size of payments to participating physicians. The demonstration also prompted the development of some new data collection systems, which was a positive, and continuing, impact. Examples of the mandatory quality threshold measures include

- documentation (quality and timeliness of medical records),
- timely medical records (no more than 5 records more than 30 days delinquent),
- consultation with social workers or discharge planners (within 24 hours of admission),
- decreased or controlled (less than 5 percent) unanticipated returns to the operating room, and
- decreased or controlled (less than 1 percent) inpatient mortality.

Documentation and timely medical records seemed to get particular emphasis, at least according to the examples cited by physicians we spoke with in the focus groups. Failure to meet one or both of these two threshold measures was also the most common reason that physicians did not receive incentive payments.

In 2010, some of the physician participants did not describe a clear change in their behavior as a result of the specific performance reports; others did. The perception that the Gainsharing demonstration actually changed physician behavior at that point in the project was mixed. BIMC moved from a 3- to a 6-month performance period for gainsharing incentives, which lengthened the time between the physician behavior and the incentive payments. One physician participant commented, "The infrequency of the feedback is not going to change behavior, which is what the hospital wants to do." Another participant commented, "I don't see the relationship between the feedback and changing my behavior." Another participating physician commented that he was willing to participate in gainsharing specifically because he was *not* required to change his behavior: "No one was going to tell me how to practice." That said, this same physician did also acknowledge that information on how his costs and LOSs compared with those of top BIMC performers was of interest to him.

By our second site visit in 2011, we noted a clearer push to look at gainsharing through a quality improvement lens, particularly among the BIMC leadership. The hospital's primary cost savings strategy continued to be reductions in LOS, which BIMC continued to see as the area of care delivery with the largest opportunity for improvement. In their efforts to reduce LOS, BIMC leadership targeted several quality areas, including improvement in medical chart completion, discharge instructions, and appropriate antibiotic use. The increased focus on quality improvement was also noted by participants in our physician focus groups. Among Petrie campus physicians, the specific emphasis on quality of care appeared to vary by department. At Kings Highway, one of the physicians mentioned that shorter LOS means fewer hospital-acquired conditions. While some physicians we spoke with (both in 2010 and again in 2011) expressed discomfort with the idea of getting paid for actions physicians should be doing anyway, many physicians at both the Petrie and Kings Highway campuses were able to point to behavioral changes. In all of our focus groups, participating physicians at both campuses adamantly maintained that they would not take part in anything that would endanger their patients and were clear that they felt no pressure to do so under the demonstration.

4.2.5 Patient Satisfaction

Our discussions with BIMC leadership during our first and second site visits found that patient satisfaction monitoring was not a focus of their gainsharing project. There were no initiatives to monitor changes in patient satisfaction related to the Gainsharing demonstration at BIMC. BIMC leadership had no sense that patients were aware of the Gainsharing demonstration, or that they were affected in any way.

4.2.6 Summary Comments

Overall, the BIMC staff we talked to during our first and second site visits felt that the Gainsharing initiative was successful. The BIMC leadership were particularly positive about gainsharing as a model for the future. They hope that the gainsharing model will continue under the Medicare program, and they plan to continue gainsharing for private payers.

- BIMC leadership initially believed that they had not previously identified the right set of tools to modify physician behavior. By 2011, BIMC felt that they had improved and focused physician feedback, particularly by meeting regularly and one on one with physicians (particularly high-volume participants).

- Over the course of the project, BIMC leadership concluded that provision of data on individual performance alone will move some physicians to improve, but something else—incentive payments—is needed to change behavior among other physicians. By 2011, BIMC leadership continued to feel that the amount of the incentive payment was not the critical factor in changing behavior for most physicians; rather, the feedback and clear communication that tapped physicians' desire to improve and perform well relative to their peers was the most powerful factor. That said, in both our visits, we did find that community-based physicians (primarily at Kings Highway) reported that the incentive payments were important to them and that they definitely contributed to improved relationships between the hospital and physicians.

- Throughout the projects, some physicians told us that they "felt uncomfortable" being paid for actions they should be taking anyway. Their discomfort didn't appear to have significant effects on the recruitment for participation in the project.

- BIMC leadership were not concerned about the PAC risk created by the 30-day PAC episode window either at the start of the demonstration or at its conclusion. Although they were not monitoring PAC expenditures, BIMC leadership simply believed that improving care and lowering LOSs would not lead to PAC problems (i.e., increased PAC payments relative to comparison hospitals).

- One-on-one meetings between BIMC leadership and individual physicians regarding their performance were critical in opening communication and interaction between physicians and the hospital. Physicians reported that these meetings were a significant improvement over the original approach of relying mostly on the dashboard document for reporting performance. Physicians and BIMC leadership agreed that one-on-one meetings improved understanding of performance data. This approach, however, is resource intensive for BIMC leadership.

- Physicians and BIMC leadership maintained that, in order to incorporate the principles of gainsharing on a wider scale, a basis for cultural change is needed. Implementing gainsharing worked well for BIMC because the demonstration was a natural extension of the quality improvements that were already a focus.

- Age seemed to play a role in physician willingness to participate. Younger physicians who were just starting out in this field seemed to have more reason to accept incentives and also seemed more open to change overall.

4.3 Comparison of Findings

The two Medicare Gainsharing Demonstration sites, CAMC and BIMC, implemented different gainsharing methodologies, each with a different clinical focus. For these reasons, direct comparisons between the findings of these two different sites should be made with caution. Still, some common themes that emerged from our site visits and physician focus groups may point to lessons learned about the overall gainsharing model. These common themes are as follows:

- *Hospitals and Physicians View Gainsharing as a promising model:* Overall, individuals we spoke with at both CAMC and BIMC felt that the gainsharing model as a whole was a promising way to improve the way care is delivered. Although each site's model was different, individuals at both sites felt that the general idea of gainsharing was a way to improve physicians' awareness of cost control. Both sites felt that the overall gainsharing model improved communication between physicians and hospital administration on issues related to lowering costs and maintaining or improving quality of care.

- *Participating physicians don't consistently understand cost and quality reporting:* Although the sites had different clinical focuses, both provided detailed quality-of-care and cost performance reports to participating physicians as a way to substantiate the payment (or nonpayment) of gainsharing incentives. However, our site visits and focus groups found that many physicians didn't understand the reports, although their understanding had improved by our second site visit to BIMC. Across both sites, we heard comments that the physician reporting was, for example, "incomprehensible" or "overly complex," despite the efforts of hospital administrators to explain the basis for gainsharing incentive payments. One-on-one meetings between hospital leadership and physicians, as implemented by BIMC, seem to be a promising way to improve this understanding.

- *The perceived direct link between the actual amount of gainsharing payments and cost and quality-of-care improvements was not always clear to the participants:* Leadership in both sites felt that the overall gainsharing model was successful, but mostly in improving the communication between hospitals and physicians and in improving awareness about the need to lower costs and maintain or improve quality of care. Less clear was their belief in the link between the specific metrics and data provided to individual physicians and any successful (or unsuccessful) changes in cost and quality trends.

- *Success of a gainsharing model depends on a strong organizational "champion":* Both sites discussed the substantial level of effort required to implement and maintain a gainsharing model. Both sites also stressed that making gainsharing work in practice requires an internal champion within the implementing organization.

SECTION 5
MEDICARE EXPENDITURE AND SAVINGS ANALYSIS

One element of the Medicare Gainsharing Demonstration evaluation is an analysis of changes in Medicare expenditures (and any associated savings) that may be attributable to the gainsharing intervention. A primary strategy of the gainsharing model was a reduction in inpatient length of stay (LOS) and associated costs. Although Medicare savings were not required under this demonstration, CMS is interested in determining whether any changes occurred in utilization and subsequent Medicare expenditures—and therefore may have produced net program savings. This section presents the complete performance years' analysis of Medicare expenditures and LOS for demonstration and comparison sites.

5.1 Data Sources and Measures

Medicare inpatient claims were obtained for the baseline calendar year (2007) for both the demonstration and comparison hospitals using CMS's Data Extraction System. Performance Years 1 through 3 data were subsequently obtained. For CAMC, the single performance year was December 1, 2008, through December 30, 2009. For BIMC, Performance Year 1 (PY 1) was defined as October 1, 2008, through September 30, 2009, Performance Year 2 (PY 2) was defined as October 1, 2009, through September 30, 2010, and Performance Year 3 (PY 3) was defined as October 1, 2010, through September 30, 2011.

Cross-referencing was performed to obtain all health insurance claims assigned to each beneficiary. RTI then made the claims to available to ARC. Using the claims in conjunction with demographic and enrollment data in the Medicare Enrollment Database, ARC determined beneficiary eligibility requirements for beneficiaries represented in the potential inpatient data. ARC identified an *index* IPPS hospitalization that was used as the basis for constructing each expenditure episode. Finally, for each demonstration site ARC created an analytic file that included a set of episode claims for the demonstration and comparison hospitals for the 2007 calendar year baseline period and the relevant intervention years. Claims for all DRGs that were covered under the BIMC demonstration were included in this analytic file. After creation of the beneficiary episode file, both RTI and ARC used this same analytic file to ensure data consistency among the various analyses for this demonstration.

Medicare Episode Expenditure Measures: Episodes included all Part A and Part B fee for service-related health care services within the 14-day pre-admission and the 30-day post-discharge windows defined by the demonstration protocols. Beneficiary co-payments are excluded. No initial adjustments were made for local area differences in Medicare payment rates (e.g., different wage indices and IME add-ons). We relied on regression methods to control for factors affecting Medicare payment rates. Although the 2008 and subsequent IPPS inpatient claims were paid under the recently implemented Medicare Severity DRG (MS-DRG) system, each of the inpatient claims was grouped using the older DRG system grouper (version 24).

For their analyses, ARC truncated the distribution of episode expenditures (Medicare payments) at the 95th and 5th percentiles using a DRG weight, tier-normalized outlier truncation

methodology.[4] RTI did not employ ARC's methodology of capping expenditures because we were interested in the composition of expenditures and did not want to bias the analysis against the study hospitals if they were constraining expenditures of the sickest beneficiaries.

In preliminary descriptive analyses, we grouped Medicare payments in the baseline and intervention years by pre-admission, index hospital, and post-discharge periods and by type of health service. Next, four types of episode payments (expenditure) variables were constructed:

- total episode payments—includes payments to all providers in the three periods, as specified in the *Budget Neutrality Analysis Reconciliation Payment* protocol
- episode payments excluding Medicare's inlier[5] and outlier payments to the index hospital
- episode payments excluding only the fixed inlier DRG payment to the index hospital
- episode payments for only the 14-day pre-admission and 30-day post-discharge periods

Table 5-1 summarizes these four types of episode payment measures. The total episode payment definition is the most inclusive of the four measures and was the definition used by ARC in its analyses. The second episode measure excludes the two forms of DRG payment (inlier and outlier) made to the index hospital. The third measure isolates the outlier payment from the inlier payment. That is, since it is not usually possible to reduce IPPS inlier payments to the index hospital, reductions in IPPS outlier payments is the major way to reduce IPPS payments to the index hospital. The fourth measure of episode payments includes Medicare payments for health services provided only during the pre-admission and post-discharge periods. The purpose of this measure is to determine whether efforts by demonstration hospitals to reduce costs might have led to increased payments for health services provided during the pre-admission and post-discharge periods. Although not examined in detail, lower inpatient costs might result in greater pre-admission testing and in discharging patients "quicker and sicker." For the most part, it was not possible for participating hospitals to reduce IPPS inlier payments because Medicare pays a flat DRG-based amount regardless of resources used by the hospital.

Cost-Related Measures: Internal cost savings were estimated by the demonstration sites (BIMC had the support of a contractor, AMS) and were validated by ARC. To determine whether Medicare claims could also detect cost reductions, we examined three cost-related

---

[4] ARC classified claims into five tiers on the basis of the DRG weight. Within each tier, expenditures below the 5th percentile were re-coded to the 5th percentile's value and expenditures above the 95th percentile were re-coded to the 95th percentile's value.

[5] The basic IPPS payment to the index hospital is the *inlier* portion of total IPPS payments to the index hospital. The basic IPPS payment is based on a formula incorporating standardized labor and nonlabor amounts, DRG payment weight, IPPS area wage index, indirect medical education (IME) adjustment (if any), and disproportionate share (DSH) adjustment (if any). The Gainsharing demonstration did not change any of these components.

measures: LOS of the index hospitalization, IPPS outlier payments for the index hospitalization, and payments for physician services provided during the index hospitalization. LOS was chosen, in part, because, reducing it was one of the primary strategies employed by BIMC to achieve internal cost savings. Reduced LOSs should result in lower nursing costs and, possibly, reduced diagnostic testing. Shorter stays can also reduce infections and other costly complications (e.g., pressure ulcers).

Table 5-1
Four expenditure measures

Payment component	(1) Total episode payment	(2) Episode payment excluding index IPPS hospital payment	(3) Episode payment excluding only inlier DRG payment	(4) Episode payment for pre-admission & post-discharge periods
14 day pre-admission	X	X	X	X
Index hospital inlier DRG	X	—	—	—
Index hospital outlier	X	—	X	—
Index hospital physician	X	X	X	—
30-day post-discharge	X	X	X	X

NOTE: DRG = Diagnosis-related group.

IPPS outlier payments are based on excess hospital costs—costs that might be influenced by physician behavior. IPPS outlier payments for the index hospitalization were examined on a per index hospitalization basis and deconstructed into two parts: incidence of outlier status and IPPS outlier payments for only those index hospitalizations with an outlier payment.

Both CAMC and BIMC shared internal cost savings with its participating physicians. These rewards were based on the premise that changes in physician behavior were responsible, in part, for these internal cost savings. If this premise was correct, then it might be possible to detect changes in physician behavior by examining whether there were changes in physician payments for services rendered during the index hospitalization.

5.2 Methods

Multivariate analysis was conducted using ordinary least-square (OLS) regressions on the four Medicare expenditure measures, the beneficiary's index hospital LOS, the beneficiary's outlier payments, and inpatient physician payments. IPPS outlier payments were examined using two approaches: (1) IPPS outlier payments per index discharge; and (2) a two-part, first-stage logistic model on the likelihood of incurrence of IPPS outlier payments in the index hospitalization followed by a second-stage OLS regression on IPPS outlier payments conditional on exceeding the outlier threshold.

In testing for Gainsharing demonstration effects on Medicare expenditures, LOS, and outliers, as well as on quality of care and patient safety in Section 6, we used the following difference-in-differences (2D) approach:

$$Y_{i,t} = \alpha + \beta \cdot X_{i,t} + \gamma_1 \cdot D_{i,t} + \gamma_2 \cdot T + \gamma_3 \cdot T \cdot D_{i,t} + \varepsilon_{i,t}, \quad (5\text{-}1)$$

where

Y	is a dependent variable (e.g., Medicare episode payment),
D	is a binary variable where 1 denotes an episode starting at a demonstration hospital and 0 denotes an episode starting at a comparison hospital,
T	is a binary variable where 0 denotes an episode in the base period and 1 denotes an episode during the demonstration period (PY 1 through PY 3),
$T \cdot D$	is an interaction term between D and T used to estimate the 2D effect,
X	is a vector of beneficiary and hospital characteristics,
ε	is an error term, and
i and *t*	are subscripts used to denote an i episode during time period t

The Gainsharing demonstration effect on changes in payments and other outcomes is measured by the interaction term $\hat{\gamma}_3$. For per-episode Medicare payments, positive values of $\hat{\gamma}_3$ indicate that per-episode payments are rising faster at participating hospitals than at comparison hospitals. Conversely, negative values of $\hat{\gamma}_3$ indicate that per-episode payments are rising more slowly at participating hospitals. Negative estimates of $\hat{\gamma}_3$ indicate that the Gainsharing demonstration was able to slow the growth in Medicare payments per episode. The $\hat{\gamma}_1$ coefficient represents the mean difference between participating and comparison hospitals in the base period after controlling for beneficiary and hospital differences. The $\hat{\gamma}_2$ coefficient estimates the growth in the dependent variables (e.g., episode payments) for comparison hospitals between the baseline and first performance year.

Although the comparison hospitals were matched on several characteristics and therefore should play little role in explaining differential rates of episode growth between participating and comparison hospitals, we did control for many of the same characteristics in the multivariate model. This has the effect of factoring out variation due to these variables and improving the precision of the models' estimates. The individual βs for the beneficiary and hospital characteristics are interpreted as the marginal effect of a unit change in a specific characteristic on payments or other outcome variables. Because the impact model (Equation 5-1) is estimated in linear form, coefficients for hospital characteristics (e.g., DRG weight) are unrealistically high. This is because Medicare discharge payments are based on a multiplicative formula. For the inlier payment, a one-unit increase in DRG weight results in a greater percentage increase if DRG case mix is correlated with, say, interns and residents per bed.

To determine whether there were year-specific demonstration effects, regressions were also estimated using an alternative specification (BIMC only). In this case T, where a value of one represented the entire demonstration period, is replaced by a set of year-specific demonstration period indicators identifying PY 1, PY 2, and PY 3 (The base period is part of the constant term, α). Similarly, the T·D 2D estimator is replaced by a set of year-specific 2D estimators: PY1·D, PY2·D, and PY3·D. The coefficients for the year-specific 2D estimators show the difference-in-differences between the base year and the specific performance year for BIMC.

Tables 5-2 and ***5-3*** display the mean values for all explanatory variables for CAMC (base year and PY1) and BIMC (base year and PY 3) respectively.

Table 5-2
CAMC means of explanatory variables by group

Name	Base year comparison	Base year CAMC	Performance Year 1 comparison	Performance Year 1 CAMC
Patient age 0–64	0.130	0.184	0.121	0.201
Patient age 65–69	0.200	0.211	0.210	0.206
Patient age 70–74	0.229	0.219	0.218	0.199
Patient age 75–79	0.215	0.194	0.213	0.202
Patient age 80 or more	0.226	0.193	0.239	0.192
Female	0.397	0.354	0.401	0.388
Nonwhite	0.091	0.026	0.091	0.016
DRG weight	4.287	4.294	4.330	4.244
HCC risk score	1.520	1.562	1.467	1.484
IPPS area wage index	0.926	0.869	0.927	0.830
Intern-/resident-to-bed ratio	0.135	0.170	0.132	0.164
Hospital beds	554	725	583	718
DSH adjustment factor (operating)	0.141	0.125	0.147	0.146
Average LOS for DRG	7.277	7.756	7.451	7.480
Number of observations	6,111	882	5,621	855

NOTE: CAMC = Charleston Area Medical Center; DRG = diagnosis-related group; DSH = disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system; LOS = length of stay.

SOURCE: RTI processing of Medicare claims.

Table 5-3
BIMC means of explanatory variables by group

Name	Base year comparison	Base year BIMC	Performance year 3 comparison	Performance year 3 BIMC
Patient age 0 to 64	0.117	0.127	0.125	0.137
Patient age 65 to 69	0.122	0.113	0.129	0.116
Patient age 70 to 74	0.148	0.136	0.148	0.140
Patient age 75 to 79	0.173	0.170	0.158	0.151
Patient age 80 plus	0.440	0.454	0.440	0.456
Female	0.593	0.614	0.583	0.601
Non-white	0.283	0.307	0.293	0.316
DRG weight	1.570	1.341	1.605	1.358
HCC risk score	2.232	2.272	2.295	2.297
Intern/Resident to bed ratio	0.543	0.347	0.532	0.425
Hospital beds	563	994	592	886
DSH adjustment factor (operating)	0.279	0.385	0.237	0.346
Average LOS for DRG	5.403	5.005	4.994	4.702
Number of observations	66,785	8,913	67,336	9,705

NOTE: BIMC = Beth Israel Medical Center; DRG = diagnosis-related group; DSH = disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system; LOS = length of stay.

SOURCE: RTI processing of Medicare claims.

Although age, gender, and race may indirectly capture the impact of patient health and access to care, the hierarchical condition category (HCC) risk score (prospectively measured) is the most comprehensive, payment-weighted measure of health status before the index hospitalization. Tables 5-2 and 5-3 show only small differences between the comparison and demonstration hospitals with regard to age, gender, race, and HCC risk score in both the baseline and performance years.

The beneficiary's DRG weight is included as a regressor because it directly influences the IPPS payment amount. It also is likely to influence health care services used in the post discharge period (e.g., rehabilitation after orthopedic surgery). In the baseline year, CAMC comparison hospitals had an average DRG weight of 4.287, which increased slightly to 4.330 in the performance year. CAMC, by comparison, had an average DRG weight of 4.294 in the baseline year, which decreased slightly to 4.224 in the performance year.

For BIMC, comparison hospitals had an average DRG weight of 1.57, which increased slightly to 1.605; BIMC had a DRG weight of 1.341, which increased slightly to 1.358.

Three hospital-specific measures, for the index hospitalization, are included among the explanatory variables: (1) the intern-/resident-to-bed (IRB) ratio, (2) the IPPS disproportionate share (DSH) adjustment factor, and (3) the number of hospital beds.[6] The first two directly affect IPPS payment amounts and are included to factor out variations in hospital payments directly due to the payment mechanism that are outside the control of the sites. They are also included because they might explain variation in payments for index admission outlier and physician services as well as services rendered during the pre-admission and post-discharge periods. Hospital beds are included because they might be a proxy for the effect of hospital size and complexity on the intensity of care during and after the index hospitalization.

CAMC had a higher IRB ratio during the base year (at 1.70) that declined in the performance year (to 1.64). The CAMC comparison sites had a lower IRB ratio (1.35), which also declined slightly (to 1.32). BIMC exhibited a lower IRB ratio in the base year (at 0.347) which increased over the performance period (to 0.425). By comparison, the IRB ratio for BIMC's comparison hospitals decreased slightly over the performance period (from 0.543 to 0.532). For CAMC, we observed that the DSH adjustment factor remained relatively constant for the comparison hospitals, but increased from 0.125 to 0.146 for the CAMC demonstration site. The DSH adjustment factors increased slightly during the performance period for both BIMC and its comparison hospitals.

For the LOS regression, the national mean LOS for the index DRG is also included as an explanatory variable because the LOS might not be able to fall much, especially for DRGs with inherently low LOSs. For CAMC, average DRG LOS decreased for the demonstration hospital but increased for the comparison hospitals. For BIMC, average DRG LOS decreased in both the demonstration and comparison hospitals.

The regressions in both this section and in Section 6 (Quality of Care) were estimated after adjusting for clustering. Clustering, when present, means that individual observations from a given hospital are not independent from each other. Specifically, each observation within a hospital provides less information about treatment in the hospital than if the observations were independent. In essence, each observation is worth less than one "full" independent observation.

Clustering can occur if Medicare payments for beneficiaries in one hospital are more alike than Medicare payments for beneficiaries in another hospital. This can occur for several reasons. First, for discharges from a given hospital, the standardized amounts, the IPPS area wage index, the IME adjustment, and the DSH adjustment are all the same. Second, extended LOSs during the index hospitalization probably systematically differs by hospital because it is likely that each hospital has its own discharge protocols. Extended LOSs can lead to higher costs and, possibly, outlier payments. Third, discharge destinations and treatment in the post-discharge period also probably differ by hospital. For instance, some hospitals might systematically discharge more beneficiaries to home health, whereas other hospitals might

[6] A fourth hospital-specific measure, the IPPS area wage index, is used in the evaluation of the Gainsharing demonstration because the demonstration and comparison hospitals are located within several IPPS area wage index payment areas. It is used for the Gainsharing demonstration only for CAMC since BIMC and all of its comparison hospitals have the same IPPS area wage index value for any given performance year.

discharge more beneficiaries to home. For types of patients commonly discharged to SNFs, rehabilitation hospitals, and long-term care hospitals, the index hospital might have a preferred set of providers to which the patients are sent. Differing styles of care by hospitals during index hospitalizations might lead to systematically different readmission rates and use of other medical services.

Adjusting for clustering does not affect the estimated regression coefficients. It does, however, usually increase the estimated standard errors. Increased standard errors result in higher p-values and can result in loss of statistical significance for individual explanatory variables.

5.3 Results—CAMC

CAMC Descriptive Findings: For CAMC, average total episode payments were $32,813 in the base period and increased to $36,614 in the performance year. CAMC comparison hospital average total episode payments increased $830 more on average, growing from $33,833 to $38,464 (***Table 5-4***). CAMC demonstration Medicare hospital payments for the index hospitalization averaged $24,930 in the base year, about 76% of average episode payments, and $28,138 in the performance year (with the proportion of episode payments increasing slightly). Similar trends were noted for the CAMC comparison sites, although inpatient hospital payments increased very slightly more for this group over the performance period. Inpatient physician payments added roughly another 12 percent to episode payments for CAMC and just under 10 percent for the CAMC comparison hospitals. The inpatient physician payment proportions of total episodes decreased for both groups during the performance year. Medicare payments for health services during the pre-admission period accounted for around 1.5 percent of episode payments for comparison hospitals, decreasing to 1.42 percent during the performance period. CAMC exhibited a higher proportion (1.79 percent) during the base pre-admission period, and decreased to 1.52 percent during the performance period. Post discharge period payments accounted for over 12 percent of episode payments for CAMC comparison hospitals and over 10 percent for CAMC; slight increases in the actual dollars and proportion of episode payments were noted for both CAMC and the comparison hospitals from the base period to the performance year. Follow-on hospital admissions (e.g., readmissions, long-term care, and rehabilitation admissions) accounted for the largest share of post discharge spending, at over 6 percent for the comparison hospitals and roughly 5.25 percent for the CAMC demonstration hospitals. We observed only increases in the spending for these post discharge payments from the base to the performance year for both CAMC and the comparison hospitals.

Table 5-4
Components of Medicare payments: CAMC

Period (Payment component)	Base year				Performance period				Comparison	Demo	
	Comparison		Demo		Comparison		Demo				
	Mean payments	Percent of episode payments	Mean payments	Percent of episode payments	Mean payments	Percent of episode payments	Mean Payments	Percent of episode payments	Change between BY and PY	Change between BY and PY	Difference in differences
14-day pre-admission period											
Physician	297	0.88	362	1.10	298	0.77	347	0.95	1	−15	−16
Outpatient	196	0.58	216	0.66	238	0.62	202	0.55	42	−15	−57
Durable medical equipment	11	0.03	7	0.02	12	0.03	8	0.02	1	1	0
Total	504	1.49	586	1.79	548	1.42	556	1.52	44	−30	−74
Index hospitalization period											
IPPS hospital inlier	25,225	74.56	24,332	74.15	28,231	73.39	27,043	73.86	3,005	2,711	−294
IPPS hospital outlier	682	2.02	598	1.82	1,033	2.68	1,095	2.99	351	497	146
IPPS hospital total	25,907	76.58	24,930	75.98	29,263	76.08	28,138	76.85	3,356	3,208	−148
Physician	3,333	9.85	3,972	12.10	3,759	9.77	4,073	11.12	426	101	−324
Total	29,240	86.43	28,902	88.08	33,022	85.85	32,211	87.97	3,782	3,309	−473
Post discharge period											
Inpatient	2,200	6.50	1,715	5.23	2,383	6.19	1,934	5.28	182	219	37
Skilled nursing facility	599	1.77	404	1.23	897	2.33	450	1.23	298	46	−252
Durable medical equipment	53	0.16	42	0.13	56	0.14	110	0.30	2	68	66
Outpatient	144	0.43	134	0.41	192	0.50	138	0.38	48	4	−44
Physician	578	1.71	716	2.18	658	1.71	750	2.05	80	34	−47

(continued)

Table 5-4 (continued)
Components of Medicare payments: CAMC

Period (Payment component)	Base year				Performance period				Comparison Change between BY and PY	Demo Change between BY and PY	Difference in differences
	Comparison		Demo		Comparison		Demo				
	Mean payments	Percent of episode payments	Mean payments	Percent of episode payments	Mean payments	Percent of episode payments	Mean payments	Percent of episode payments			
Home health agency	513	1.52	313	0.95	709	1.84	465	1.27	196	151	−45
Total	4,088	12.08	3,325	10.13	4,895	12.72	3,847	10.51	806	522	−284
Total episode payments	33,833	—	32,813	—	38,464	—	36,614	—	4,632	3,801	−830

NOTE: BY = base year; CAMC = Charleston Area Medical Center; IPPS = inpatient prospective payment system; PY = performance year.

SOURCE: RTI processing of Medicare claims.

Mean episode payments for each of the four expenditure measures are shown in ***Table 5-5*** for CAMC and its comparison hospitals. Average total episode payments for CAMC and the comparison hospitals were similar in the baseline year. However, by the end of the performance year, CAMC hospital payments had grown at a slower rate on average than those of the comparison hospitals. Average total episode payments for the comparison group hospitals increased by $4,632 from the baseline through Year 1 and for CAMC by $3,801, a difference of $830 less for CAMC relative to their comparison sites. Payment growth rates varied somewhat under the alternative episode definitions. Excluding the IPPS inlier and outlier payments for index hospitalizations, the comparison hospital payments increased by $1,276 between baseline and Year 1 but CAMC's payments increased by only $594, resulting in a difference in growth rates of $682. Using the third episode definition that excludes only the inlier payment, average total payments grew $536 less at CAMC. Finally, considering average payments for only the pre- and post discharge windows, average payments increased for the comparison hospitals by $850 and for CAMC by $492, a $358 difference.

Both CAMC and the comparison hospitals exhibited very high average LOSs during the baseline and performance years, substantially exceeding the national Medicare average LOS (5.6 days in 2008 and 5.5 days in 2009).[7] Part of this finding can be explained by the high average HCC scores exhibited by these groups (over 1.5; see Table 5-2) and by the fact that the analysis was limited to a subset of cardiac cases. However, whereas the CAMC average LOS decreased slightly from 9.77 to 9.75 days, the comparison group LOS actually increased from the base to the performance year—to 8.28 from 7.81.

We do observe a large change in outlier payments for beneficiaries incurring an outlier among both CAMC demonstration and the comparison hospitals. The share of discharges with an outlier payment nearly doubled for both CAMC and its comparison hospitals. At $14,247 per outlier discharge, CAMC had slightly higher outlier payments per outlier patient than the $14,079 for the comparison hospitals during the base year. Outlier payments per outlier beneficiary fell an average of $2,378 per discharge for comparison hospitals between the baseline and first performance year, and increased $151 for CAMC.

CAMC Episode Payments*: *Table 5-6 presents estimates for the full model (Equation 5-1) for four payment measures for CAMC. We note that these results should be interpreted with caution because the CAMC site operated for a relatively short time period (1 year) and limited its intervention to cardiac-related care. The 2D coefficients ranged from $385 lower to $307 higher for episodes originating at CAMC than at comparison hospitals, but none of these coefficients were statistically significant, even at the 10 percent level. This result indicates that the Gainsharing demonstration did not have a significant effect on per-episode Medicare payments for CAMC.

---

7 https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/DataCompendium/2011_Data_Compendium.html, accessed on June 7, 2012.

Table 5-5
Mean episode payments, length of stay, and IPPS outliers by group and time period: CAMC

Name	Base year		Performance year 1		Change between BY and PY comparison	Change between BY and PY demo	Difference in differences
	Comparison	Demo	Comparison	Demo			
Total episode Medicare payments*	$33,833	$32,813	$38,464	$36,614	$4,632	$3,801	−$830
Total episode payments other than the payment to the index hospital*	$7,925	$7,883	$9,201	$8,476	$1,276	$594	−$682
Total episode payments other than the flat inlier DRG payment to the index hospital. This includes any outlier payments made to the index hospital.*	$8,607	$8,481	$10,234	$9,571	$1,626	$1,090	−$536
Total episode payments for the 14-day pre-admission period plus the 30-post discharge period (same as second type except physician payments during the index hospitalization are excluded)*	$4,592	$3,911	$5,442	$4,403	$850	$492	−$358
Length of stay (days)	7.81	9.77	8.28	9.75	0.47	−0.02	−0.48
IPPS outlier (index hospitalization [discharge])							
Outlier payments overall index discharges	$682	$598	$1,033	$1,095	$351	$497	$146
Percentage of index discharges with outlier payments	4.84%	4.20%	8.82%	7.60%	3.98%	3.41%	−0.57%
Outlier payments per index discharges with an outlier	$14,079	$14,247	$11,701	$14,398	−$2,378	$151	$2,528
Maximum number of observations	6,111	882	5,621	855	n/a	n/a	n/a

NOTE: *Excludes beneficiary co-payments. BY = base year; CAMC = Charleston Area Medical Center; DRG = diagnosis-related group; IPPS = inpatient prospective payment system; PY = performance year.

SOURCE: RTI processing of Medicare claims.

Table 5-6
CAMC episode payment regressions

Explanatory variable	1. Total episode payments			2. Total except the IPPS index			3. #2 plus IPPS outlier payments			4. Pre-adm & post discharge		
	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	−82.27	383.82	0.835	−61.40	261.55	0.819	38.77	250.58	0.880	−35.31	219.46	0.875
Patient age 70–74	675.91	419.43	0.138	384.33	229.49	0.125	555.36	246.85	0.048	372.18	202.73	0.096
Patient age 75–79	1,347.79	411.71	0.008	1,062.69	251.49	0.002	1,178.32	294.57	0.003	1,126.71	230.01	0.001
Patient age 80 or more	2,265.14	535.96	0.002	1,696.88	272.98	0.000	1,870.35	365.14	0.000	1,854.04	235.89	0.000
Female	1,486.36	225.22	0.000	1,316.60	186.71	0.000	1,462.20	180.52	0.000	1,207.07	175.85	0.000
Nonwhite	−118.89	933.34	0.901	261.07	684.66	0.711	−92.44	820.76	0.913	555.16	576.46	0.358
DRG weight	8,128.26	200.12	0.000	1,433.50	93.45	0.000	1,864.21	102.89	0.000	731.43	91.24	0.000
HCC risk score	410.09	132.19	0.011	283.67	82.45	0.006	303.53	94.08	0.009	296.19	72.61	0.002
Participating hospital indicator	291.16	438.27	0.521	1,145.82	242.09	0.001	788.73	226.86	0.006	659.21	176.53	0.004
Performance period indicator	3,832.01	540.82	0.000	1,249.53	268.62	0.001	1,511.13	348.87	0.001	896.55	206.56	0.001
2D estimator	307.30	555.31	0.592	−384.92	284.92	0.206	−206.01	355.10	0.575	−213.23	229.34	0.374
IPPS area wage index	22,062.79	1,489.17	0.000	6,884.37	1,635.56	0.002	6,321.42	1,584.39	0.003	5,880.75	1,450.41	0.002
Intern-/resident-to-bed ratio	8,397.94	2,548.69	0.008	−2,974.00	2,405.66	0.245	−1,711.83	1,898.76	0.388	−2,603.25	1,827.96	0.185
Hospital beds	1.64	2.86	0.580	−3.73	0.98	0.004	−2.44	1.23	0.074	−4.72	1.09	0.001
DSH adjustment factor (operating)	25,691.65	3,283.62	0.000	−6,469.77	2,030.54	0.010	−4,043.50	2,111.46	0.084	−3,554.95	1,772.76	0.073
Constant term	−28,704.25	2,597.62	0.000	−2,809.32	1,656.26	0.121	−4,835.27	1,933.21	0.031	−2,176.48	1,640.25	0.214
R^2	0.4453	—	—	0.0828	—	—	0.0857	—	—	0.0454	—	—
Number of observations	13,187	—	—	13,187	—	—	13,187	—	—	13,187	—	—

NOTE: CAMC = Charleston Area Medical Center; 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

The participating hospital indicator was positive for all four measures of episode payments but statistically significant only for measures 2–4. These results suggest that CAMC patients had higher PAC costs than patients at the comparison hospitals. The performance period indicator shows that episode payments for comparison hospitals increased from $897 (model 4) to $3,832 (model 1). All four coefficients are statistically significant at conventional levels. Even after controlling for gender, race, DRG, and HCC score, beneficiary age still shows a strong positive effect on all four payment measures. The coefficients for gender are positive and statistically significant, indicating that women with cardiac conditions were more expensive for Medicare than male cardiac patients. The coefficients for race are not statistically significant.

The two variables most directly associated with payments and the health condition of beneficiaries—DRG weight and the patient's prospective HCC risk score—both have positive, statistically significant effects. The coefficient for the DRG weight is over $8,000 in the total episode payment regression (model 1), but it falls to under $2,000 in the other three episode payment regressions. The DRG weight coefficient is artificially higher in the total payments regression that includes the DRG inlier payment because these payments are determined in a compound, nonlinear fashion, whereas the wage index, IRB ratio, and DSH ratio were estimated in linear form. In the other three models, DRG weight clearly plays a cost-increasing role for outlier and physician payments as well as adding significantly ($731) to pre- and post-discharge payments, even after controlling for other beneficiary characteristics. As expected, the IPPS area wage index and IRB ratio are major contributors to total episode payments because of their role in determining inlier hospital payments. The wage index may be acting as a proxy for prices and utilization in the ambulatory setting. Greater teaching intensity also adds considerably to the DRG inlier payment, but not to extra physician and outlier payments. The coefficient on the HCC risk score ranges from $410 (model 1) down to $283 (model 2).

CAMC Length of Stay and IPPS Outlier Payments: The 2D coefficients for LOS were positive but not statistically significant for CAMC (***Table 5-7***). The coefficient for the HCC risk score is positive and statistically significant for CAMC. The patient age variables have the same signs and patterns of coefficients as in the episode payment regressions. Two specifications of hospital outlier payments were analyzed. The estimated 2D effect for total outlier payments per beneficiary episode was $214 for CAMC, but was not statistically significant. For CAMC, outlier payments were nearly $5,500 higher for patients with outlier payments.

CAMC Inpatient Physician Payments: For CAMC, the 2D coefficient in the inpatient physician payments regression (***Table 5-8***) was −$171.69 and was statistically significant ($p = 0.033$) with an adjusted R^2 of 0.30. This indicates that Medicare Part B physician payments increased by $171.69 less at CAMC than at the comparison hospitals.

5.4 Results—BIMC

BIMC Descriptive Findings: BIMC had average total episode payments of $22,127 in the base period, which increased to $23,634 at the end of the third performance year. BIMC's comparison hospitals average total episode payments also increased during the performance period, from $23,413 to $26,122, an average of $202 more than BIMC (***Table 5-9***). BIMC's Medicare hospital payments for the index hospitalization averaged $12,523 in the base year, over half of average episode payments, and increased to $13,409 in the performance period (with the

Table 5-7
CAMC length of stay and IPPS outlier payment regressions

Explanatory variable	Length of stay			IPPS outlier payment amount			Deconstruction of IPPS outlier payments: Logit results for the likelihood of an IPPS outlier hospitalization			Deconstruction of IPPS outlier payments: OLS results on IPPS outlier payments for discharges with an outlier payment		
	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	0.131	0.168	0.454	122.23	145.36	0.420	0.20	0.16	0.199	611.59	2,151.94	0.782
Patient age 70–74	0.301	0.125	0.037	185.84	165.97	0.289	0.16	0.09	0.067	413.17	1,417.06	0.777
Patient age 75–79	0.447	0.116	0.003	146.14	95.16	0.156	0.10	0.10	0.316	606.57	1,528.99	0.700
Patient age 80 or more	0.734	0.162	0.001	216.50	137.48	0.146	0.11	0.05	0.028	165.73	1,761.63	0.927
Female	0.451	0.085	0.000	98.19	66.76	0.172	−0.07	0.07	0.304	1,529.18	1,169.09	0.220
Nonwhite	0.441	0.206	0.058	−303.99	139.48	0.054	−0.29	0.19	0.132	−1,724.29	1,714.99	0.338
DRG weight	0.278	0.080	0.006	233.35	25.62	0.000	0.30	0.06	0.000	832.10	441.78	0.089
HCC risk score	0.112	0.042	0.025	30.25	22.31	0.205	0.03	0.01	0.024	39.04	360.61	0.916
Participating hospital indicator	1.732	0.471	0.004	−407.77	129.23	0.010	−0.28	0.22	0.193	−4,114.41	1,458.65	0.018
Performance period indicator	0.226	0.223	0.334	245.74	123.84	0.075	0.60	0.13	0.000	−3,293.61	992.27	0.008
2D estimator	0.095	0.248	0.710	213.96	137.79	0.152	−0.18	0.14	0.202	5,487.78	1,175.96	0.001
IPPS area wage index	4.039	2.055	0.078	−682.28	484.07	0.189	−2.84	1.16	0.014	9,860.41	11,380.09	0.407
Intern-/resident-to-bed ratio	−2.598	2.841	0.382	1,254.73	848.76	0.170	1.78	1.15	0.123	13,406.84	9,221.94	0.177
Hospital beds	0.002	0.002	0.462	1.23	0.84	0.176	0.00	0.00	0.945	13.26	6.93	0.085
DSH adjustment factor (operating)	1.258	2.287	0.594	2,918.55	972.65	0.013	4.28	0.98	0.000	−14,600.43	10,693.21	0.202
Average LOS for DRG	0.811	0.036	0.000	134.71	24.84	0.000	0.04	0.05	0.444	910.23	305.99	0.014
Constant term	−4.545	2.581	0.109	−2,107.47	688.29	0.012	−3.08	1.12	0.006	−15,572.76	12,288.30	0.234
R^2 (pseudo for logit)	0.2796	—	—	0.0262	—	—	0.0804	—	—	0.0925	—	—
Number of observations	13,187	—	—	13,187	—	—	13,187	—	—	874	—	—

NOTE: CAMC = Charleston Area Medical Center; 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system; LOS = length of stay; OLS = ordinary least-squares.

SOURCE: RTI processing of Medicare claims.

Table 5-8
CAMC physician inpatient payments regression

Explanatory variable	Coefficient	Standard error	P > | t |
Patient age 0–64	−26.09	64.19	0.693
Patient age 70–74	12.15	51.03	0.817
Patient age 75–79	−64.02	48.12	0.213
Patient age 80 or more	−157.16	56.69	0.020
Female	109.53	28.07	0.003
Nonwhite	−294.09	119.45	0.034
DRG weight	702.07	28.55	0.000
HCC risk score	−12.52	15.60	0.441
Participating hospital indicator	486.61	186.47	0.026
Performance period indicator	352.98	84.52	0.002
2D estimator	−171.69	69.22	0.033
IPPS area wage index	1,003.62	855.72	0.268
Intern/resident to bed ratio	−370.75	907.10	0.691
Hospital beds	0.99	0.81	0.251
DSH adjustment factor (operating)	−2,914.82	995.89	0.015
Constant term	−632.84	625.00	0.335
R^2	0.3009	—	—
Number of observations	13,187	—	—

NOTE: Physician inpatient payments are for the index hospitalization only.
CAMC = Charleston Area Medical Center; 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

Table 5-9
Components of Medicare payments: BIMC

Period (Payment component)	Mean payments				Percent of total episode payments			
	Base year comparison	Base year BIMC	Performance year 3 comparison	Performance year 3 BIMC	Base year comparison	Base year BIMC	Performance year 3 comparison	Performance year 3 BIMC
14-day pre-admission period								
Physician	276	242	293	242	1.18	1.15	1.12	1.02
Outpatient	81	69	101	94	0.34	0.32	0.39	0.40
Durable medical equipment	16	18	15	19	0.07	0.09	0.06	0.08
Total	373	329	409	355	1.59	1.56	1.57	1.50
Index hospitalization period								
IPPS hospital inlier	14,067	12,432	14,844	13,095	60.08	58.84	56.83	55.41
IPPS hospital outlier	326	91	560	314	1.39	0.43	2.14	1.33
IPPS hospital total	14,394	12,523	15,404	13,409	61.48	59.28	58.97	56.74
Physician	1,531	1,469	1,713	1,727	6.54	6.95	6.56	7.31
Total	15,925	13,992	17,117	15,136	68.02	66.23	65.52	64.04
Post-discharge period								
Inpatient	3,391	3,196	3,465	3,163	14.48	15.13	13.26	13.38
Skilled nursing facility	1,712	1,610	2,839	2,767	7.31	7.62	10.87	11.71
Durable medical equipment	73	81	75	75	0.31	0.38	0.29	0.32
Outpatient	142	156	197	178	0.61	0.74	0.76	0.75
Physician	971	942	1,093	1,054	4.15	4.46	4.18	4.46
Home health agency	825	821	928	905	3.52	3.89	3.55	3.83
Total	7,115	6,806	8,597	8,143	30.39	32.21	32.91	34.45
Total episode	23,413	21,127	26,122	23,634	—	—	—	—
Number of observations	66,785	8,913	67,336	9,705	—	—	—	—

NOTE: BIMC = Beth Israel Medical Center; IPPS = inpatient prospective payment system

SOURCE: RTI processing of Medicare claims.

share of episode payments decreasing by about 2.5 percentage points). BIMC's comparison hospitals' payment levels were about $2,000 higher than BIMC's during both the base year and the third performance year, reflecting, in part, their higher average casemix (DRG weights). The index hospitals payments share of total episode payments for the comparison hospitals also fell about 2.5 percentage points. Inpatient physician payments were nearly seven percent of BIMC's total episode payment and about 6.5 percent for the comparison hospitals. The proportion of inpatient physician payments relative to total episode payments increased for both groups over the performance period—but more so for BIMC. Medicare payments for health services during the pre-admission period accounted for just under 1.6 percent for both BIMC and the comparison hospitals, and these proportions fell slightly for both groups during the performance period. Post-discharge period payments accounted for over 30-32 percent of total episode payments for BIMC and the comparison hospitals during the base year. BIMC and the comparison hospital proportions of post-discharge payments relative to total episode payments increased over the performance period—from 32.2 percent to 34.5 percent at BIMC and from 31.4 to 32.9 percent at the comparison hospitals. Follow-up hospital and SNF admissions (e.g., readmissions, long-term care, and rehabilitation admissions) accounted for the largest share of post-discharge spending at over 22 percent of the episode for the comparison hospitals and roughly 23 percent for BIMC in the base period; these figures increased about 2.3 percentage points for both groups during the performance years.

Mean episode payments for each of the four expenditure measures are shown in ***Table 5-10*** for BIMC and its comparison hospitals. Average total episode Medicare payments for BIMC were about $2,300 lower than for its comparison hospitals during the baseline year. As can be seen in the upper left panel of ***Figure 5-1***, BIMC's average total episode payments were lower than those for its comparison hospitals not only in the base year but also in each of the performance years. Average total Medicare payments for BIMC's comparison hospitals increased $2,709 between the base year and PY 3 while for BIMC the increase was $2,507. Thus, the difference between BIMC and its comparison hospitals increased $202 to nearly $2,500 during the third performance period. Another way viewing changes in average episode payments over time is the upper left panel of ***Figure 5-2***. The line plotted in this panel shows the unadjusted 2D (difference in differences) between the base year and each performance year and is calculated as follows:

= (average BIMC PY payments minus average BIMC base year payments) –
(average comparison PY payments minus average comparison base year payments).

A positive value for the 2D estimate indicates the amount BIMC payments increased more than at its comparison hospitals whereas a negative 2D value indicates the amount BIMC payments increased less than at its comparison hospitals. The upper left panel of Figure 5-2 shows that largest unadjusted 2D estimate was about -$500 for PY 1 and that it became smaller in each succeeding year so that it was down to about -$200 in PY 3. This line approaches zero difference (heavy red gridline) over time. This suggests that the largest impact of the Gainsharing demonstration on total episode Medicare payments was during its first year and that it became smaller thereafter.

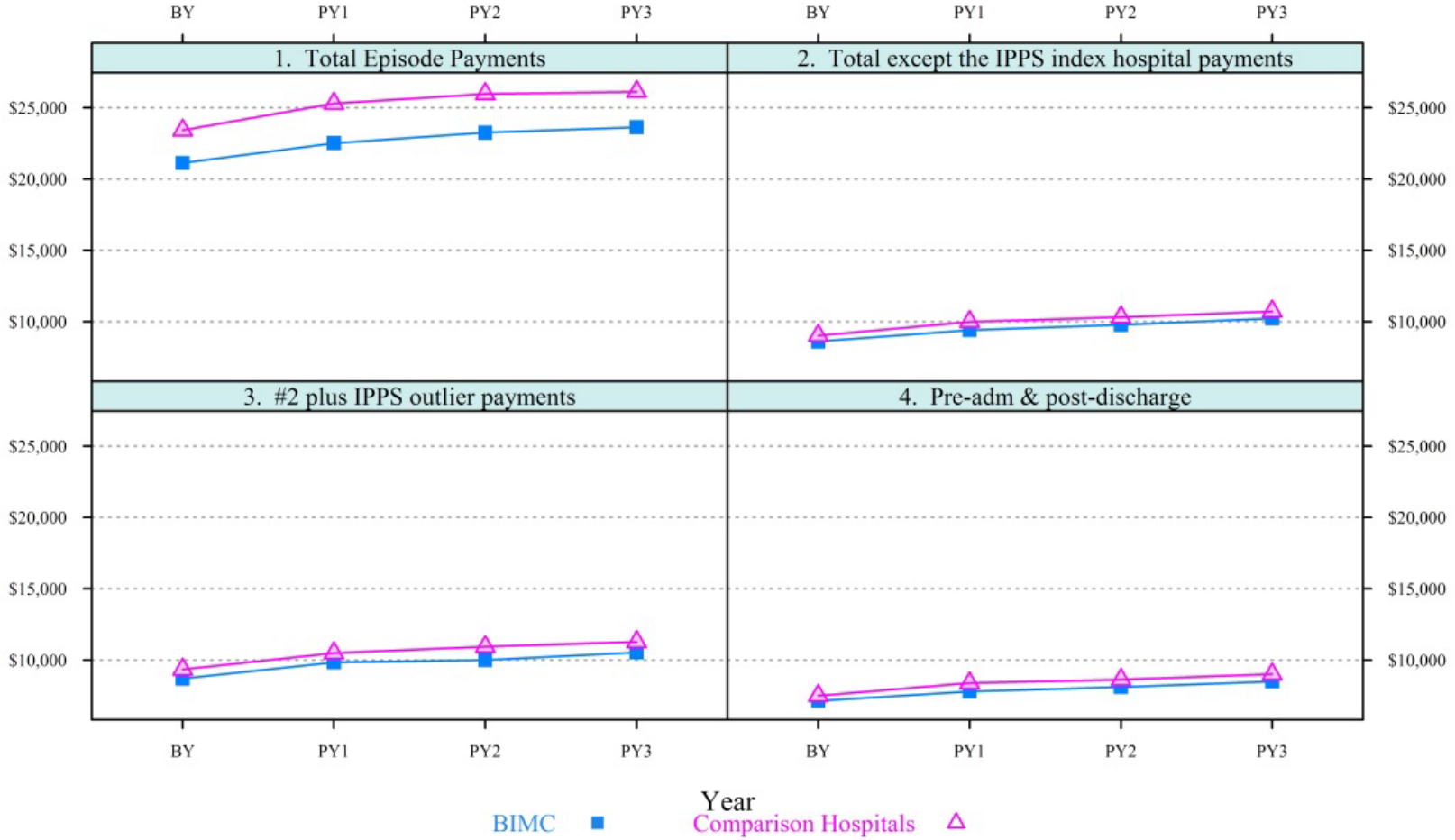
Table 5-10
Mean episode payments, length of stay, and IPPS outliers by group and time period: BIMC

Name	Base year comparison	Base year BIMC	Performance year 3 comparison	Performance year 3 BIMC	Change between BY and performance year 3 comparison	Change between BY and performance year 3 BIMC	Difference in differences
Total episode Medicare payments*	$23,413	$21,127	$26,122	$23,634	$2,709	$2,507	-$202
Total episode payments other than the payment to the index hospital*	$9,020	$8,604	$10,718	$10,225	$1,699	$1,621	-$78
Total episode payments other than the flat inlier DRG payment to the index hospital. This includes any outlier payments made to the index hospital.*	$9,346	$8,695	$11,278	$10,539	$1,932	$1,845	-$87
Total episode payments for the 14-day pre-admission period plus the 30-day post-discharge period (same as second type except physician payments during the index hospitalization are excluded)*	$7,488	$7,135	$9,006	$8,498	$1,518	$1,363	-$154
Length of stay (days)	7.13	6.69	6.70	6.75	-0.43	0.05	0.48
IPPS outlier (index hospitalization [discharge])							
Outlier payments overall index discharges	$326	$91	$560	$314	$233	$223	-$10
Percent of Index discharges with outlier payments	1.66%	0.73%	2.56%	2.06%	0.90%	1.33%	0.43%
Outlier payments per index discharges with an outlier	$19,625	$12,467	$21,817	$15,255	$2,193	$2,788	$595
Physician payments during index discharge	$1,531	$1,469	$1,713	$1,727	$181	$258	$77
Number of observations	66,785	8,913	67,336	9,705	n/a	n/a	n/a

NOTE: *Excludes beneficiary co-payments. BIMC = Beth Israel Medical Center; BY = base year; DRG = diagnosis-related group; IPPS = inpatient prospective payment system; and n/a = not applicable.

SOURCE: RTI processing of Medicare claims.

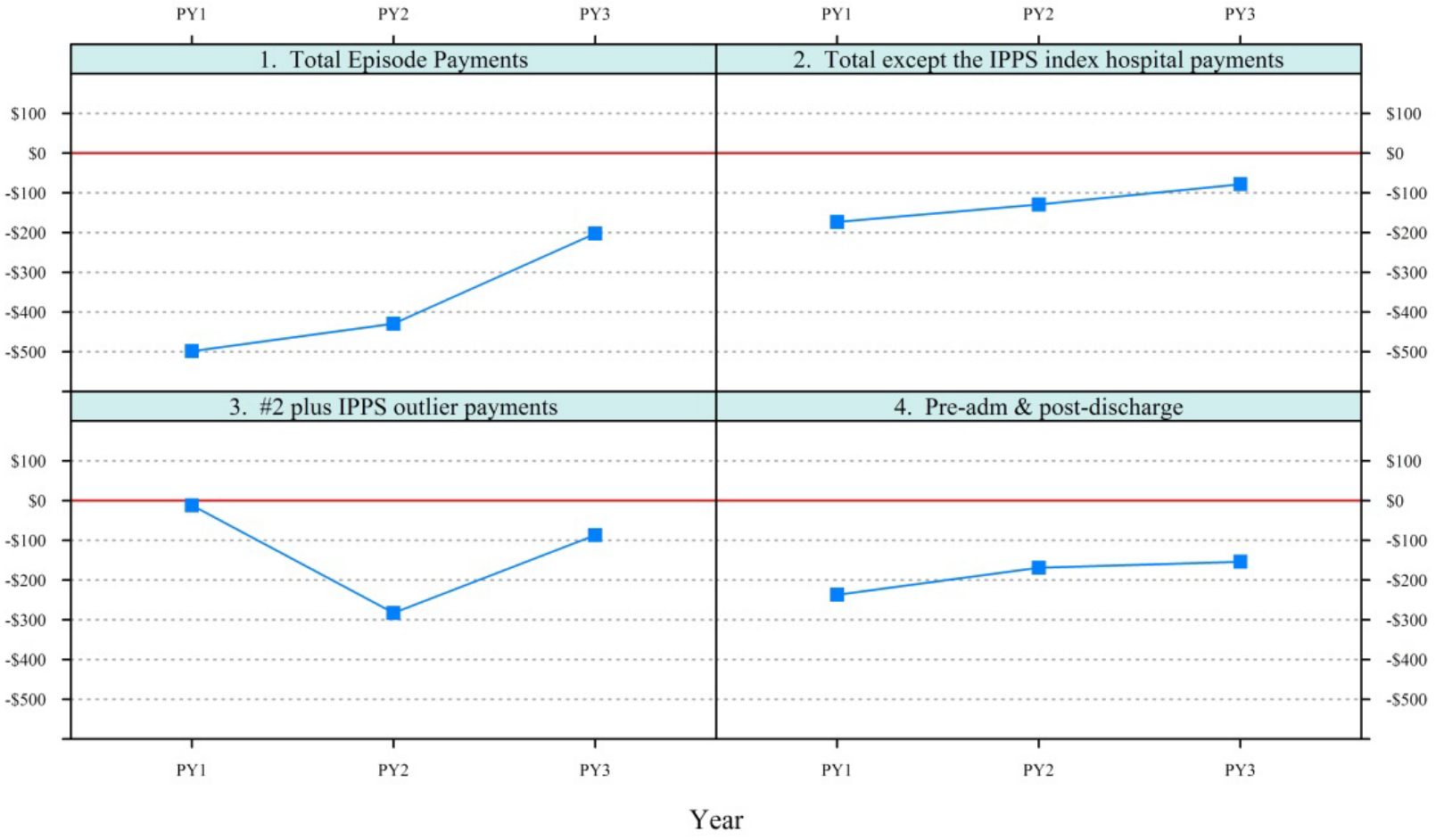
Figure 5-1
Average total episode payments for four types of payment measures by performance year and demonstration status



NOTES:
Panel 1: total episode payments—includes payments to all providers;
Panel 2: episode payments excluding Medicare's inlier and outlier payments to the index hospital;
Panel 3: episode payments excluding only the fixed inlier DRG payment to the index hospital; and
Panel 4: episode payments for only the 14-day pre-admission and 30-day post discharge periods.
The panel numbers conform to the four episode payment measures shown in Table 5-1.

SOURCE: RTI processing of Medicare claims.

Figure 5-2
Unadjusted difference in differences for four types of payment measures by performance year



NOTE: See notes to Figure 5-1. Difference in differences for each performance year was based on the differences between the performance and the base year.

SOURCE: RTI processing of Medicare claims.

Payment growth rates varied somewhat under the alternative episode definitions. Excluding the IPPS inlier and outlier payments for index hospitalizations, the comparison hospital payments increased by $1,699 from baseline through performance period 3 and by $1,621 for BIMC, resulting in a difference in growth rates of $78 in BIMC's favor. As with total episode payments, the largest change in BIMC's favor for the second payment measure was in PY 1 (upper right panel in Figure 5-2). Using the third episode definition that excludes only the inlier payment, average total payments grew $87 less at BIMC. Unlike the two previous episode payment measures, there was little difference in the change in payments between the base year and PY 1 (lower left panel of Figure 5-2). There was a large change between the base year and PY 2 in BIMC's favor but was transitory as difference narrowed in PY 3. Finally, considering average payments for only the pre- and post-discharge windows, average payments increased for the comparison hospitals by $1,518 and for BIMC by $1,363, a $154 difference. The unadjusted 2D estimates for this last measure decreased slightly over the course of the demonstration (lower right panel in Figure 5-2).

Both BIMC and its comparison hospitals exhibited relatively high average LOSs during the baseline and performance years, substantially exceeding the national Medicare average LOS (5.6 days in 2008 and 5.5 days in 2009).[8] Part of this finding can be explained by the high average HCC scores exhibited by these groups (over 2.2; see Table 5-3). Whereas average LOS fell for BIMC's comparison hospitals from 7.13 to 6.7 days, the BIMC's average LOS increased slightly from 6.69 to 6.75 days. As can be seen in the left panel of ***Figure 5-3***, while average LOS continually fell for its comparison hospitals, BIMC's went up, then down, and then back up again. BIMC's average LOS did not exhibit the reductions self-reported by BIMC during our site visit discussions.

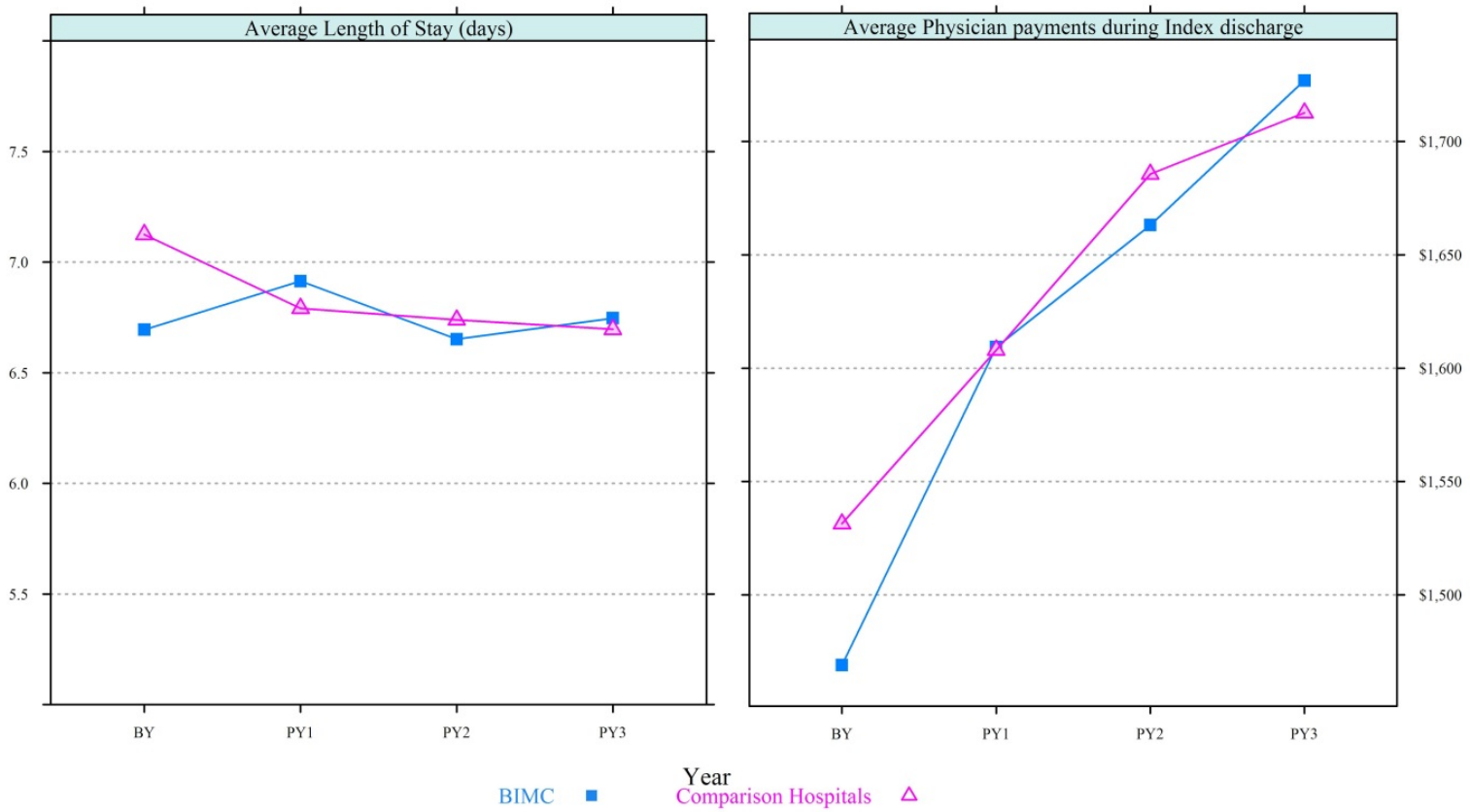
During the base year, BIMC had a much lower share of the discharges with an outlier payment (0.73 percent) than its comparison hospitals (1.66 percent). Both had large changes in the share of discharges with an outlier payment, but BIMC's increase was 0.43 percentage point higher. As can be seen in the left panel of ***Figure 5-4***, there was considerable variation in year to year changes. Average outlier payments per index discharge were $326 for BIMC's comparison hospitals during the base year while they were $91 for BIMC. Average outlier payment increased for both (right panel of Figure 5-4). They increased $233 between the base year and PY 3 for BIMC's comparison hospitals while BIMC's increase was $223. During the base year, average outlier payments per outlier discharge were $19,625 for the comparison hospitals, much higher than BIMC's $12,467. By PY 3, average outlier payments per outlier discharge increased $2,193 at the comparison hospitals and increased $2,788 at BIMC. Even so, average outlier payments per outlier discharge at BIMC during PY 3 were about $6,500 lower than at the comparison hospitals.

Average inpatient physician payments during the index discharge were $1,531 at BIMC's comparison hospitals, about $62 higher than at BIMC. By PY 3, however, average physician payments at BIMC were $14 higher than at its comparison hospitals, an unadjusted 2D effect of

---

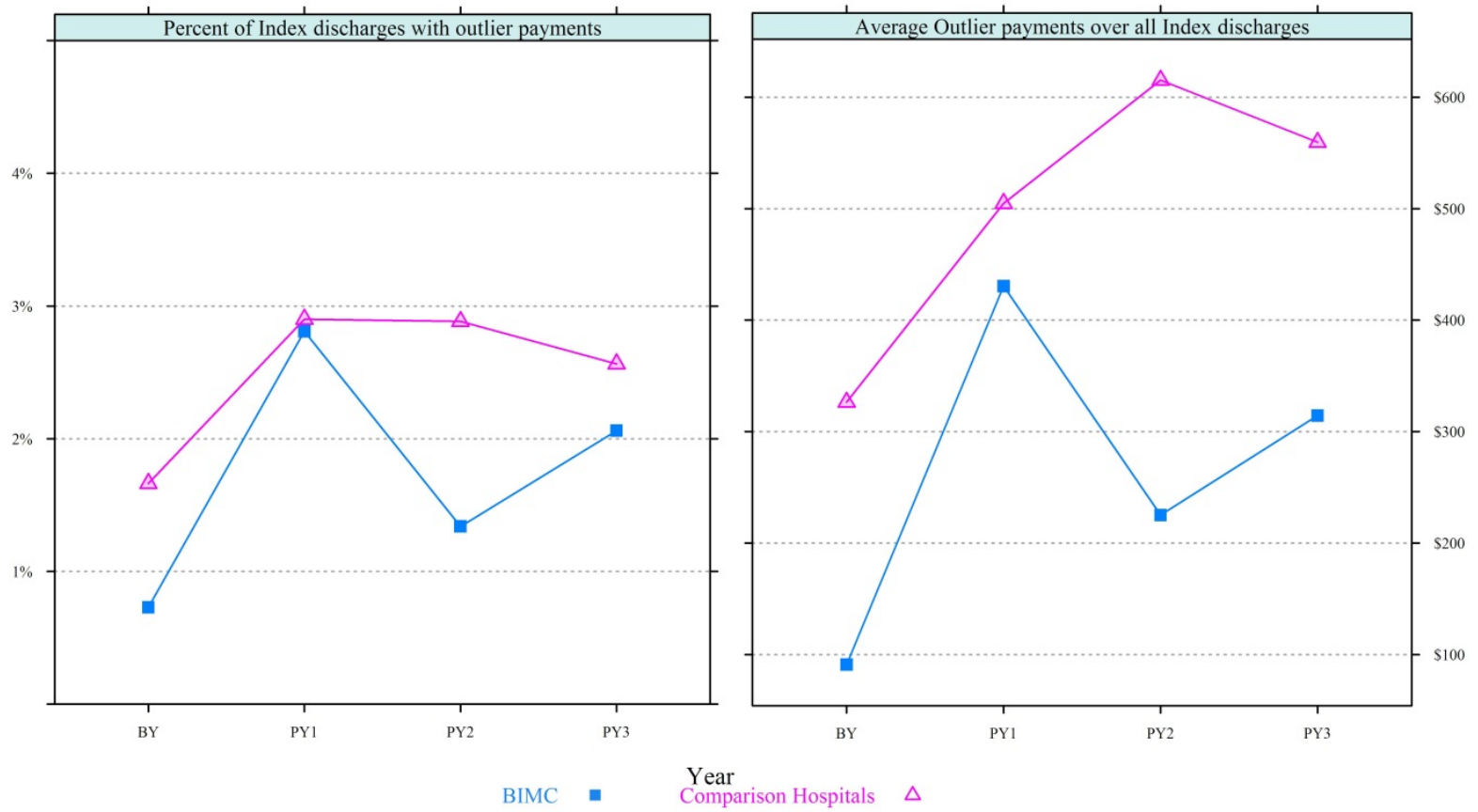
8 https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/DataCompendium/2011_Data_Compendium.html, accessed on June 7, 2012.

Figure 5-3
Average length of stay and average inpatient physician payments during index hospital stay by performance year and demonstration status



SOURCE: RTI processing of Medicare claims.

Figure 5-4
Share of index discharges with outlier payments and average outlier payments over all index discharges by performance year and demonstration status



SOURCE: RTI processing of Medicare claims.

$77. Most of BIMC's *relative* increase occurred between the base year and PY 1 (right panel in Figure 5-3).

Although descriptive statistics are informative, both study and comparison hospitals lost admissions between the base year and PY 1 and then increased each of the two following years. Between the base year and PY 3, there was an overall 0.8 percent increase for the comparison hospitals and an 8.9 percent increase for BIMC.[9] These volume changes might have affected the relative cost performance of the two groups, especially BIMC's as volume increases lessened the reliance of the demonstration on improving financial performance through cost cutting. Therefore, our analytic approach relies on a multivariate 2D analysis that is described and presented in the next section.

BIMC Episode Payments*: *Table 5-11 presents estimates for the model based on Equation 5-1 for the four payment measures for BIMC. The 2D coefficients ranged from $676 to $250 lower for episodes originating at BIMC than at comparison hospitals, but none of these coefficients were statistically significant. The 2D coefficients presented in Table 5-11 are an average for the combined three performance years. The participating hospital indicator was positive but not statistically significant for all four measures of episode payments. The performance period indicator shows that episode payments for comparison hospitals increased from $2,394 (measure 1) to $1,218 (measure 4). The coefficients were statistically significant at conventional levels for all four measures.

Year-specific 2D estimates (***Table 5-12***) indicate that BIMC had $832 smaller increase in average total episode payments (measure 1) for PY 1—it is the only year-specific 2D estimate that was statistically significant. Although mostly statistically insignificant, the 2D estimates for episode payment measures 1, 2, and 4 continually become smaller in absolute value over time. The exception was for PY 2 for measure 3, the measure that reintroduces index outlier payments back into the episode payment measure definition. The participating hospital indicator was positive but not statistically significant for all four measures of episode payments—the magnitudes were similar to those in Table 5-11. The performance period indicators shows that episode payments for comparison hospitals increased from annually for all four episode payment measures and were all statistically significant at conventional levels.

Together with Table 5-9's 2D estimates, the year-specific 2D results indicate the impact of the demonstration on Medicare payments was inconsistent across years. The 2D PY 1 effect on total episode payments suggests there was an initial degree of practice behavior changed was achieved but then not sustained.

As the coefficients for the other variables are similar in both tables, the results from Table 5-11 are presented here. The two variables most directly associated with payments and the health condition of beneficiaries—DRG weight and the patient's prospective HCC risk score—both have positive, statistically significant effects. The coefficient for the DRG weight is nearly

[9] BIMC received a large share of patients previously admitted to a nearby hospital (St. Vincent's) that closed in 2010.

Table 5-11
BIMC episode payment regressions

Explanatory variable	1. Total episode payments			2. Total except the IPPS index hospital payments			3. #2 plus IPPS outlier payments			4. Pre-adm & post discharge		
	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	-11.27	242.10	0.964	-257.24	156.92	0.125	-56.82	238.09	0.815	-151.87	154.96	0.345
Patient age 70–74	400.66	127.89	0.008	475.69	106.23	0.001	445.74	129.91	0.004	479.05	113.01	0.001
Patient age 75–79	1,370.93	153.05	0.000	1,282.82	129.28	0.000	1,364.69	143.74	0.000	1,248.88	131.06	0.000
Patient age 80 or more	2,681.82	157.77	0.000	2,472.46	161.00	0.000	2,632.94	164.40	0.000	2,431.61	165.36	0.000
Female	387.53	164.03	0.034	426.12	174.66	0.030	430.21	175.62	0.029	361.97	162.69	0.044
Nonwhite	-1,013.65	194.43	0.000	-866.76	177.51	0.000	-1,015.29	194.28	0.000	-738.63	180.70	0.001
DRG weight	11,910.05	350.64	0.000	1,846.85	119.12	0.000	2,448.05	259.35	0.000	1,031.33	99.87	0.000
HCC risk score	242.15	23.71	0.000	194.49	13.62	0.000	229.05	21.91	0.000	182.62	13.55	0.000
Intern-/resident-to-bed ratio	4,249.95	2,592.08	0.125	-741.91	2,214.14	0.743	69.12	2,789.21	0.981	-771.37	2,004.91	0.707
Hospital beds	-0.80	1.03	0.451	-1.40	1.04	0.200	-1.55	1.08	0.176	-1.26	0.97	0.216
DSH adjustment factor (operating)	11,309.53	2,604.91	0.001	1,067.65	1,910.66	0.586	1,186.76	2,910.80	0.690	1,273.16	1,773.45	0.486
Participating hospital indicator	394.05	955.36	0.687	331.95	825.99	0.694	442.58	1,034.26	0.676	116.07	756.66	0.880
Performance period indicator	2,394.46	185.75	0.000	1,320.98	124.78	0.000	1,549.06	181.16	0.000	1,218.11	130.14	0.000
2D estimator	-676.51	425.01	0.135	-250.02	344.55	0.481	-322.03	429.16	0.466	-309.32	327.68	0.362
Constant term	-2,255.45	1,208.69	0.085	5,222.37	985.02	0.000	4,073.82	1,201.58	0.005	4,889.11	927.99	0.000
R^2	0.6539	—	—	0.0695	—	—	0.0894	—	—	0.0322	—	—
Number of observations	303,295	—	—	303,295	—	—	303,295	—	—	303,295	—	—

NOTE: 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

Table 5-12
BIMC episode payment regressions, year-specific 2D effects

Explanatory variable	1. Total episode payments			2. Total except the IPPS index hospital payments			3. #2 plus IPPS outlier payments			4. Pre-adm & post discharge		
	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	-21.80	243.49	0.930	-265.35	157.04	0.115	-65.11	239.04	0.790	-158.97	155.22	0.324
Patient age 70–74	397.16	127.97	0.008	474.15	105.72	0.001	443.75	129.54	0.005	477.80	112.47	0.001
Patient age 75–79	1,377.20	153.47	0.000	1,286.13	129.46	0.000	1,368.39	143.78	0.000	1,251.69	131.45	0.000
Patient age 80 or more	2,677.12	157.25	0.000	2,469.15	161.28	0.000	2,629.39	164.22	0.000	2,428.74	165.63	0.000
Female	391.43	163.74	0.033	428.91	174.28	0.029	433.27	175.10	0.028	364.38	162.38	0.043
Nonwhite	-1,042.83	196.17	0.000	-883.82	180.51	0.000	-1,033.80	196.41	0.000	-753.27	184.08	0.001
DRG weight	11,912.15	350.42	0.000	1,848.15	119.35	0.000	2,449.41	259.27	0.000	1,032.45	100.18	0.000
HCC risk score	241.34	23.59	0.000	193.91	13.59	0.000	228.49	21.86	0.000	182.11	13.54	0.000
Intern-/resident-to-bed ratio	4,122.44	2,480.56	0.120	-830.43	2,144.72	0.705	-17.90	2,720.29	0.995	-848.99	1,942.43	0.669
Hospital beds	-0.83	1.00	0.423	-1.41	1.03	0.192	-1.56	1.07	0.169	-1.27	0.96	0.207
DSH adjustment factor (operating)	11,697.59	2,657.40	0.001	1,282.13	1,942.15	0.521	1,415.49	2,953.86	0.640	1,457.26	1,799.87	0.433
Participating hospital indicator	340.89	934.80	0.721	298.18	810.64	0.719	407.55	1,025.69	0.698	86.78	742.29	0.909
Performance period indicator												
Performance year 1	1,658.72	155.79	0.000	908.72	86.20	0.000	1,077.14	165.34	0.000	868.26	85.16	0.000
Performance year 2	2,678.58	224.75	0.000	1,323.14	138.38	0.000	1,609.22	216.34	0.000	1,206.27	147.82	0.000
Performance year 3	2,864.56	222.44	0.000	1,739.85	176.83	0.000	1,969.67	196.34	0.000	1,587.33	186.20	0.000
2D estimator												
Performance year 1	-832.21	334.05	0.027	-356.15	280.03	0.226	-247.10	344.69	0.486	-426.26	260.19	0.125
Performance year 2	-786.35	465.15	0.115	-279.14	344.96	0.433	-503.91	477.87	0.311	-311.73	329.54	0.361
Performance year 3	-484.40	472.94	0.324	-158.53	400.42	0.699	-249.38	460.90	0.598	-229.26	388.51	0.565
Constant term	-2,271.48	1,193.96	0.079	5,222.61	967.79	0.000	4,069.41	1,193.09	0.005	4,890.26	912.52	0.000
R^2	0.6543	—	—	0.0701	—	—	0.0899	—	—	0.0326	—	—
Number of observations	303,295	—	—	303,295	—	—	303,295	—	—	303,295	—	—

NOTE: 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

$12,000 in the total episode payment regression (measure 1), but falls to under $2,500 in the other three episode payment regressions. The DRG weight coefficient is artificially higher in the total payments regression that includes the DRG inlier payment because these payments are determined in a compound, nonlinear fashion. In the other three episode payment measures, DRG weight clearly plays a cost-increasing role for outlier and physician payments as well as adding significantly ($1,031) to pre- and post-discharge payments, even after controlling for other beneficiary characteristics. The IRB ratio and the DSH operating adjustment factor are major contributors to total episode payments because of their role in determining inlier hospital payments. However, only the DSH coefficient is statistically significant. Greater teaching intensity does not add extra physician and outlier payments. The coefficient on the HCC risk score ranges from $242 (measure 1) to $14 (measure 4).

Even after controlling for gender, race, DRG, and HCC score, beneficiary age still shows a strong positive effect on all four payment measures. The coefficients for females are positive and statistically significant. The coefficients for race are negative and statistically significant, indicating that nonwhites were less expensive for Medicare, possibly because they received less physician and post-acute care.

BIMC Length of Stay: The 2D coefficient for LOS was negative but not statistically significant using the Equation 5-1 specification (***Table 5-13***). The year-specific 2D estimates are positive for PY 1 and 3 but negative for PY 2—none were statistically significant (***Table 5-14***). The coefficient for the DRG weight in BIMC's LOS regression was negative instead of positive, probably because of the strong positive effect of the national average LOS in the model. The coefficient for the HCC risk score is positive and statistically significant. The patient age variables have the same signs and patterns of coefficients as in the episode payment regressions. Female and nonwhite patients have longer LOSs than males and whites. The IPPS DSH adjustment factor was positive and statistically significant. This finding may suggests that beneficiaries treated in hospitals with lower-income patients are possibly more difficult to place after inpatient hospital discharge.

BIMC IPPS Outlier Payments: The two specifications of hospital outlier payments were analyzed. The estimated 2D effect, based on the Equation 5-1 specification, for total outlier payments per beneficiary episode was −$77 but was not statistically significant (***Table 5-13***). Furthermore, the adjusted R^2 of 0.04 was extremely low. The year-specific 2D estimates jumped around but none were statistically significant (***Table 5-14***). Next, we deconstructed outlier payments into two parts: (1) the likelihood of an outlier and (2) outlier payments conditional on being an outlier. The 2D estimate of the likelihood of incurring an outlier payment was positive but not significant at conventional levels for the Equation 5-1 specification (Table 5-13). The year-specific 2D estimates of the likelihood of incurring an outlier payment were positive for PY 1 and PY 3 and negative for PY 2, but none were significant at conventional levels (Table 5-14). No effect was found on outlier payments for beneficiaries actually incurring an outlier payment (Table 5-13). While the year-specific 2D estimates were not statistically significant, the coefficients still displaced an interesting pattern of moving from a positive value in PY 1 to large negative values during the last two performance years.

BIMC Inpatient Physician Payments: The 2D coefficient in the inpatient physician payments regression based on Equation 5-1 (***Table 5-15***) was $59.30 and not statistically

Table 5-13
BIMC length of stay and IPPS outlier payment regressions

							Decomposition of IPPS outlier payments					
	Length of stay			IPPS outlier payment amount			Logit results for the likelihood of an IPPS outlier hospitalization			OLS results on IPPS outlier payments for discharges with an outlier payment		
Explanatory variable	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	0.036	0.067	0.597	116.03	101.63	0.274	0.196	0.129	0.128	1,773.98	1,510.38	0.261
Patient age 70–74	0.089	0.048	0.088	-35.00	33.33	0.313	0.004	0.049	0.937	-914.60	1,001.91	0.378
Patient age 75–79	0.356	0.052	0.000	50.82	30.24	0.117	0.144	0.076	0.056	-341.70	689.78	0.629
Patient age 80 or more	0.737	0.074	0.000	47.49	62.41	0.460	0.144	0.124	0.244	674.76	697.39	0.351
Female	0.244	0.033	0.000	-25.16	25.15	0.335	-0.152	0.025	0.000	-147.55	230.72	0.534
Nonwhite	0.293	0.085	0.004	-156.56	143.46	0.295	-0.347	0.205	0.091	-1,702.93	1,044.23	0.127
DRG weight	-0.120	0.051	0.034	232.07	61.00	0.002	-0.199	0.169	0.237	654.78	456.28	0.175
HCC risk score	0.058	0.008	0.000	19.19	6.86	0.015	0.023	0.002	0.000	354.90	189.76	0.084
Intern-/resident-to-bed ratio	0.355	0.552	0.531	863.23	1,042.97	0.423	1.687	1.723	0.328	11,585.03	12,261.76	0.362
Hospital beds	-0.002	0.000	0.000	-0.04	0.74	0.953	0.000	0.001	0.918	4.69	7.49	0.542
DSH adjustment factor (operating)	1.981	0.567	0.004	-256.11	1,585.95	0.874	-1.707	3.411	0.617	5,452.24	6,760.13	0.434
Average LOS for DRG	1.238	0.035	0.000	190.04	66.48	0.013	0.199	0.072	0.006	1,383.35	322.74	0.001
Participating hospital indicator	0.721	0.230	0.008	110.46	526.69	0.837	-0.150	1.074	0.889	-3,697.77	5,737.87	0.530
Performance period indicator	0.049	0.103	0.641	274.47	152.94	0.096	0.657	0.163	0.000	2,064.61	982.91	0.056
2D estimator	-0.021	0.121	0.862	-76.99	226.74	0.740	0.366	0.382	0.338	-518.85	2,490.85	0.838
Constant term	0.231	0.334	0.501	-1,458.83	923.49	0.138	-5.490	1.434	0.000	-7,690.45	5,853.48	0.212
R2 (pseudo for logit)	0.3867	—	—	0.0393	—	—	0.076	—	—	0.175	—	—
Number of observations	303,295	—	—	303,295	—	—	303,295	—	—	7,307	—	—

NOTE: 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system; LOS = length of stay; OLS = ordinary least-squares.

SOURCE: RTI processing of Medicare claims.

Table 5-14
BIMC length of stay and IPPS outlier payment regressions, year-specific 2D effects

	Length of stay			IPPS outlier payment amount			Decomposition of IPPS outlier payments					
							Logit results for the likelihood of an IPPS outlier hospitalization			OLS results on IPPS outlier payments for discharges with an outlier payment		
Explanatory variable	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |	Coef.	Std. Err.	P > | t |
Patient age 0–64	0.031	0.067	0.653	115.01	102.79	0.283	0.198	0.133	0.136	1,747.70	1,487.26	0.261
Patient age 70–74	0.088	0.048	0.091	-35.56	33.50	0.308	0.004	0.049	0.935	-979.47	1,034.34	0.361
Patient age 75–79	0.357	0.052	0.000	51.24	30.36	0.115	0.145	0.076	0.056	-339.91	692.76	0.632
Patient age 80 or more	0.733	0.075	0.000	46.55	61.60	0.463	0.146	0.122	0.234	632.98	675.59	0.366
Female	0.245	0.033	0.000	-24.85	25.06	0.339	-0.152	0.024	0.000	-127.34	224.63	0.580
Nonwhite	0.285	0.085	0.005	-158.91	145.31	0.294	-0.345	0.207	0.096	-1,780.67	1,057.46	0.116
DRG weight	-0.127	0.051	0.027	230.48	61.44	0.002	-0.196	0.174	0.260	559.44	417.34	0.203
HCC risk score	0.057	0.008	0.000	19.11	6.73	0.014	0.023	0.002	0.000	350.55	188.70	0.086
Intern-/resident-to-bed ratio	0.318	0.572	0.588	861.01	1,039.91	0.423	1.700	1.727	0.325	12,364.03	12,391.10	0.337
Hospital beds	-0.002	0.000	0.000	-0.04	0.74	0.953	0.000	0.001	0.922	3.09	7.33	0.681
DSH adjustment factor (operating)	2.072	0.572	0.003	-234.02	1,606.53	0.886	-1.743	3.429	0.611	5,696.79	7,022.79	0.432
Average LOS for DRG	1.242	0.035	0.000	190.92	66.73	0.013	0.197	0.075	0.008	1,423.66	314.87	0.001
Participating hospital indicator	0.706	0.230	0.009	107.55	529.96	0.842	-0.145	1.077	0.893	-2,869.71	5,671.35	0.621
Performance period indicator												
Performance year 1	-0.158	0.070	0.041	192.16	124.44	0.147	0.673	0.195	0.001	-120.72	925.32	0.898
Performance year 2	0.080	0.127	0.539	328.10	181.74	0.094	0.677	0.219	0.002	3,346.13	1,396.46	0.032
Performance year 3	0.231	0.126	0.091	304.10	173.81	0.104	0.613	0.144	0.000	3,456.34	922.78	0.002
2D estimator												
Performance year 1	0.043	0.087	0.632	118.68	201.96	0.567	0.724	0.337	0.032	2,085.00	2,284.39	0.378
Performance year 2	-0.169	0.142	0.255	-233.40	259.19	0.384	-0.131	0.453	0.772	-2,060.19	2,838.86	0.481
Performance year 3	0.048	0.146	0.749	-107.19	229.70	0.648	0.402	0.373	0.281	-2,953.87	2,355.09	0.232
Constant term	0.224	0.343	0.525	-1,465.05	926.94	0.138	-5.490	1.431	0.000	-7,343.51	6,004.20	0.243
R2 (pseudo for logit)	0.3871	—	—	0.0394	—	—	0.077	—	—	0.1775	—	—
Number of observations	303,295	—	—	303,295	—	—	303,295	—	—	7,307	—	—

NOTE: 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system; LOS = length of stay; OLS = ordinary least-squares.

SOURCE: RTI processing of Medicare claims.

Table 5-15
BIMC physician inpatient payments regression

Explanatory variable	Coefficient	Standard error	P > | t |
Patient age 0–64	-105.36	19.44	0.000
Patient age 70–74	-3.36	23.08	0.886
Patient age 75–79	33.94	23.48	0.172
Patient age 80 plus	40.85	30.64	0.205
Female	64.15	16.86	0.002
Nonwhite	-128.13	19.72	0.000
DRG weight	815.53	39.71	0.000
HCC risk score	11.87	3.98	0.011
Intern-/resident-to-bed ratio	29.46	289.12	0.920
Hospital beds	-0.14	0.11	0.216
DSH adjustment factor (operating)	-205.51	210.61	0.347
Participating hospital indicator	215.89	96.81	0.044
Performance period indicator	102.88	27.49	0.002
2D estimator	59.30	33.67	0.102
Constant term	333.26	127.37	0.021
R^2	0.4608	—	—
Number of observations	303,295	—	—

NOTE: Physician inpatient payments are for the index hospitalization only. 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

significant (p = 0.102) with an adjusted R^2 of 0.46. Year-specific 2D coefficients ranged from $32.59 for PY 2 to $70 for both PY 1 and PY 3—only the PY 1 coefficient was statistically significant (***Table 5-16***). These results indicate the Gainsharing demonstration did not seem to have an effect on inpatient physician payments. Physician payments for the youngest (mostly persons with disabilities) Medicare beneficiaries were lower than for other beneficiary age categories. Physician payments were lower for nonwhite beneficiaries but higher for females. The DRG weight and HCC scores both had positive coefficients.

Table 5-16
BIMC physician inpatient payments regression, year-specific 2D effects

Explanatory variable	Coefficient	Standard error	P > | t |
Patient age 0–64	-106.38	19.58	0.000
Patient age 70–74	-3.64	23.18	0.878
Patient age 75–79	34.44	23.32	0.164
Patient age 80 plus	40.40	30.67	0.210
Female	64.53	16.82	0.002
Nonwhite	-130.55	19.61	0.000
DRG weight	815.70	39.70	0.000
HCC risk score	11.80	3.99	0.011
Intern-/resident-to-bed ratio	18.56	285.32	0.949
Hospital beds	-0.14	0.11	0.213
DSH adjustment factor (operating)	-175.13	218.43	0.437
Participating hospital indicator	211.40	96.15	0.047
Performance period indicator			
Performance year 1	40.46	18.91	0.052
Performance year 2	116.87	29.81	0.002
Performance year 3	152.52	39.94	0.002
2D estimator			
Performance year 1	70.12	25.64	0.017
Performance year 2	32.59	35.30	0.373
Performance year 3	70.72	41.96	0.116
Constant term	332.36	127.54	0.022
R^2	0.4612	—	—
Number of observations	303,295	—	—

NOTE: Physician inpatient payments are for the index hospitalization only. 2D = difference-in-difference; DRG = diagnosis-related groups; DSH= disproportionate share hospital; HCC = hierarchical condition category; IPPS = inpatient prospective payment system.

SOURCE: RTI processing of Medicare claims.

5.5 Discussion

One goal of the Gainsharing demonstration was to reduce the hospital's internal costs at a level sufficient to generate savings that could be shared with physicians. Medicare savings were not required, although these changes in incentives were expected to change physician behavior in ways that might reduce Medicare's outlays per episode through reduced physician Part B charges. Since we only had one performance year for CAMC, we focus on findings from BIMC which operated for the full three years envisioned for the demonstration.

We found a transitory significant impact of the first year of the Gainsharing demonstration on BIMC's per-episode inpatient physician payments. The 2D coefficients for BIMC's four episode payment measures all had negative signs, but only one was statistically significant (total episode payments during PY 1). Despite BIMC's self-reported emphasis on reducing LOS as a source for reducing internal costs, BIMC's average LOS were slightly higher in PY 3 (possibly due to a slightly complex casemix) whereas the average LOS for the comparison hospitals continually fell during the course of the demonstration. This suggests that the internal cost savings were either driven by factors other than changes in the LOS or changes in physician billing behavior. Internal cost savings self-reported by BIMC did not reduce overall Medicare episode payments.

This page left intentionally blank.

SECTION 6
QUALITY OF CARE—CLAIMS-BASED AND MEDICAL RECORDS-BASED INDICATORS

One goal of the Medicare Gainsharing Demonstration is to evaluate mechanisms for hospitals and physicians to join forces to improve quality and efficiency of care. Evaluation of the demonstration therefore requires an assessment of the impact of gainsharing on the quality of care. To do this, it is necessary to understand the incentives that each gainsharing methodology offers for changes in hospital and physician behavior.

Strategies introduced by CAMC and BIMC to reduce internal hospital costs include reduced LOS, reduced inpatient diagnostic testing, and reduced use of specialist consultations. Other tactics may include increased coordination of care, improved transitions of patients across care settings, and the development of targeted case management for high-risk patients. These activities, however, should not result in a decline in quality of care. We examine three data sources in this report: (1) Medicare claims, (2) medical records abstractions, and (3) beneficiary surveys (in Section 7). The quality measures presented below and in the following chapter analyze three performance years for BIMC and one performance year for CAMC.

6.1 Data Sources and Measures

Administrative claims are a cost-effective means of measuring provider quality. Claims data are routinely collected as part of the delivery of hospital services and do not require additional data collection. These data include information on diagnoses, procedures, age, gender, admission source, and discharge status (Agency for Healthcare Research and Quality [AHRQ], 2007a,b). AHRQ developed four quality indicator (QI) modules that rely solely on inpatient claims data, in order to measure quality of care in inpatient or outpatient settings.

Two QI modules are relevant to the current analysis: inpatient quality indicators (IQIs) and patient safety indicators (PSIs). IQIs include inpatient mortality for selected medical conditions and surgical procedures, utilization rates for selected procedures (where there may be a question of over-, under-, or misuse), and volume rates for selected procedures (where a high volume may be associated with lower mortality). PSIs are rates of potentially avoidable complications and iatrogenic events that are adjusted with diagnosis related group (DRG) relative weights (e.g., postoperative complications, death in low-mortality DRGs, and decubitus ulcers).

Data used for the quality outcomes and analyses come from Medicare Part A inpatient claims from January 2007 through September 2011. The base year evaluation period is October 2007 through December 2007, the year 1 evaluation period is October 2008 through September 2009, the year 2 evaluation period is October 2009 through September 2010 and the year 3 evaluation period is October 2010 through September 2011. The level of analysis is the episode of care. We built quality analytical files from the episode of care finder files jointly developed with Medicare claims (standardized to CMS DRG Version 24 codes) from the Data Extraction System pulls by RTI, and based on the core analytic file prepared by the Actuarial Research Corporation (ARC). An episode of care is defined as the period beginning 14 days before the date of a qualifying admission and ending 30 days after discharge (thus requiring some data from

September 2007 and through December 2011.[10] Claims data were pulled for beneficiaries receiving care from the 12 intervention hospitals or the hospitals in the comparison group.

The sample consists of Medicare fee-for-service beneficiaries who have been continuously enrolled in both Medicare Part A and Part B and who have Medicare as their primary payer. Excluded from the analysis are beneficiaries enrolled in Medicare Part C; beneficiaries with end-stage renal disease; and beneficiaries receiving hospice care. Using the ID established for each episode of care, as well as the associated admission and discharge dates, we merged additional data needed to construct the quality analytical files. This includes information such as beneficiary race, state and county of residence, discharge status, details of admission, diagnoses coded, and procedures performed from Standard Analytical File (SAF) claims. We also merged data containing hierarchical condition category (HCC)-based risk scores (Pope et al., 2011). A number of variables were then constructed, including LOS, 30- and 90-day mortality, and discharge quarter. Certain variables, such as race, admissions source, and primary payer, were then recoded to match the AHRQ QI software specifications. Once constructed and validated, the quality analytic file was then processed with the APR-DRG grouper followed by the QI software to risk-adjust the data and calculate the individual QIs. The AHRQ software creates flags to indicate whether an admission counts toward the numerator for a given indicator. We appended these flags to the quality analytic file for use in the difference-in-differences analyses.

Quality Indicators From the Agency for Healthcare Research and Quality: Administrative claims are a cost-effective means of measuring provider quality. Claims data are routinely collected, are widely available, and do not require additional data collection. These data include information on diagnoses, procedures, age, gender, admission source, and discharge status.

AHRQ developed quality indicators (QIs), four modules that rely solely on inpatient claims data, to measure quality of care in inpatient or outpatient settings. Two QI modules are relevant to the current analysis: inpatient quality indicators (IQIs) and patient safety indicators (PSIs). IQIs include inpatient mortality for selected medical conditions and surgical procedures, utilization rates for selected procedures (where there may be a question of over-, under-, or misuse), and volume rates for selected procedures (where a high volume may be associated with lower mortality). PSIs are risk-adjusted rates of potentially avoidable complications and iatrogenic events (e.g., postoperative complications, death in low-mortality DRGs, decubitus ulcers). The QI software is calibrated to risk adjust on the basis of a large proportion of the U.S. population. To reflect the population affected by this demonstration, and to risk adjust the quality measures more accurately, we recalibrated the software to use the 2007–2010 Medicare population as the reference population.[11] In addition to inpatient mortality (IQIs) and

---

[10] In the case of a beneficiary who is an inpatient of a hospital or skilled nursing facility, or who is covered by home health on the date that an episode of care would otherwise begin, the episode will begin on the day after discharge. Same-day transfers in from another IPPS hospital are excluded. Transfers from a skilled nursing facility or home health create a new episode. Same-day IPPS transfers out terminate the episode of care.

[11] See http://www.qualitynet.org for more information.

complications (PSIs), we also examined two simpler measures: 30-day mortality and readmissions.

The AHRQ IQIs are a set of measures providing rates of volume of specific high-technology, or highly complex, procedures; mortality indicators for certain inpatient procedures; mortality indicators for certain inpatient conditions; and utilization rates for certain procedures that vary greatly across hospitals. We used the AHRQ IQI software to calculate the rate of mortality for each of the conditions or procedures below.[12]

- AMI
- CHF
- stroke
- gastrointestinal hemorrhage (GI hemorrhage)
- hip fracture
- pneumonia
- CABG

The AHRQ PSIs are a set of measures providing rates of potentially preventable complications and other iatrogenic events that occur in the hospital setting. These are limited to cases in which a secondary diagnosis code indicates a potentially preventable complication. The PSIs include 20 provider-level indicators and 7 area-level indicators. We focused on the following 13 indicators that are appropriate for the Medicare population. (An asterisk shows that the indicator is included in CMS' hospital-acquired condition [HAC]/present on admission [POA] payment penalty program.)

- physiologic and metabolic derangements
- postoperative respiratory failure
- postoperative pulmonary death in low-mortality DRGs
- pressure ulcer*
- death among surgical patients

[12] Because CAMC chose to participate in the demonstration for only a select group of DRGs, only the IQI measures for AMI, CHF, and CABG are relevant. Because of small numbers, we did not measure the CABG mortality rate for BIMC.

- iatrogenic pneumothorax
- central venous catheter-related bloodstream infections*
- postoperative hip fracture*
- postoperative hemorrhage or hematoma
- postoperative embolism or deep vein thrombosis*
- postoperative sepsis
- postoperative wound dehiscence
- accidental puncture or laceration

The level of analysis was the episode of care. We built quality-of-care analytic files from the episode of care finder files linked to Medicare claims (standardized to CMS DRG Version 24 codes). An episode of care is generally defined as the period beginning 14 days before the date of a qualifying admission and ending 30 days after discharge.[13]

Claims data were pulled for beneficiaries receiving care from either the demonstration hospitals (CAMC and BIMC) or their comparison hospitals. At CAMC and its comparison hospitals, the episodes were limited to those with specific cardiac DRGs. The sample consisted of Medicare fee-for-service beneficiaries who had been continuously enrolled in both Medicare Part A and Part B and who had Medicare as their primary payer. Excluded from the analysis were beneficiaries enrolled in Medicare Part C, beneficiaries with end-stage renal disease, and beneficiaries receiving hospice care. Using the ID established for each episode of care, as well as the associated admission and discharge dates, we merged additional data needed to construct the quality analytical files (such as information on beneficiary race, state and county of residence, discharge status, details of admission, diagnoses coded, and procedures performed) from Standard Analytical File claims.

A number of variables were then constructed (including LOS, 30-day mortality, and discharge quarter). Certain variables (such as race, admissions source, and primary payer) were then recoded to match the AHRQ QI software specifications. Once constructed and validated, the quality analytic file was then processed with the QI software to risk-adjust the data and calculate the quality measures.

Medical Record Abstractions: Although claims data are able to provide measures of various patient outcomes that result from the provision of health care, they offer only limited

[13] In the case of a beneficiary who is an inpatient of a hospital or SNF, or who is covered by home health on the date that an episode of care would otherwise begin, the episode begins on the day after discharge. Same-day transfers in from another IPPS hospital are excluded. Transfers from a SNF or home health create a new episode. Same-day IPPS transfers out terminate the episode of care.

insight into *how* that care was provided. To fully assess the impact of the Medicare Gainsharing Demonstration on quality of care, it was also necessary to examine possible changes in how care was delivered in the demonstration and comparison hospitals. The level of detail necessary to generate information on the process of care is available in patient medical records.

Currently, CMS has a number of hospital-based quality initiatives that yield data sets that were applicable for this evaluation. The first of these initiatives is the current Hospital Inpatient Quality Reporting (IQR) program,[14] which collects data from hospitals on designated quality measures. Hospitals that successfully report designated quality measures are eligible for a higher annual update to their payment rates. Reported to CMS quarterly, the Hospital IQR data include 27 measures related to process of care for three conditions that are common to Medicare beneficiaries and often require hospitalization, as well as on processes relevant to the Surgical Care Improvement Project (SCIP).[15]

The three conditions covered by the Hospital IQR are AMI, heart failure (HF), and pneumonia. There are eight measures related to AMI care, four measures related to HF care, seven measures that address pneumonia care, and seven measures related to SCIP. Each of these evidence-based measures assesses treatment processes that are related to positive outcomes. Data from a sample of patient charts are converted to rates. The construct of each measure is such that more is better (e.g., achieving a rate of 100 percent indicates that a particular process of care was followed for each patient in the sample). The Hospital IQR data submissions are validated and standardized, allowing for comparison between hospitals.

The Hospital IQR data measures for each of the three conditions and SCIP are as follows.

- AMI (8 measures):
  - aspirin at arrival
  - aspirin prescribed at discharge
  - angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) for left ventricular systolic dysfunction (LVSD)
  - adult smoking cessation advice and counseling
  - beta blocker prescribed at discharge
  - beta blocker at arrival
  - fibrinolytic therapy received within 30 minutes of hospital arrival

---

[14] Hospital IQR was formerly known as the Reporting Hospital Quality Data for Annual Payment Update or RHQDAPU program.

[15] See http://www.qualitynet.org for an overview of the Hospital IQR program.

  - PCI received within 90 minutes of hospital arrival

- HF (4 measures):
  - evaluation of left ventricular systolic (LVS) function
  - ACE inhibitor or ARB for LVSD
  - smoking cessation advice and counseling
  - discharge instructions

- Pneumonia (7 measures):
  - oxygenation assessment
  - pneumococcal vaccination
  - blood cultures performed in the ER before initial antibiotic received in hospital
  - smoking cessation advice and counseling
  - initial antibiotic received within 6 hours of hospital arrival
  - initial antibiotic selection for community-acquired pneumonia (CAP) in immunocompetent patient
  - influenza vaccination

- SCIP (7 measures):
  - prophylactic antibiotic received within 1 hour before surgical incision
  - prophylactic antibiotic selection for surgical patients
  - prophylactic antibiotic discontinued within 24 hours after surgery end time
  - surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours before surgery to 24 hours after surgery
  - cardiac surgery patients with appropriate hair removal

We analyzed baseline and demonstration period data for the above measures. For each measure a ratio was created in which the numerator was the number of patients receiving a specific intervention (e.g., aspirin at arrival for AMI patients) and the denominator was the total number of adult patients eligible to be included in the numerator (i.e., AMI patients with a known aspirin allergy would be excluded in the AMI example above). The ratio was interpreted as the percentage of eligible patients who received the intervention. We also calculated a composite measure for each topic (AMI, HF, pneumonia, and SCIP). Annual Hospital IQR data

were used. The comparison group percentage was calculated by summing the data from the relevant period across the comparison hospitals for each measure. Composite measures were created by summing the numerators and denominators for each measure in a topic and using those sums as a numerator and denominator.

6.2 Methods

Thirty-Day Postdischarge Methodology: The Medicare Gainsharing Demonstration could potentially provide an incentive to reduce patients' LOS. There are likely many cases in which a patient's LOS is longer than medically necessary because of hospital inefficiencies (e.g., the physician not being available to sign discharge orders) that could be improved as a result of the gainsharing agreement between hospitals and physicians. There may be cases in which a shorter LOS may not be medically appropriate (a quality consideration) and could lead to readmission or shifting of care to another facility (which would affect the cost to Medicare).

The demonstration design utilizes a 30-day postdischarge period in defining an episode of care to account for costs of readmissions to the same hospital and costs associated with postdischarge care. Because quality of care is also an issue, changes in indicators such as LOS, 30-day mortality, readmissions within 30 days of discharge, and IQIs need to be examined as well.[16]

Risk Adjustment: The AHRQ QI software uses the APR-DRG risk adjustment grouper developed by 3M Corporation to risk-adjust all data on patients in DRGs relevant to IQI measures for patient severity (mortality risk). The grouper generates a severity score for each episode of care. There are four subclasses of mortality risk, 1 through 4, with 1 representing minor risk of mortality and 4 representing extreme risk of mortality.[17]

Difference-in-Differences Analysis: To estimate the impact of hospital gainsharing on hospital quality, we used a 2D analysis following the methodology described below, with one difference over time (the base year and the performance year) and one across subjects (the demonstration hospitals and the comparison groups). Subtracting the baseline difference in

---

16 Thirty-day mortality is calculated on the basis of date of admission, not date of discharge.

17 APR-DRGs are an enhanced extension of the basic DRG concept developed by 3M's Clinical Research Group, the National Association of Children's Hospitals and Research Institutes, and several physician groups.

Whereas DRGs focus on the Medicare population, APR-DRGs describe a complete cross-section of acute care patients and are specifically designed to adjust data for severity of illness (How sick is the patient?) and risk of mortality (How likely is it that the patient will die?). The fundamental principle of APR-DRGs is that the severity of illness and risk of mortality are both dependent on the patient's underlying condition. High severity of illness and risk of mortality are characterized by multiple serious diseases and the interactions between the disorders.

The 3M APR-DRG methodology is the most widely used severity-of-illness and risk-of-mortality adjustment tool available today. It has become the standard for adjusting large volumes of data to account for differences related to the individual's severity of illness or risk of mortality. As a result, the focus can be on the differences in clinical care, thus providing equitable comparisons of quality and cost of care. APR-DRGs are also recognized as the tool of choice by commissions, state agencies, and others who disseminate comparative performance data to regulators, payers, and the general public.

hospital quality from the demonstration difference eliminates any selection bias caused by the observable differences in hospitals, as long as the differences are fixed in time.

Nonlinear models will be estimated to determine the impacts on hospital quality and patient safety. For nonlinear models (e.g., logit, probit, and Poisson count models), the estimated coefficient on the interaction term, $\hat{\gamma}_3$, cannot simply be exponentiated to estimate the 2D effect because the model is, in fact, nonlinear. Because the patient and hospital characteristics interact in a multiplicative rather than a linear way, the mean of the differences between groups is not equal to the difference in mean differences. The standard method to derive 2D numerical estimates involves simulations (described in detail below). In these simulations, four dependent variables are estimated for each observation (episode) in the sample. Aside from the demonstration status (D) and the pre/post (T) variables, actual values for all of the other explanatory variables are used. Because there are 2 values each for D and T, four separate estimates of the dependent variable are calculated as follows:

- For each observation i, a simulated dependent variable ($\hat{P}_i$) is calculated as if the observation is for an episode in the pre-period by setting D to one and T to zero—see Cell 1 in ***Figure 6-1***.

- For each observation i, $\hat{P}_i$ is calculated as if the observation is for an episode in the post period by setting D to one and T to one—see Cell 2 in Figure 6-1.

- For each observation i, $\hat{P}_i$ is calculated as if the observation is for an episode in the pre period by setting D to zero and T to zero—see Cell 3 in Figure 6-1.

- For each observation i, $\hat{P}_i$ is calculated as if the observation is for an episode in the post period by setting D to zero and T to one—see Cell 4 in Figure 6-1.

The estimated probabilities, $\hat{P}_i$, are derived from the logit regression by the following transformation:

$$\hat{P}_i = \frac{e^{X_i'\hat{\beta}}}{1+e^{X_i'\hat{\beta}}} \qquad (6\text{-}1)$$

where X represents all explanatory variables and not just the patient and hospital characteristics. The X-characteristics (outlined in Section 5 above) are specific to a beneficiary and hospital in each time period.

Figure 6-1
Components for difference-in-differences calculations for nonlinear models

Group	Time period		Changes for each group (post minus pre)
	T = 0 (pre)	T = 1 (post)	
Demo participant (D = 1)	1 $\hat{P}_{i,D=1,T=0}$	2 $\hat{P}_{i,D=1,T=1}$	$\Delta PP_{i,D=1}$ Cell 2 minus Cell 1
Comparison (D = 0)	3 $\hat{P}_{i,D=0,T=0}$	4 $\hat{P}_{i,D=0,T=1}$	$\Delta PP_{i,D=0}$ = Cell 4 minus Cell 3

NOTES:

D denotes the dummy variable used to distinguish between demonstration participants and the comparison population, and T denotes the dummy variable used to distinguish between the pre and post periods.

The numbers in the shaded boxes are cell numbers.

For each observation, pre/post changes (ΔPP) are calculated as if the observation were for an episode at a demonstration hospital (ΔPP_i,D=1) and as if the observation were for an episode at a comparison hospital (ΔPP_i,D=0). The demonstration 2Ds effect for each episode is calculated by subtracting the comparison hospital pre/post change from the participating hospital pre/post change:

$$2D_i = \Delta PP_{i,D=1} - \Delta PP_{i,D=0}. \quad (6\text{-}2)$$

The average demonstration effect is then estimated by calculating the mean of the individual observation demonstration effects.

Limitations: Claims provide a cost-effective, easily accessible source of quality data, but they are not without limitations. Claims data are a poor source of information for measuring many process measures (and patient risk factors) because of their limited clinical information. We addressed this concern by balancing use of claims-based quality measures with those based on medical chart abstraction (presented below) and patient surveys (presented in the following chapter). In addition, the measures generated by the QI software are not standardized to account for variation in hospital volume, and therefore they are not appropriate for direct comparison *between* hospitals. The software generates observed, expected, and risk-adjusted rates of mortality or other outcome (described below). In the Chapter 6 Appendix tables, we employed an indirect standardization (by focusing on the ratio of observed to expected rates) to address the lack of standardization and allow for some comparisons. Because the Medicare population is not

comparable to the Healthcare Cost and Utilization Project (HCUP) population used to calculate the expected and risk-adjusted rates, we employed a recalibration technique to align the reference group with the Medicare population.[18] Finally, many conditions and procedures have only a small number of observations at the provider level. Thus, the resulting confidence intervals of the estimates are wide and the estimates may not be very precise.

6.3 Results

We analyzed baseline and performance year measures of two patient outcomes: 30-day mortality and readmissions. Hospitals trying to achieve savings may target reducing LOS. It is possible that some patients may be discharged too soon, which could result in a readmission to the hospital or even death. Therefore we considered 30-day mortality (mortality that occurs within 30 days of the relevant admission) and readmissions to the hospital as well. The measure of 30-day mortality is a flag (yes or no) indicating whether the patient died within 30 days of the admission that triggered the qualifying episode of care.

An all-cause readmission is based on the discharge associated with the qualifying episode of care admission. It is defined as any inpatient hospital admission for any condition, to any facility that occurs at least 1 day after and within 30 days of the related discharge. Therefore same-day transfers to another facility are not counted as readmissions in these analyses.

6.3.1 Charleston Area Medical Center

Counts of episodes of care from the base period and demonstration period are presented in ***Table 6-1***. These are used in the following analyses for CAMC and the hospitals that compose the comparison group. The demonstration period for CAMC and its comparison hospitals is 1 year, October 1, 2008, through September 30, 2009. Both CAMC and the group of comparison hospitals had fewer qualifying episodes of care during the demonstration period than in the base period. The number of episodes at CAMC declined 3.1 percent, from 882 in the base period to 855 in the demonstration period. The comparison hospitals had 8 percent fewer episodes in the demonstration period (5,622) than in the base period (6,112).

Table 6-1
Base year and Year 1 episodes of care: CAMC and comparison hospitals

Hospital	Base period episodes	Demo period episodes	Percentage change in episodes
Charleston Area Medical Center (CAMC)	882	855	−3.1%
Comparison hospitals	6,112	5,622	−8.0%

SOURCE: 2007–2010 Medicare inpatient prospective payment system claims.

[18] See www.qualitynet.org for detailed instructions on recalibrating the reference population for IQIs and PSIs.

Thirty-Day Mortality: Rates of 30-day mortality for the base period and demonstration period are presented in ***Table 6-2***. At CAMC the 30-day mortality rate was 2.61 percent in the base period and 2.46 percent in the demonstration period. This is an almost 6 percent decrease in the rate in the demonstration period relative to the base period. Among the comparison hospitals, the 30-day mortality rate also decreased from 2.54 percent in the base period to 2.42 percent during the demonstration, almost a 5 percent decrease in the rate between the two periods.

Table 6-2
Base year and Year 1 rates of 30-day mortality: CAMC and comparison hospitals

Hospital	Base period deaths	Base period 30-day mortality rate	Demo period deaths	Demo period 30-day mortality rate	Demo period rate/Base period rate
Charleston Area Medical Center (CAMC)	23	2.61	21	2.46	0.94
Comparison hospitals	155	2.54	136	2.42	0.95

SOURCE: 2007–2010 Medicare IPPS claims.

Readmissions: Table 6-3 presents all-cause readmissions for CAMC and for the group of comparison hospitals in the base and demonstration periods. Readmissions were counted if they occurred within 30 days of discharge from the qualifying hospital stay, regardless of where the readmission occurred. The readmission rate at the comparison hospitals was nearly identical in both periods, at 15.30 percent in the base year and 15.26 percent in the demonstration period. In contrast, during the base year, CAMC had a 15 percent lower rate of readmissions (13.95 percent) relative to the demonstration period (16.02 percent).

Table 6-3
Base year and Year 1 rates of 30-day readmissions: CAMC and comparison hospitals

Hospital	Base period readmissions	Base period readmission rate	Demo period readmissions	Demo period readmission rate	Demo period rate/base period rate
Charleston Area Medical Center (CAMC)	123	13.95	137	16.02	1.15
Comparison hospitals	935	15.30	858	15.26	1.00

SOURCE: 2007–2010 Medicare IPPS claims.

Inpatient Quality Indicator Findings: The AHRQ software generates observed, expected, risk-adjusted, and smoothed mortality rates. We focused our analysis on the

components of the observed rates. For each condition and procedure listed above (***Section 6.1***), the observed rate is the actual number of deaths per 1,000 patients admitted for that condition or procedure.

CAMC chose to participate in the demonstration for only a select group of cardiovascular DRGs; as a result only the IQI measures for AMI, CHF, and CABG are examined here. Because the IQIs measure mortality rates among patients treated for only these two conditions and one procedure, the population at risk for any measure may be quite small. To address this in the 2D analysis of impact of the demonstration, we used a simple composite variable to measure whether a patient died while being treated for any one of the relevant conditions or procedure during the time period (base period or demonstration period). ***Table 6-4*** presents base year and demonstration period mortality rates across the two conditions (AMI and CHF) and one procedure (CABG) for CAMC and the comparison hospitals. "Population at risk" refers to any patient who meets all exclusion criteria and is treated for one of the conditions or with the procedure above. "Deaths" refers to deaths of patients in the population at risk. Thus, the mortality rate is the observed mortality rate among patients treated for AMI or CHF, or who had a CABG, during the measurement period. We then divided the rate in the demonstration period by the base period rate. If the ratio was greater than one, the hospital had a higher mortality rate across the measures in the first year of the demonstration than in the base year.

Both CAMC and the group of comparison hospitals had higher rates of mortality among patients treated for AMI or CHF, or who received a CABG, during the demonstration than in the base period. The ratio of the demonstration rate to the base rate was 1.37 for CAMC and 1.28 for the comparison group, indicating that for both groups the demonstration period mortality rate for this subset of patients was approximately 30 percent higher than the rate in the base year. It is important to note that year-to-year change in rates may appear large because the population size for an individual hospital is relatively small. Therefore these numbers should be interpreted only along with the 2D analyses.

Table 6-4
Base year and Year 1 mortality rates per 1,000 episodes for selected cardiac conditions and procedures: CAMC and comparison hospitals

Hospital	Base period population at risk	Base period deaths	Base period mortality rate	Demo period population at risk	Demo period deaths	Demo period mortality rate	Demo period rate/base period rate
Charleston Area Medical Center (CAMC)	548	123	22.45	447	137	30.65	1.37
Comparison hospitals	2,955	935	31.64	2,125	858	40.38	1.28

SOURCE: 2007–2010 Medicare IPPS claims.

Patient Safety Indicator Findings: The AHRQ software generates observed, expected, and risk-adjusted rates of complications. For each complication listed above (***Section 6.1***), the observed rate is the actual number of occurrences per 10,000 patients.

Because the PSIs measure the rate of occurrence of adverse events, the number of any type events at a single hospital is likely to be quite small. To address this in the 2D analysis of the demonstration, we used a simple composite variable to measure whether a patient experienced at least one adverse event during an episode (thus each episode is counted only once) during the time period (base year or performance year 1). ***Table 6-5*** presents base year and demonstration period rates of adverse events in CAMC and its comparison hospitals. "Population at risk" refers to any patient who meets all exclusion criteria (essentially, all patients). "Events" refers to an occurrence of an adverse event. We divided the demonstration period rate by the base year rate. If the ratio was greater than one, the hospital had a higher rate of adverse events occurring during the demonstration than during the base period.

As shown in the table below, the rates of adverse events occurring was higher during the demonstration period than the base period for both CAMC and the group of comparison hospitals. The ratio of demonstration rate to base rate was 1.38 at CAMC and 1.24 for the comparison hospitals, indicating that (without accounting for any other factors), adverse events occurred around 30 percent more often during the demonstration period. The adverse events captured by the PSIs are particularly rare, and therefore the number of occurrences for a single category is very low. As such, the counts and rates presented here are extremely sensitive to large changes. These results should therefore be viewed with caution and in conjunction with the 2D analyses below.

Table 6-5
Base year and Year 1 rates of adverse events: CAMC and comparison hospitals

Hospital	Base period population at risk	Base period adverse events	Base period mortality rate	Demo period population at risk	Demo period adverse events	Demo period mortality rate	Demo period rate/base period rate
Charleston Area Medical Center (CAMC)	852	26	3.05	855	36	4.21	1.38
Comparison hospitals	5,797	156	2.69	5,622	188	3.34	1.24

SOURCE: 2007–2010 Medicare IPPS claims.

Medical Records-Based Measures: As noted in Section 6-1, we analyzed baseline and demonstration period data for a range of medical records based measures. These are presented by clinical grouping.

AMI Care: In the base year, the AMI composite measure for CAMC was 98 percent, as shown in ***Table 6-6***. This increased to 99 percent in the performance year. The measures aspirin at arrival, aspirin prescribed at discharge, and adult smoking cessation advice and counseling all remained the same at either 99 or 100 percent. ACE inhibitor or ARB for LVSD increased from 95 to 98 percent, and primary PCI received within 90 minutes of arrival increased the most, from 58 to 78 percent. Beta blocker prescribed at discharge decreased to 99 percent from 100 percent; beta blocker at arrival was 95 percent in base year, but there were no data for that measure in the demonstration period.

Table 6-6
Hospital process of care measures—AMI: CAMC and comparison hospitals

	Base year				PY1			
	CAMC		Comparison group		CAMC		Comparison group	
AMI care measure	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Aspirin at arrival	447	99%	2,705	98%	423	99%	2,559	98%
Aspirin prescribed at discharge	1,082	99%	6,393	98%	1,083	99%	6,182	98%
ACEI or ARB for LVSD	289	95%	1,402	90%	215	98%	1,164	93%
Adult smoking cessation advice/counseling	582	100%	2,718	99%	526	100%	2,655	99%
Beta blocker prescribed at discharge	1,077	100%	6,902	98%	1,038	99%	6,475	98%
Beta blocker at arrival[a]	244	95%	1,641	95%	N/A	N/A	N/A	N/A
Fibrinolytic therapy received within 30 minutes of hospital arrival[d]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Primary PCI received within 90 minutes of hospital arrival	55	58%	579	79%	65	78%	526	81%
AMI composite score	3,776	98%	22,340	97%	3,350	99%	19,561	97%

NOTE: ACEI = angiotensin-converting enzyme inhibitor; AMI = acute myocardial infarction; ARB = angiotensin receptor blocker; CAMC = Charleston Area Medical Center; LVSD = left ventricular systolic dysfunction; PCI = percutaneous coronary intervention; PY = performance year.

SOURCE: 2008–2010 Medicare Inpatient Quality Reporting data.

The AMI composite rate within the comparison group was 97 percent in both the base year and the performance year. Four measures—aspirin at arrival, aspirin prescribed at discharge, adult smoking cessation advice and counseling, and beta blocker at arrival—also remained the same for both years at either 98 or 99 percent. ACE inhibitor or ARB for LVSD and primary PCI received within 90 minutes of hospital arrival both increased, from 90 to 93 percent and from 79 to 81 percent, respectively. Beta blocker at arrival was 95 percent in the base year, but there were no data for the performance year for that measure.

HF Care: The HF process of care measures reported by CAMC and the comparison hospitals are presented in ***Table 6-7***. The composite score for all of the HF measures in the base year for CAMC was 95 percent. It decreased to 90 percent in the performance year. ACE inhibitor or ARB for LVSD was the only measure that did not decrease; it remained the same at 88 percent for both the base and demonstration periods. The other three measures all went down. Discharge instructions decreased the most, to 81 percent from 94 percent. Evaluation of LVS function and adult smoking cessation advice and counseling both dropped one percentage point from 98 to 97 percent and 100 to 99 percent, respectively.

Table 6-7
Hospital process of care measures—HF: CAMC and comparison hospitals

	Base year				PY1			
	CAMC		Comparison group		CAMC		Comparison group	
HF care measure	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Discharge instructions	648	94%	4,480	70%	691	81%	4,105	72%
Evaluation of LVS function	734	98%	5,347	94%	763	97%	4,995	96%
ACE inhibitor or ARB for LVSD	316	88%	2,155	87%	300	88%	1,834	91%
Adult smoking cessation advice/counseling	174	100%	1,071	97%	177	99%	992	96%
HF composite score	1,872	95%	13,053	85%	1,931	90%	11,926	87%

NOTE: ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; CAMC = Charleston Area Medical Center; HF = heart failure; LVSD = left ventricular systolic dysfunction; PY = performance year.

SOURCE: 2008–2010 Medicare Inpatient Quality Reporting data.

The comparison group did better at increasing the composite score from 85 to 87 percent and decreasing in only one measure, adult smoking cessation advice and counseling, which fell to 96 percent from 97 percent. The other three measures—discharge instructions, evaluation of LVS function, and ACE inhibitor or ARB for LVSD—increased from 70 to 72 percent, 94 to 96 percent, and 87 to 91 percent, respectively.

Pneumonia Care: As shown in ***Table 6-8***, CAMC's composite score for pneumonia care increased from 91 to 92 percent from the base year to the performance year. Four of its seven measures increased: pneumococcal vaccination from 87 to 92 percent, blood cultures performed before initial antibiotic from 93 to 96 percent, initial antibiotic received within 6 hours of hospital arrival from 88 to 90 percent, and influenza vaccination from 81 to 89 percent. Two measures (oxygenation assessment and adult smoking cessation advice and counseling) remained at 100 percent and 97 percent, respectively, whereas one measure, initial antibiotic selection for CAP in immunocompetent patients, decreased from 86 to 84 percent.

The composite score for the comparison group decreased to 89 percent from 90 percent. Pneumococcal vaccination increased from 83 to 88 percent and initial antibiotic received within 6 hours of hospital arrival increased from 88 to 90 percent. Four measures remained the same: oxygenation assessment at 100 percent, adult smoking cessation advice and counseling at 96 percent, initial antibiotic selection for CAP in immunocompetent patient at 88 percent, and influenza vaccination at 82 percent. Blood cultures performed in the ER decreased from 85 to 81 percent.

Table 6-8
Hospital process of care measures—Pneumonia: CAMC and comparison hospitals

	Base year				PY1			
	CAMC		Comparison group		CAMC		Comparison group	
Pneumonia care measure	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Oxygenation assessment	643	100%	3,340	100%	534	100%	2,300	100%
Pneumococcal vaccination	515	87%	2,918	83%	619	92%	2,678	88%
Blood cultures performed in the emergency department before initial antibiotic received in hospital	514	93%	2,361	85%	608	96%	2,275	81%
Adult smoking cessation advice/counseling	285	97%	1,468	96%	378	97%	1,397	96%
Initial antibiotic received within 6 hours of hospital arrival	249	88%	1,038	88%	618	90%	2,267	90%
Initial antibiotic selection for CAP in immunocompetent patient	352	86%	1,679	88%	444	84%	1,439	88%
Influenza vaccination	184	81%	1,144	82%	726	89%	3,201	82%
Pneumonia composite score	2,742	91%	13,948	90%	3,927	92%	15,557	89%

NOTE: CAMC = Charleston Area Medical Center; CAP = community-acquired pneumonia.
SOURCE: 2008–2010 Medicare Inpatient Quality Reporting data.

Surgical Care Improvement Project: The composite score for CAMC for the SCIP measures increased from 94 percent in the base year to 96 percent in the demonstration period. These scores are shown in ***Table 6-9***. Of the eight measures, one, surgery patients on beta-blocker therapy before arrival who received a beta-blocker during a preoperative period, had no data for either period. Four increased: prophylactic antibiotic received within 1 hour before surgical incision from 91 to 99 percent, prophylactic antibiotic selection for surgical patients from 97 to 99 percent, surgery patients with recommended venous thromboembolism from 94 to 95 percent, and surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours before surgery to 24 hours after surgery from 92 to 94 percent. One measure, surgery patients with appropriate hair removal, remained at 97 percent. Two measures decreased: prophylactic antibiotics discontinued within 24 hours after surgery end time from 95 to 94 percent and cardiac surgery patients with controlled 6:00 a.m. postoperative blood glucose from 93 to 90 percent.

The comparison group had a composite score of 89 percent in the base year, which increased to 92 percent in the first performance year. There were no data for surgery patients on beta-blocker therapy before arrival who received a beta-blocker during a preoperative period. All but one of the other measures—surgery patients with appropriate hair removal, which remained at 96 percent—increased from the base year to the performance year: prophylactic antibiotic received within 1 hour before surgical incision from 87 to 89 percent, prophylactic antibiotic selection for surgical patients from 95 to 97 percent, prophylactic antibiotics discontinued within 24 hours after surgery from 85 to 88 percent, surgery patients with

recommended venous thromboembolism from 92 to 94 percent , surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours before surgery to 24 hours after surgery from 85 to 89 percent, and cardiac surgery patients with controlled 6:00 a.m. postoperative blood glucose from 90 to 91 percent.

Table 6-9
Hospital process of care measures—SCIP: CAMC and comparison hospitals

	Base year				PY1			
	CAMC		Comparison group		CAMC		Comparison group	
Surgical care improvement measure	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Prophylactic antibiotic received within 1 hour before surgical incision	1,443	91%	10,286	87%	2,351	99%	10,570	89%
Prophylactic antibiotic selection for surgical patients	1,364	97%	7,909	95%	2,431	99%	10,644	97%
Prophylactic antibiotics discontinued within 24 hours after surgery end time	1,242	95%	9,711	85%	1,987	94%	10,039	88%
Surgery patients with recommended venous thromboembolism prophylaxis ordered	762	94%	5,425	92%	1,123	95%	7,198	94%
Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours before surgery to 24 hours after surgery	762	92%	5,425	85%	1,122	94%	7,196	89%
Cardiac surgery patients with controlled 6 a.m. postoperative blood glucose	309	93%	801	90%	951	90%	2,200	91%
Surgery patients with appropriate hair removal	787	97%	3,684	96%	2,361	97%	10,457	96%
Surgery patients on beta-blocker therapy before arrival who received a beta-blocker during the perioperative period	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
SCIP composite score	6,669	94%	43,241	89%	12,326	96%	58,304	92%

NOTE: CAMC = Charleston Area Medical Center; PY = performance year; SCIP = Surgical Care Improvement Project.

SOURCE: 2008–2010 Medicare Inpatient Quality Reporting data.

Multivariate Claims-Based Measures: We present the results of a 2D analysis of the following dependent variables: 30-day mortality, 30-day readmissions, IQI numerator (mortality), and PSI numerator (event). Each of these is a binary variable equal to one if the patient met the criteria in question (e.g., if the patient died within 30 days of admission). As

outlined above, we modeled logit regressions to estimate the impact of the demonstration on each of these outcome measures. The sample size for each measure is indicated below.

We used two measures that were based on the AHRQ QI measures presented in ***Section 6.1***. The first, IQI numerator (mortality), is equal to one if the patient was eligible for inclusion in the numerator of any of the IQI measures described above. The second, PSI numerator (adverse event), is equal to one if the patient was eligible for inclusion in the numerator of any of the PSI measures described above.

As described above, the IQIs are mortality rates for selected conditions or procedures. The relevant IQIs for CAMC are AMI, CHF, and CABG. The denominator consists of all patients treated for these conditions or receiving this procedure who met additional exclusion criteria imposed by the AHRQ methodology. A patient who died while being treated for any of these conditions was counted in the numerator for the overall IQI measure. Aggregating to this level addressed the fact that the individual denominators for any single IQI measure may be small.

Similarly, for the PSIs, we calculated one overall measure to indicate whether a patient experienced any of the 13 preventable complications captured by the PSIs we detailed in ***Section 6.1***. For example, the PSI numerator was equal to one if a patient who met all exclusion criteria developed a pressure ulcer while in the hospital. Aggregating to this level addressed the fact that the individual numerators for any single PSI measure may be small.

The focus of the Gainsharing demonstration was to give hospitals and physicians incentives to collaborate in an effort to generate cost savings. Although the demonstration did not specifically reward improvements in quality of care outside of any improvements that might generate savings, any strategies employed by the hospital should not have resulted in a decline in hospital quality. A decline in hospital quality would appear as a positive coefficient (i.e., each of the measures is a negative event). We calculated the percent 2Ds by dividing the coefficient from the logit equation by the mean outcome (e.g., 30-day mortality) for the demonstration hospitals as a group in the base year. We tested the significance of our estimates by constructing 95 percent confidence intervals using the standard errors calculated from a simple OLS regression model.

We found no statistically significant impact of the demonstration on any of the four measures (***Table 6-10***). Our results did indicate some decline in each of the measures, but none were statistically significant. However, it is unlikely that many changes resulting from the demonstration would have been measurable after only one performance year.

Table 6-10
Difference-in-differences estimates of claims-based quality and patient safety measures: Charleston Area Medical Center and comparison hospitals

Measure	Difference-in-differences logit	Percent difference-in-differences logit
30-day mortality (*N* = 13,189)	0.0037	14.49%
30-day readmissions (*N* = 13,189)	0.0143	9.40%
Inpatient quality indicators numerator (mortality; *N* = 5,919)	0.0003	0.81%
Patient safety indicators numerator (event; *N* = 12,862)	0.0013	3.89%

NOTE: None of the measures were statistically significant.

Medical Records-Based Measures: We present the CAMC results of a simple 2D analysis of the four IQR composite scores in ***Table 6-11***. The differences presented in the table represent the difference across time and the difference between each hospital and the comparison group. Across each of the composite scores, the differences are small, generally within +/−1 percent although as large as −7 percent for HF. Because these estimates are based on a single observation for each hospital in each period, we cannot test the statistical significance of these estimates.

Table 6-11
Difference-in-differences estimates of medical records-based measures: Charleston Area Medical Center and comparison hospitals

Measure	Base year vs. performance year
Acute myocardial infarction composite score	0%
Heart failure composite score	−7%
Pneumonia composite score	1%
Surgical Care Improvement Project composite score	−1%

6.3.2 Beth Israel Medical Center

Table 6-12 below presents the base year and year 3 counts of episodes of care used in the following intervention analyses and the hospitals that comprise the comparison group. BIMC had 9 percent more episodes in year 3 over the base year while the comparison group had an increase of 1 percent between the base year and year 3.

Table 6-12
Episodes of care for BIMC and comparison group during the base year and third implementation year

Hospital	Base period episodes	Year 3 episodes	Percent change in episodes
Beth Israel Medical Center	8,913	9,705	9%
Comparison hospitals	66,785	67,336	1%

SOURCE: 2007–2011 Medicare IPPS Claims.

30- and 90-Day Mortality: Rates of 30-day mortality for the base year and third implementation year are presented in ***Table 6-13***. In the base year 30-day mortality rates at BIMC was 4.38 percent. The comparison group hospitals had a 30-day mortality rate of 5.14 percent in the base year. The 30-day mortality rate at BIMC in the third implementation year was 4.40 percent; this was an overall increase of 0.46 percent between the base year and the third performance year. Across the comparison hospitals, 30-day mortality decreased by 4.05 percent between the base year and the third year of the demonstration.

Table 6-13
Base year and Year 3 rates of 30-day mortality for BIMC and its comparison group

Hospital	Base year deaths	Base year 30-day mortality rate	Year 3 deaths	Year 3 30-day mortality rate	Percentage change in mortality rate
Beth Israel Medical Center	390	4.38%	427	4.40%	0.46%
Comparison hospitals	3,431	5.14%	3,327	4.94%	-4.05%

SOURCE: 2007–2011 Medicare inpatient prospective payment system (IPPS) claims.

Table 6-14 shows both base year and year 3 rates of 90-day mortality for each intervention hospital and for the comparison hospitals as a group. During the base year, the rate of 90-day mortality was 8.59 percent for the group of comparison hospitals. At BIMC, the 90-day mortality rate was 8.59 percent in the base year. In the final year of the demonstration, the rate of 90-day mortality was 9.38 percent at the comparison hospitals and 8.71 percent at BIMC. The average percentage change between the base year and final implementation year was 1.4 percent at BIMC and a decline of 2.56 percent at the comparison hospitals.

Table 6-14
Base year and Year 3 rates of 90-day mortality for BIMC and its comparison group

Hospital	Base year deaths	Base year 90-day mortality rate	Year 3 deaths	Year 3 90-day mortality rate	Percentage change in mortality rate
Beth Israel Medical Center	766	8.59%	845	8.71%	1.40%
Comparison hospitals	6,425	9.26%	6,315	9.38%	2.56%

SOURCE: 2007–2011 Medicare inpatient prospective payment system (IPPS) claims.

Readmissions: Table 6-15 presents all cause readmissions for BIMC and for the group of comparison hospitals in the base year and third implementation year. Readmissions were counted if they occurred within 30 days and at least one day after discharge from the qualifying hospital stay, regardless of where the readmission occurred. The readmission rate at the comparison hospitals was 18.25 percent in the base year and declined to 17.19 percent in the third performance year, an overall decline of 5.81 percent. BIMC had a higher 30-day readmission rate in the base year, 19.03 percent, and declined to 17.96 in the third performance year, a decline of 5.62 percent.

Table 6-15
Base year and Year 3 rates of all cause 30-day readmissions for BIMC and its comparison group

Hospital	Base year counts	Base year readmission rate	Year 3 counts	Year 3 readmission rate	Percentage change in readmission rate
Beth Israel Medical Center	1,696	19.03%	1,743	17.96%	-5.62%
Comparison hospitals	12,186	18.25%	11,572	17.19%	-5.81%

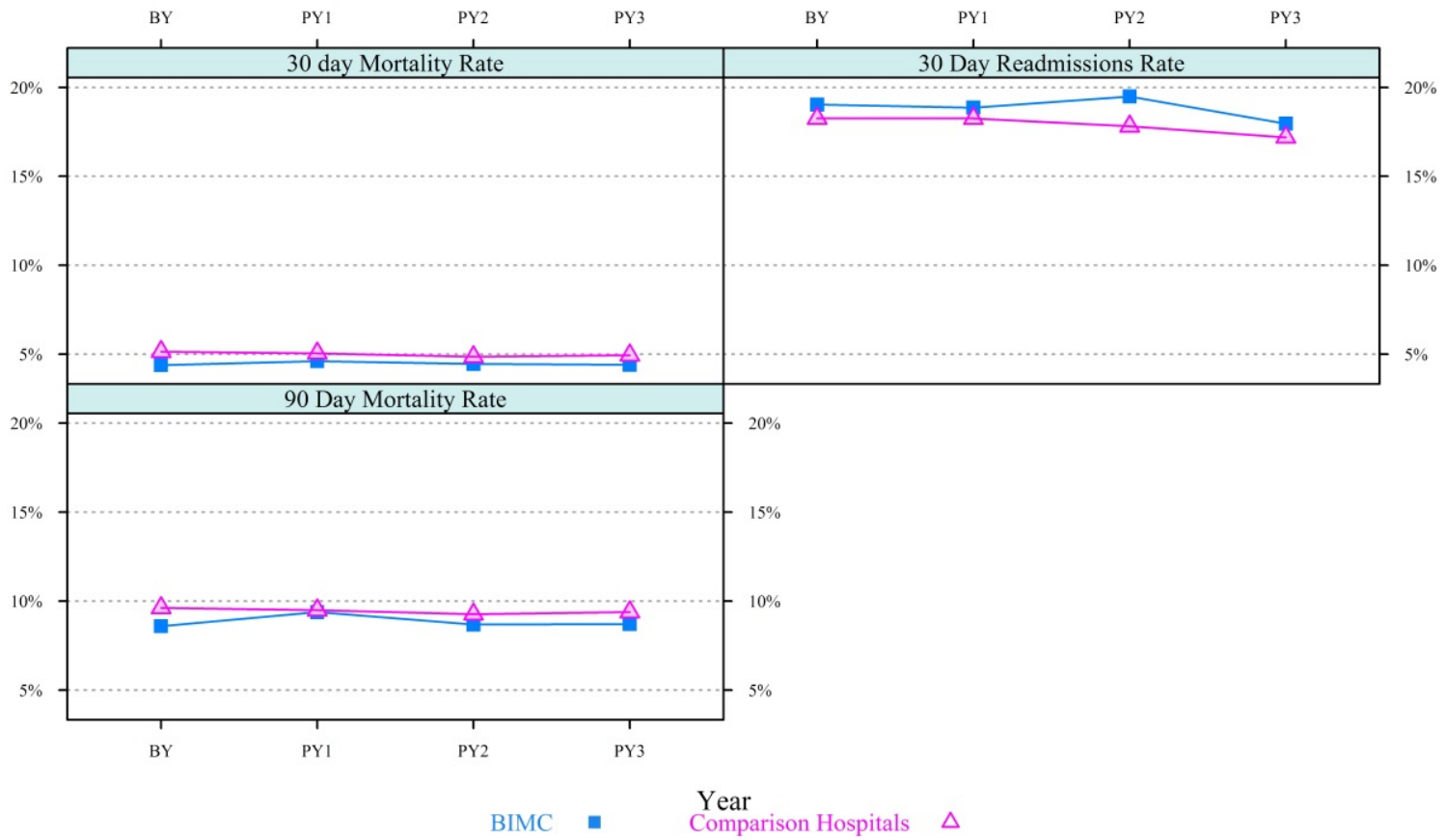
SOURCE: 2007–2011 Medicare inpatient prospective payment system (IPPS) claims.

Figure 6-2 shows a graphical presentation of the change from the base year, first, second and third performance years for 30-day mortality, 90-day mortality and 30-day readmissions. The year to year trend is relatively unchanged between BIMC and its controls.

Inpatient Quality Indicator Findings: Unlike CAMC, the BIMC intervention included a full range of DRG conditions and we present this more extensive analysis. ***Appendix Table 6-1*** presents the base year and year three rates generated by the AHRQ IQI software for each of the specific conditions below:

- Acute myocardial infarction (AMI)
- Congestive heart failure (CHF)

Figure 6-2
Change from the base year, first, second and third performance years for 30-day mortality, 90-day mortality and 30-day readmissions



- Stroke
- Gastrointestinal hemorrhage (GI hemorrhage)
- Hip fracture
- Pneumonia

Because the IQIs measure mortality rates among patients treated for only the six conditions above, the population at risk for any measure at a single hospital may be quite small. To address this in the 2D analysis of impact of the demonstration, we use a simple composite variable to measure whether a patient died while being treated for any one of the six conditions during the time period (base year or demonstration year 2). ***Table 6-16*** presents base year and year 3 mortality rates per 1,000 episodes across all six conditions for each intervention hospital and the comparison hospitals. "Population at risk" refers to any patient who meets all exclusion criteria and is treated for at least one of the conditions above. "Occurrences" refers to deaths of patients in the population at risk. Thus, the mortality rate is the observed mortality rate among patients treated for AMI, CHF, stroke, GI hemorrhage, hip fracture, or pneumonia during the measurement period. We then compare the percentage change in the year 3 rate from the base year rate. If the percentage is positive, the hospital had a higher mortality rate across the six conditions in the second year of the demonstration relative to the base year. It is important to note that year-to-year change in rates may appear large because the population size for each hospital is relatively small. Therefore these numbers should only be interpreted along with the 2D analyses presented below. Detailed information for the 6 components of the composite measure are presented in Appendix Table 6-1.

Table 6-16
Base year and year 3 mortality rates per 1,000 episodes for selected conditions: BIMC and its comparison group

Hospital	Base year population at risk	Base year occur-rences	Base year rate	Year 3 population at risk	Year 3 occur-rences	Year 3 rate	Percent change in mortality rate
Beth Israel Medical Center	1,507	80	53.09	1,368	94	68.71	23%
Comparison hospitals	11,178	701	62.71	10,154	636	62.64	0%

SOURCE: 2007-2011 Medicare IPPS claims.

BIMC had an increased mortality rate of 23 percent between the base year and year 3. The comparison group, on the other hand, performed nearly the same in both years (-0.10%). BIMC, however, started from a lower absolute rate (53.08) than the comparison hospitals (62.71), but the comparison hospitals had a lower rate of mortality after year 3.

Patient Safety Indicator Findings: As noted in section 6-1, the AHRQ PSIs are a set of measures providing rates of potentially preventable complications and other iatrogenic events that occur in the hospital setting. These are limited to cases in which a secondary diagnosis code

indicates a potentially preventable complication. For each complication, the observed rate is the actual number of occurrences per 10,000 patients. Detailed descriptions of the complications are presented for the base year and third intervention year in ***Appendix Table 6-2***.

Because the PSIs measure the rate of occurrence of adverse events, the number these events at a single hospital is likely to be quite small. To address this in the 2D analysis of the demonstration, we use a simple composite variable to measure whether a patient experienced at least one adverse event during an episode during the time period under consideration. ***Table 6-17*** presents base year and year 3 rates of adverse events per 10,000 episodes for each intervention hospital and the comparison hospitals. "Population at risk" refers to any patient who meets all exclusion criteria; almost all of the patients qualified to be in the denominator. "Occurrences" refers to an occurrence of an adverse event. We compare the percentage change in the year 3 rate from the base year rate. If the percentage is positive, the hospital had a higher rate of adverse events in the third year of the demonstration relative to the base year.

Table 6-17
Base year and year 3 rates of adverse events per 10,000 episodes: BIMC and its comparison group

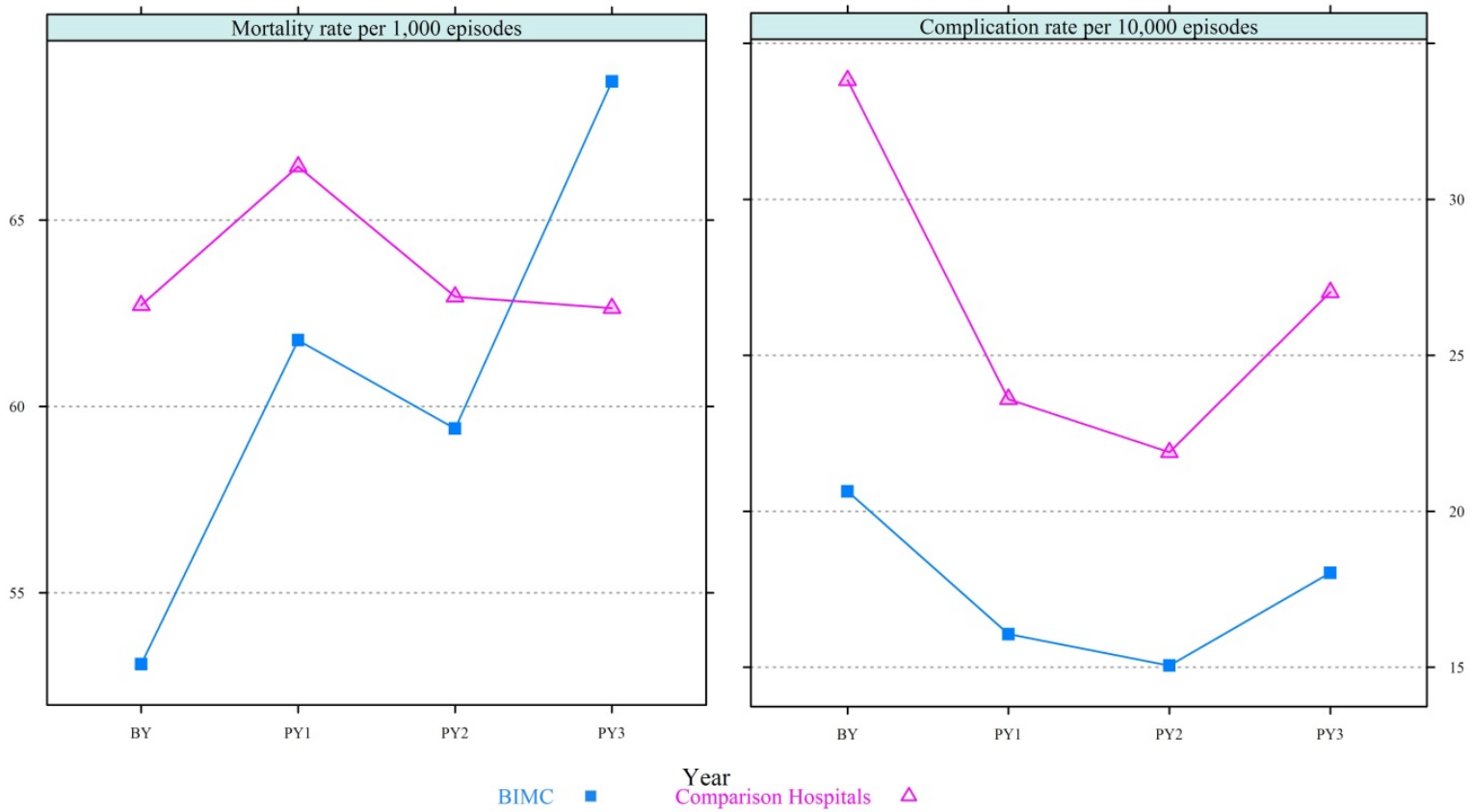
Hospital	Base year population at risk	Base year occur-rences	Base year rate	Year 3 population at risk	Year 3 occur-rences	Year 3 rate	Percent change in rate of adverse events
Beth Israel Medical Center	37,308	77	20.63901	39,930	72	18.03156	-13%
Comparison hospitals	295,117	998	33.8171	294,533	796	27.02583	-20%

SOURCE: 2007-2011 Medicare IPPS claims.

In most cases the change in rates of adverse events between the base year and year 3 seem large. It is important to note that year-to-year change in rates may appear large because the number of occurrences for each hospital is relatively small. Therefore these numbers are informational and should only be interpreted along with the 2D analyses presented below. BIMC experienced a decline of 13 percent in the rate of adverse events between the base year and the third performance year. The comparison hospitals also had a large drop, declining 20 percent between the base year and performance year 3. BIMC, however, had a much lower absolute rate in year 3 (18 per 10,000) versus the comparison hospitals (27 per 10,000).

Figure 6-3 shows a graphical presentation of the change from the base year, first, second and third performance years for the aggregated IQI and PSI rates. BIMC starts at a much lower rate in both and remains below the comparison hospitals in the PSIs, but exceeds the comparison hospitals in the IQIs in year 3.

Figure 6-3
Change from the base year, first, second and third performance years for the aggregated IQI and PSI rates



Medical Record Based Data Findings: The Inpatient Quality Reporting data measure adherence to process of care standards for three conditions and process measures for surgery from the surgical care improvement project (SCIP).

We used quarterly IQR data provided by CMS; annual data from 2008 through 2011 are presented below. To generate a rate for the comparison group, we summed the numerator and denominator for each measure across all four quarters and across each hospital. We then divided the numerator by the denominator to calculate the rate, which can be interpreted as the percentage of eligible patients across all of the comparison hospitals who received the intervention. Results for each topic are presented in ***Tables 6-18*** through ***6-21.***

Acute Myocardial Infarction (Heart Attack) Care: The AMI process of care measures for each intervention hospital and the comparison hospitals as a group can be found in ***Table 6-18***. In the third implementation year the composite score for BIMC was 97 percent as compared to the comparison hospitals at 98 percent. While the range across the measures was narrow, there was substantial variation in the percentage of patients receiving primary PCI within 90 minutes of arriving at the hospital. Only 72 percent of AMI patients treated at BIMC received this intervention in the specified amount of time while the comparison hospitals provided this 90percent of the time. BIMC, however, improved from 33 percent in the base year. The percentage of AMI patients receiving the remaining interventions is above 90 percent. There was some progress between the base year and the third performance year in smoking cessation counseling.

Heart Failure Care: Table 6-19 presents hospital process of care measures for treating patients with heart failure. The composite score for the third intervention year is 88 percent at BIMC and 95 percent at the control hospitals. The majority of scores for each intervention range between 90 and 100 percent although in some cases the percentage of patients receiving a particular intervention was significantly lower. Discharge instructions were only received by 73 percent of the heart failure patients at BIMC in the third intervention year which brought down their composite score. This was an improvement over the 60 percent of heart failure who received discharge instructions in the base year.

Pneumonia Care: The process of care measures for pneumonia patients are presented in ***Table 6-20***. In the intervention year, composite scores range from 85 to 99 percent. Other than the influenza vaccination, BIMC and the comparison hospitals provided the recommended interventions to more than 90 percent of patients. The composite scores of the comparison hospitals increased from 88 to 93 percent between the base year and the third performance year and BIMC increased from 91 to 94 percent over the same time period.

Surgical Care Improvement Project: The surgical care improvement project (SCIP) process of care measures for the intervention hospitals and the comparison group are presented in ***Table 6-21***. In year 3, the SCIP composite scores were 96 percent at BIMC and 97 percent at the comparison hospitals. There is very little variation in the SCIP measures outside of the 90^{th} percentile. A notable exception includes the percent of surgery patients on a beta-blocker prior to arrival who received a beta-blocker during the perioperative period. Only 78 percent of patients at BIMC received appropriate care in this domain.

Table 6-18
Year 3 hospital process of care measures: AMI, BIMC and its comparison group

	Base year				PY3			
	BIMC		Comparison group		BIMC		Comparison group	
	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
AMI care								
Aspirin at arrival	197	94%	2,435	96%	286	99%	2,826	98%
Aspirin prescribed at discharge	207	97%	2,756	97%	284	98%	2,927	99%
ACE inhibitor or ARB for LVSD	56	89%	791	89%	61	90%	644	95%
Adult smoking cessation advice/counseling	49	86%	735	92%	61	98%	692	100%
Beta blocker prescribed at discharge	221	96%	2,785	96%	276	98%	2,886	98%
Beta blocker at arrival	131	94%	1,394	94%	N/A	N/A	N/A	N/A
Fibrinolytic therapy received within 30 minutes of hospital arrival	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Primary PCI received within 90 minutes of hospital arrival	18	33%	237	73%	25	72%	321	90%
AMI composite score	879	93%	11,133	95%	993	97%	10,296	98%

NOTE: The composite score is calculated by summing the numerator and denominator for each measure in a topic and then dividing numerator by denominator to get the rate. ACE inhibitor = angiotensin-converting enzyme inhibitor; AMI = acute myocardial infarction; ARB = angiotensin receptor blocker; BIMC = Beth Israel Medical Center; LVSD = left ventricular systolic dysfunction; PCI = percutaneous coronary intervention; PY = performance year.

SOURCE: 2008–2011 Medicare Inpatient Quality Reporting data.

Table 6-19
Year 3 hospital process of care measures: heart failure, BIMC and its comparison hospitals

	Base year				PY3			
	BIMC		Comparison group		BIMC		Comparison group	
	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
HF care								
Discharge instructions	243	60%	4,445	70%	274	73%	3,588	91%
Evaluation of LVS function	295	97%	5,557	97%	322	98%	4,513	99%
ACE inhibitor or ARB for LVSD	109	90%	2,120	87%	105	93%	1,564	94%
Adult smoking cessation advice/counseling	39	92%	741	91%	39	100%	499	98%
HF composite score	686	83%	12,863	85%	740	88%	10,164	95%

NOTE: The composite score is calculated by summing the numerator and denominator for each measure in a topic and then dividing numerator by denominator to get the rate. ACE = angiotensin-converting enzyme; ARB = angiotensin receptor blocker; BIMC = Beth Israel Medical Center; HF = heart failure; LVSD = left ventricular systolic dysfunction; PY = performance year.

SOURCE: 2008–2011 Medicare Inpatient Quality Reporting data.

Table 6-20
Year 3 hospital process of care measures: pneumonia, BIMC and its comparison group

	2007				PY3			
	BIMC		Comparison group		BIMC		Comparison group	
	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Pneumonia care								
Oxygenation assessment	186	99%	3,166	100%	N/A	N/A	N/A	N/A
Pneumococcal vaccination	117	71%	2,508	80%	157	85%	2,214	89%
Blood cultures performed in the emergency department prior to initial antibiotic received in hospital	170	96%	2,585	84%	193	99%	2,447	95%
Adult smoking cessation advice/counseling	52	85%	619	93%	58	97%	463	97%
Initial antibiotic received within 6 hours of hospital arrival	79	89%	1,279	89%	193	95%	2,280	94%
Initial antibiotic selection for CAP in immunocompetent patient	107	94%	1,559	90%	113	94%	1,199	95%
Influenza vaccination	41	85%	826	90%	59	93%	752	86%
Pneumonia composite score	752	91%	12,542	88%	773	94%	9,355	93%

NOTE: The composite score is calculated by summing the numerator and denominator for each measure in a topic and then dividing numerator by denominator to get the rate. BIMC = Beth Israel Medical Center; CAP = community-acquired pneumonia; PY = performance year.

SOURCE: 2008–2011 Medicare Inpatient Quality Reporting data.

Table 6-21
Year 3 hospital process of care measures: surgical care improvement project, BIMC and its comparison group

	Base year				PY3			
	BIMC		Comparison group		BIMC		Comparison group	
	# patients	% receiving	# patients	% receiving	# patients	% receiving	# patients	% receiving
Surgical care improvement project								
Prophylactic antibiotic received within 1 hour prior to surgical incision	455	92%	6,269	91%	524	99%	5,800	97%
Prophylactic antibiotic selection for surgical patients	354	99%	5,041	95%	529	96%	5,881	97%
Prophylactic antibiotics discontinued within 24 hours after surgery end time	438	90%	5,932	83%	517	94%	5,566	94%
Surgery patients with recommended venous thromboembolism prophylaxis ordered	260	97%	5,002	93%	235	98%	3,771	98%
Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery	260	94%	5,002	89%	235	98%	3,761	96%
Cardiac surgery patients with controlled 6 a.m. postoperative blood glucose	37	95%	251	89%	155	91%	1,337	94%
Surgery patients with appropriate hair removal	166	99%	2,037	91%	737	99%	9,039	100%
Surgery patients on beta-blocker therapy prior to arrival who received a beta-blocker during the perioperative period	N/A	N/A	N/A	N/A	244	78%	2,849	92%
SCIP composite score	1,970	94%	29,534	90%	3,176	96%	38,004	97%

NOTE: The composite score is calculated by summing the numerator and denominator for each measure in a topic and then dividing numerator by denominator to get the rate. BIMC = Beth Israel Medical Center; PY = performance year; SCIP = Surgical Care Improvement Project.

SOURCE: 2008–2010 Medicare Inpatient Quality Reporting data.

Multivariate Claims Based Measures: We present the results of a 2D analysis of the following dependent variables for BIMC: 30-day mortality, 90-day mortality, 90-day readmissions, IQI numerator (mortality), and PSI numerator (event). Each of these is a binary variable equal to one if the patient met the criteria in question (i.e., if the patient died within 30 days of admission). As outlined previously in the methods section, we used logistic regressions to estimate the impact of the demonstration on each of these outcome measures. The sample size for each measure is indicated below.

We use two measures that are based on the AHRQ Quality Indicators measures presented above. The first, IQI numerator (mortality), is equal to one if the patient is eligible for inclusion in the numerator of at least one of the Inpatient Quality Indicators (IQI) measures described above. As described above, the IQIs are mortality rates for selected conditions: acute myocardial infarction (AMI), congestive heart failure (CHF), stroke, gastrointestinal (GI) hemorrhage, hip fracture, or pneumonia. The denominator consists of all patients treated for these conditions and who meet additional exclusion criteria imposed by the AHRQ methodology. A patient who died while being treated for any of these conditions will be counted in the numerator for the overall IQI measure. Aggregating to this level addresses the fact that the individual denominators for any single IQI measure may be too small.

Similarly, for the Patient Safety Indicators (PSIs) we calculate one overall measure to indicate whether a patient experienced any of the 13 preventable complications captured by the PSIs we calculate above. For example, the PSI numerator will be equal to one if a patient who meets all exclusion criteria develops a pressure ulcer while in the hospital. Aggregating to this level addresses the fact that the individual numerators for any single PSI measure may be small.

The focus of the Gainsharing Demonstration is to incentivize hospitals and physicians to collaborate in an effort to generate cost savings. Although the demonstration did not specifically incentivize improvements in quality of care outside of any improvements that might generate savings, any strategies employed by the hospital should not lead to declines in hospital quality. A decrease in hospital quality appears as a positive coefficient (i.e., each of the measures are negative events). We calculated the percent difference-in-differences by dividing the coefficient from the logit equation by the mean outcome (e.g., 30-day mortality) for the demonstration hospitals as a group in the base year. We test the significance of our estimates by constructing 95 percent confidence intervals using the standard errors calculated from a simple OLS regression model.

We found no statistically significant impact of the BIMC demonstration on any of the five measures except for the patient safety indicators (***Table 6-22***). The PSIs are significant at the 10 percent level. The percent change from the demonstration hospitals in the base year looks large because the PSIs are relatively rare events. In fact, BIMCs PSI rate is almost half of the rate of the comparison hospitals in absolute value. There was some amount of positive change in 30-day mortality rate against the BIMC demonstration hospitals. However, these changes between the pooled performance years and the base year are not significant and of very small magnitude. We also tested year specific 2D estimators and a few years were significant, but there was no noticeable pattern.

Table 6-22
Difference-in-differences (2D) estimates of quality and patient safety measures

Measure	Difference-in-differences logit	Percent difference-in-differences logit
30-day mortality (N = 303,295)	-0.00030	-0.69%
90-day mortality (N = 303,295)	0.00057	0.34%
30-day readmissions (N = 303,295)	0.00532	2.96%
Inpatient quality indicators numerator (mortality; N = 56,475)	0.00447	6.70%
Patient safety indicators numerator (event; N = 254,325)	0.01259*	69.45%

*** = significant at 1% level; ** = significant at 5% level, * = significant at 10% level

SOURCE: 2007-2011 Medicare inpatient prospective payment system (IPPS) claims.

Medical Record Based Measures: We present the results of a simple 2D analysis of the four IQR composite scores for BIMC in ***Table 6-23***. The differences presented in the table represent the difference across time and the difference between BIMC and the comparison group. Three of the four 2D estimators are negative, indicating a decline against the comparison hospitals. Because these estimates are based on a single observation for each hospital in each period, we cannot test the statistical significance of these estimates.

Table 6-23
Difference-in-differences (2D) estimates of medical record-based measures

Hospital	AMI composite score	HF composite score	Pneumonia composite score	SCIP composite score
Beth Israel Medical Center	1%	-3%	-2%	-4%

SOURCE: 2008-2011 Quarterly Inpatient Quality Reporting (IQR) data. AMI, acute myocardial infarction; HF, heart failure; SCIP, surgical care improvement project.

6.4 Discussion

Our results indicate small, and where we are able to test, statistically insignificant impacts of the Gainsharing Demonstration on the quality indicators measured above. This is likely the result of convergence in quality across hospitals given the emphasis on quality improvement over the past two decades. It does not appear that the gainsharing demonstration had any significantly negative effects on the quality of care that was received at either CAMC or BIMC over the period of the demonstration. That said, BIMC had an increased mortality rate of 23 percent between the base year and year 3; performance for the comparison group remained nearly constant. There were also some troubling trends in the descriptive findings (particularly

increases in 30- and 90-day mortality rates) that indicate the Gainsharing demonstration at BIMC did not result in improved quality of care.

This page left intentionally blank.

SECTION 7
PATIENT SATISFACTION SURVEY-BASED INDICATOR FINDINGS

Patient experience is an important component of quality of care. In this section we present patient satisfaction measures from beneficiary surveys. The measures presented represent findings for the base year (calendar year 2007) as well as one performance year for Charleston Area Medical Center (CAMC) and three performance years for Beth Israel Medical Center (BIMC). The specific time periods are October 1, 2008, through September 30, 2009 [PY1]; October 1, 2009, through September 30, 2010 [PY2]; and October 1, 2010, through September 30, 2011[PY3]. These analyses focus less on differences in the actual proportions of responses regarding patient satisfaction and more on divergent trends between demonstration and comparison group hospitals relative to the baseline findings. The gainsharing initiatives implemented under the Medicare Gainsharing Demonstration were intended to be transparent to patients. Evidence of decreased levels of patient satisfaction in demonstration hospitals relative to the comparison group during the intervention years may indicate that this goal was not accomplished.

7.1 Data Sources, Measures, and Methods

In addition to understanding how care is delivered in the hospital setting, gaining insight into the patient experience is crucial to seeing a complete picture of hospital quality. As consumers of health care, patients may have concerns in addition to those addressed by measures of outcome and process of care. Because of this, we analyzed patient experience measures on 10 topics.

Endorsed by the National Quality Forum in 2005, the HCAHPS survey was developed through a partnership between CMS and AHRQ. HCAHPS data have been collected since 2006 and were first publicly reported in 2008. In each month of the year, the survey is administered to a random sample of adult patients across medical conditions. Although the data are collected by vendors hired by reporting hospitals, CMS provides quality oversight (e.g., inspecting survey administration procedures, analyzing submitted data). Four methods are available to hospitals for collecting data; CMS adjusts for this variance when standardizing scores for comparison across hospitals.

Survey-based quality measures provide details about hospital quality that cannot be garnered from claims or medical records data, but these measures do have limitations. The data are drawn from a sample of patients that has the potential for bias. Although strict standards are upheld to ensure the quality of data and minimize the impact of bias, these methods may not succeed in eliminating all bias from the data.

The HCAHPS data are also not specific to particular service lines. Any relationship between the demonstration and these quality measures is incomplete at best, and the results must be analyzed with that caveat.

The HCAHPS survey contains 27 questions that result in 10 survey-based quality indicator measures: 6 summary measures, 2 individual measures, and 2 global measures. The categories are as follows:

- summary measures
  - communication with nurses
  - communication with doctors
  - responsiveness of hospital staff
  - pain management
  - communication about medication
  - discharge information
- individual measures
  - cleanliness of hospital environment
  - quietness of hospital environment
- global measures
  - overall rating of hospital
  - willingness to recommend hospital

With the exception of the discharge information measure, which requires only two (yes or no) questions, each measure uses at least three questions to develop the rating. We used HCAHPS data downloaded from the Hospital Compare Web site. Reporting of HCAHPS data is voluntary, although the Medicare Gainsharing Demonstration sites were required, as a condition of participation, to report HCAHPS data.

We also conducted a simple 2D analysis of the 10 HCAHPS measures for CAMC and BIMC. The differences presented in the table represent the difference across time and the difference between each hospital and the comparison group. For each of the 9 measures that consist of a 3-part question, we used the percentage of patients that chose the two most positive answers. For example, "Communication with nurses" refers to the percentage of patients who answered that nurses always or usually communicated well.

7.2 Results

Charleston Area Medical Center: The HCAHPS results for CAMC and the group of comparison hospitals can be found in ***Table 7-1***.

Table 7-1
HCAHPS results: CAMC and comparison hospitals

	Base year, CAMC	Base year, comparison group	Demo, CAMC	Demo, comparison group	Difference, CAMC	Difference, comparison group
Communication with nurses*						
Nurses always communicated well	73%	78%	72%	77%	−1%	−2%
Nurses usually communicated well	20%	17%	22%	19%	2%	2%
Nurses sometimes or never communicated well	7%	5%	6%	5%	−1%	0%
Communication with doctors*						
Doctors always communicated well	80%	83%	79%	79%	−1%	−4%
Doctors usually communicated well	15%	14%	16%	17%	1%	3%
Doctors sometimes or never communicated well	5%	4%	5%	4%	0%	0%
Responsiveness of hospital staff						
Patients always received help as soon as they wanted	60%	63%	57%	62%	−3%	−1%
Patients usually received help as soon as they wanted	27%	27%	27%	27%	0%	−1%
Patients sometimes or never received help as soon as they wanted	13%	10%	16%	12%	3%	2%
Pain management						
Pain was always well controlled	68%	70%	65%	69%	−3%	−1%
Pain was usually well controlled	24%	23%	26%	24%	2%	1%
Pain was sometimes or never well controlled	8%	7%	9%	7%	1%	0%
Communication about medicines						
Staff always explained	60%	62%	53%	60%	−7%	−2%
Staff usually explained	17%	16%	21%	18%	4%	2%
Staff sometimes or never explained	23%	22%	26%	22%	3%	0%
Discharge information						
Yes, staff did give patients this information	80%	83%	79%	83%	−1%	0%
No, staff did not give patients this information	20%	17%	21%	17%	1%	0%
Cleanliness of hospital environment						
Room was always clean	65%	67%	61%	65%	−4%	−2%
Room was usually clean	25%	21%	26%	22%	1%	1%
Room was sometimes or never clean	10%	12%	13%	13%	3%	1%

(continued)

Table 7-1 (continued)
HCAHPS results: CAMC and comparison hospitals

	Base year, CAMC	Base year, comparison group	Demo, CAMC	Demo, comparison group	Difference, CAMC	Difference, comparison group
Quietness of hospital environment						
Always quiet at night	46%	57%	48%	56%	2%	−1%
Usually quiet at night	36%	30%	33%	32%	−3%	2%
Sometimes or never quiet at night	18%	13%	19%	13%	1%	0%
Overall rating of hospital*						
Patients who gave a rating of 9 or 10 (high)	66%	68%	66%	67%	0%	−1%
Patients who gave a rating of 7 or 8 (medium)	24%	24%	24%	24%	0%	0%
Patients who gave a rating of 6 or lower (low)	10%	8%	10%	9%	0%	1%
Willingness to recommend hospital						
Yes, patients would definitely recommend the hospital	74%	78%	74%	75%	0%	−3%
Yes, patients would probably recommend the hospital	21%	17%	20%	21%	−1%	4%
No, patients would not recommend the hospital (they probably would not or definitely would not recommend it)	5%	4%	6%	4%	1%	0%

NOTES: The base year 2008 comparison group data are based on a simple average of data from 7 of the 10 comparison hospitals, except as noted. The 2009 ("demo") comparison group data are based on all 10 comparison hospitals. "Difference" is demo – base year; discrepancies are due to rounding.

* Indicates measure is based on data from 6 of the 10 comparison hospitals.

SOURCE: 2008 and 2009 Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

Patients were asked how often nurses communicated well. Those patients treated at CAMC reported a slight improvement (a 1-percentage-point increase, from 93 to 94 percent, reporting that nurses always or usually communicated well). Patients treated at the group of comparison hospitals also reported a 1-percentage-point increase in nurses always or usually communicating well, from 95 to 96 percent. In contrast, patient responses to the question about how well doctors communicated with patients indicated a slight decline for the comparison group (1 percentage point, decreasing from 97 to 96 percent of doctors always or usually communicating well) and no change at CAMC (95 percent).

Patient respondents were asked how often they received help quickly from hospital staff. Of patients surveyed after receiving care at CAMC, 84 percent reported always or usually receiving help as soon as they wanted (a decrease of 3 percentage points from 87 percent in the base year). Patients treated at the comparison hospitals also reported a slight decrease in the

responsiveness of hospital staff, from 90 to 89 percent. The HCAHPS survey also asked patients about how well their pain was controlled during their hospital stay. Pain was always or usually well controlled for 91 percent of patients treated at CAMC during PY1 (down from 92 percent in the base year); this measure did not change for patients treated at the comparison hospitals.

When asked about hospital staff providing explanation about medication before giving it to patients, 74 percent of patients treated at CAMC during PY1 reported that staff always or usually explained about medicine (a 3-percentage-point decrease from 77 percent in the base year). Patients treated at comparison hospitals reported no change from base year to PY1 (in both years, 78 percent reported that staff always or usually explained medication).

Patients were also asked whether they were given information about what was required to continue their recovery at home. Although there was no change among patients treated at comparison hospitals (in both years, 83 percent indicated that they always or usually received information before discharge), the share of patients who always or usually received discharge information at CAMC decreased modestly (from 80 to 79 percent). Patient responses to a question about the cleanliness of patient rooms and bathrooms indicated declines at both CAMC (from 90 to 87 percent) and the comparison hospitals (88 to 87 percent).

Patients were asked how often the area around their room was kept quiet at night. The hospital was usually or always quiet for 82 percent of patients treated at CAMC during the base year and for 81 percent treated during PY1. Slightly more patients treated at the comparison hospitals reported that the hospital was always or usually quiet at night (87 percent in the base year and 88 percent in PY1). In addition, hospital patients surveyed were asked to rate the facility using a 10-point scale. In both the base year and PY1, 90 percent of patients treated at CAMC gave the hospital a medium (7 or 8) or high (9 or 10) rating, whereas 90 percent of patients rated the comparison hospitals medium or high in PY1, compared with 91 percent in the base year.

Finally, we present results of a question about patients' willingness to recommend the hospital to friends and family. Among patients treated at CAMC, those who would definitely or probably recommend the hospital decreased 1 percentage point, from 95 to 94 percent. Patients treated at the comparison hospitals reported that more (96 percent) would probably or definitely recommend the hospital in PY1; the rate was 95 percent in the base year.

For the majority of measures, in both the base year and PY1, the comparison hospitals have a better overall rating than CAMC. In the base year the only measure for which CAMC had a better rating was cleanliness of the hospital environment; CAMC's rooms scored "usually or always clean" 2 percentage points higher than the comparison group, but in PY1 both CAMC and the comparison groups' rooms were given the same score for cleanliness. In PY1, CAMC did not receive any higher scores than the comparison group. Patients were equally as likely to definitely or probably recommend the hospital in the base year but were 2 percentage points more likely to recommend the comparison hospitals in PY1.

The 2D analysis presented in ***Table 7-2*** shows that between the base year and the performance year and across CAMC and its comparison hospitals, CAMC does better than the comparison hospitals in only two measures (communication with doctors and overall rating of

the hospital). In both cases CAMC was rated higher by 1 percent. In one measure, communication with nurses, there was no difference between CAMC and the comparison. In all other measures the comparison group did better than CAMC.

Table 7-2
Charleston Area Medical Center difference-in-difference estimates of HCAHPS measures

Measure	Performance year 1
Communication with nurses	0%
Communication with doctors	1%
Responsiveness of staff	−2%
Pain management	−1%
Communication about medication	−3%
Discharge information	−1%
Cleanliness	−2%
Quietness	−2%
Overall rating	1%
Willingness to recommend	−2%

SOURCE: 2008 and 2009 Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

Beth Israel Medical Center: Because we have 3 years of performance data for this site, we report a wider range of analyses. ***Figure 7-1*** and ***Figure 7-2*** illustrate the trends in the HCAPHS measures from the base year to the third year after implementation. The data presented is from the September 2008, 2009, 2010, and 2011 releases. Ratings for the comparison group were calculated as a simple average across the comparison hospitals.

Figures 1 and 2 illustrate the changes in positive responses to each HCAHPS question from the base year through the demo period. For the purpose of graphing we considered a positive response as either always or usually. In many cases (8 measures), though not all, patient satisfaction at BIMC improved from the base year to performance year 3. In the group of comparison hospitals only half of the measures showed that patient satisfaction improved from the base year to performance year 3. The graphs show that at BIMC for the majority of measures the largest increase in positive responses was between the first and second performance years and in half the measures there is a decrease in positive response between the second and third performance years. At the comparison hospitals, eight measures saw a decrease in positive responses between the base year and the first performance year. After performance year one all but one measure at the comparison hospitals either increased slightly or remained the same each year.

Figure 7-1
BIMC versus comparison hospitals for 6 HCAPHSH measures

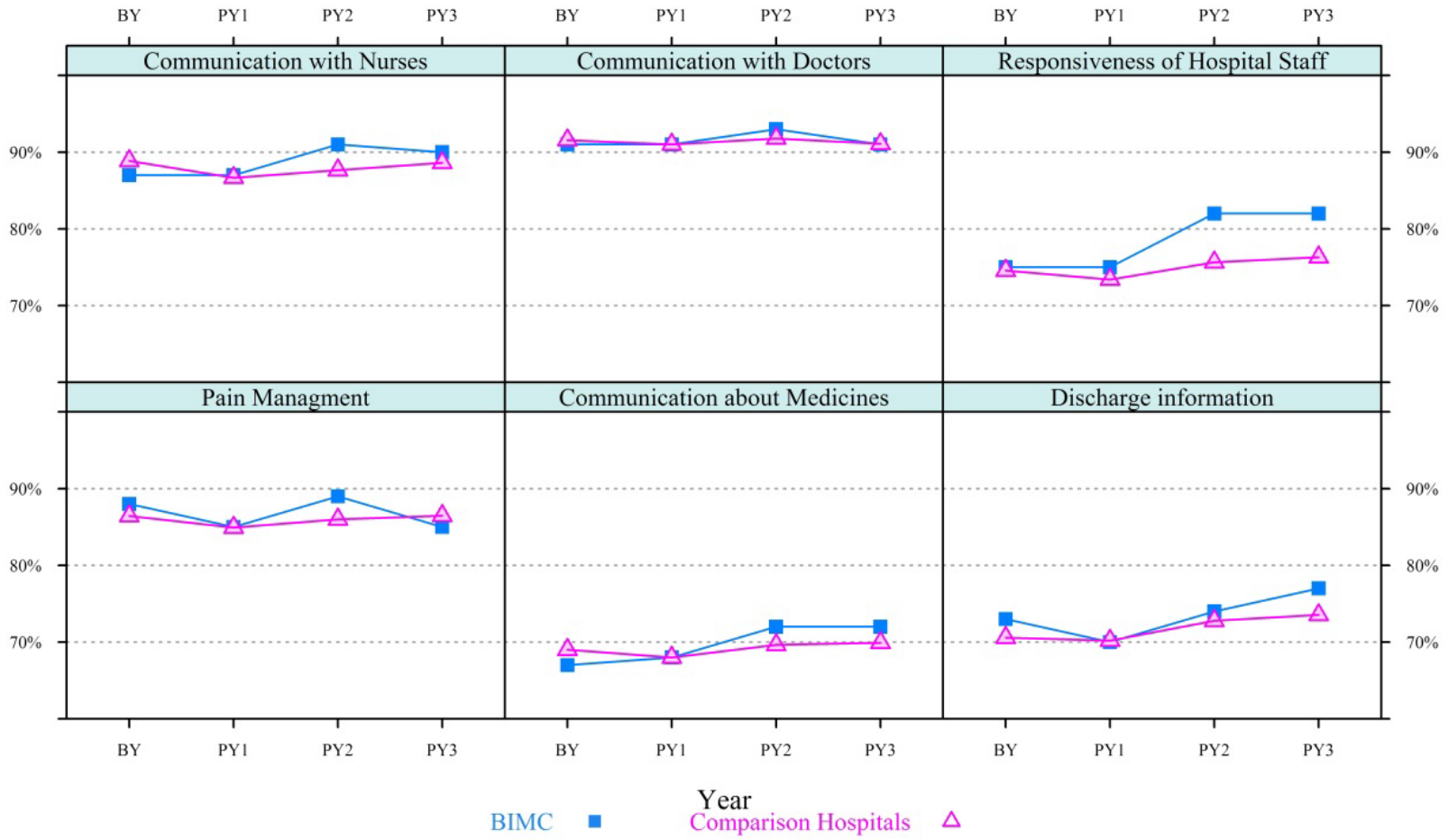
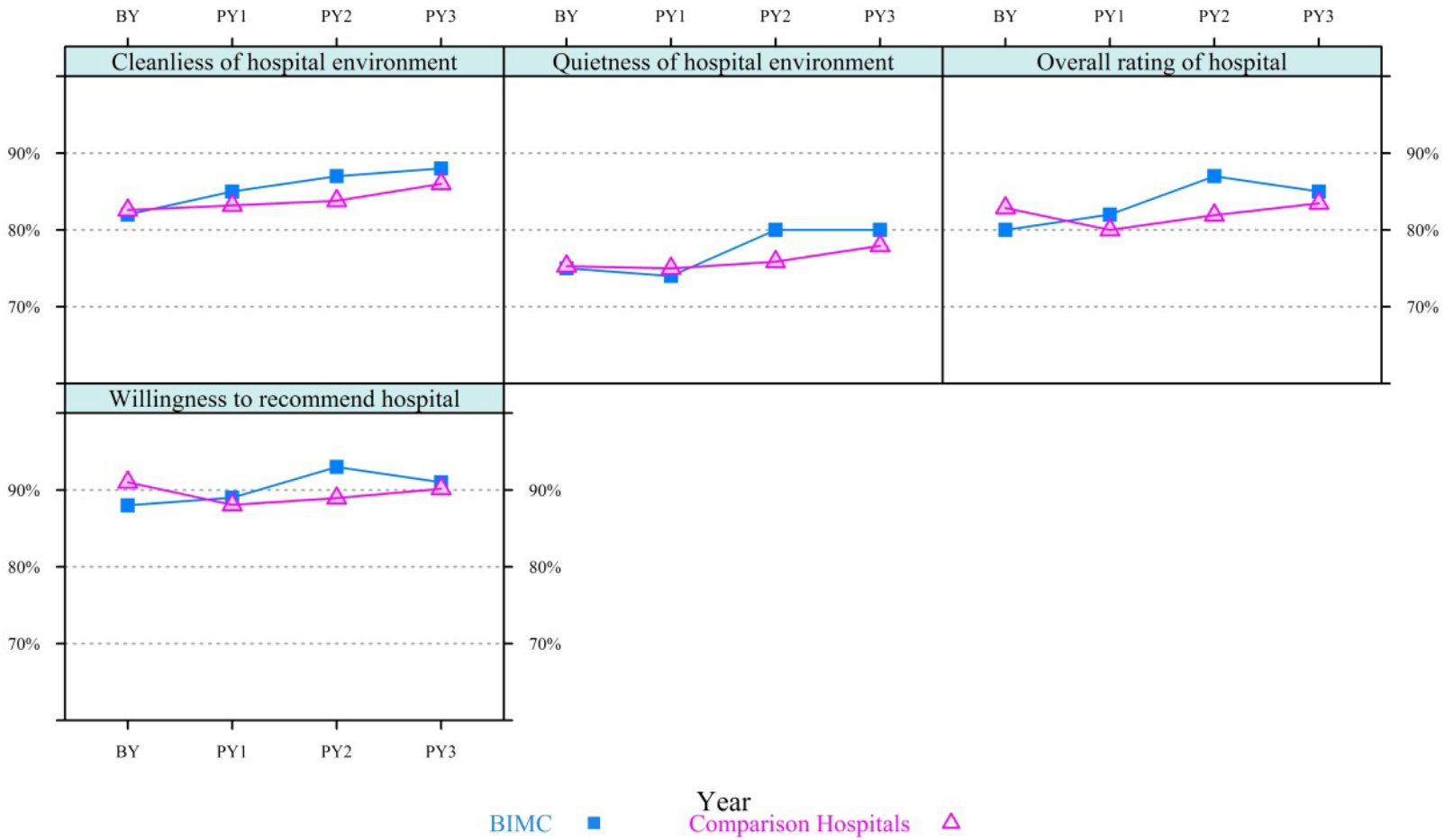


Figure 7-2
BIMC versus comparison hospitals for 4 HCAPHSH measures



We show additional detail for these BIMC HCAHPS results in ***Table 7-3.*** When asked how often nurses communicated well, patients treated at BIMC reported a slight improvement from the base year (87 percent) to performance year 3 (90 percent reporting that nurses always or usually communicated well). There was a slight increase in patients treated at the group of comparison hospitals who indicated an improvement in the third performance year (88 percent) compared with the base year (89 percent). Patient responses to the question about how well doctors communicated with patients indicated remained the same for both the comparison group (91 percent of doctors always or usually communicating well) and at BIMC (91).

In response to a question about how often they received help quickly from hospital staff, 75 percent of patients surveyed after receiving care at BIMC reported always or usually receiving help as soon as they wanted during the base year (with an increase to 82 percent during performance year 3). There was a slight change in how patients treated at the comparison hospitals answered this question (74 to 76 percent). When asked about pain management, 85 percent of BIMC patients indicated that pain was always or usually well controlled during the performance year (down from 88 percent in the base year); the comparison hospitals experienced a slight decrease in the number of patients responding favorably (87 percent to 86 percent).

Patients were asked how often hospital staff provided explanation about medication before giving it to patients. Of patients treated at BIMC during performance year three, 72 percent reported that staff always or usually explained about medicine (up from 67 percent in the base year). Patients treated at comparison hospitals reported 69 percent in both the base year and during the third year of the demonstration.

Patients responded to a question of whether they were given information about what was required to continue their recovery at home. At BIMC there was an increase in the number of patients reporting that they had received information about discharge (from 73 to 77 percent). There was a similar increase at the comparison hospitals (from 71 to 74 percent). Patient responses to a question about the cleanliness of patient rooms and bathrooms indicated improvement at BIMC (from 82 percent to 88 percent) and an increase at the comparison hospitals as well (from 82 to 86 percent).

When asked how often the area around their room was kept quiet at night, 75 percent of patients at BIMC responded that hospital was usually or always quiet during the base year; 80 percent reported this way during performance year three. Among patients treated at the comparison hospitals, 75 percent always or usually found the hospital to be quiet in the base period and 77 percent during performance year 3. Patients surveyed were asked to rate the facility using a 10-point scale. In the base year 80 percent of patients treated at BIMC gave the hospital a medium (7 or 8) or high (9 or 10) rating; 85 percent of patients gave the same ratings in the performance years. Among patients treated at the comparison hospitals, 83 percent gave the hospital a medium or high rating during performance year three (up from 82 percent in the base year).

Table 7-3
HCAHPS results: BIMC versus comparison hospitals

	Base year		PY3		Difference (PY3 – Base Year)	
	BIMC	Comparison group	BIMC	Comparison group	BIMC	Comparison group
Communication with nurses						
Nurses always communicated well	57%	63%	64%	65%	7%	1%
Nurses usually communicated well	30%	25%	26%	24%	-4%	-2%
Nurses sometimes or never communicated well	13%	11%	10%	11%	-3%	0%
Communication with doctors						
Doctors always communicated well	70%	71%	71%	71%	1%	-1%
Doctors usually communicated well	21%	20%	20%	20%	-1%	0%
Doctors sometimes or never communicated well	9%	8%	9%	9%	0%	0%
Responsiveness of hospital staff						
Patients always received help as soon as they wanted	43%	46%	55%	48%	12%	2%
Patients usually received help as soon as they wanted	32%	28%	27%	28%	-5%	0%
Patients sometimes or never received help as soon as they wanted	25%	25%	18%	24%	-7%	-2%
Pain management						
Pain was always well controlled	58%	58%	58%	58%	0%	1%
Pain was usually well controlled	30%	29%	27%	28%	-3%	-1%
Pain was sometimes or never well controlled	12%	14%	15%	14%	3%	0%
Communication about medicines						
Staff always explained	49%	49%	51%	50%	2%	1%
Staff usually explained	18%	20%	21%	19%	3%	0%
Staff sometimes or never explained	33%	31%	28%	30%	-5%	-1%
Discharge information						
Yes, staff did give patients this information	73%	71%	77%	74%	4%	3%
No, staff did not give patients this information	27%	29%	23%	26%	-4%	-3%

(continued)

Table 7-3 (continued)
HCAHPS results: BIMC versus comparison hospitals

	Base year		PY3		Difference (PY3 – Base Year)	
	BIMC	Comparison group	BIMC	Comparison group	BIMC	Comparison group
Cleanliness of hospital environment						
Room was always clean	53%	54%	65%	60%	12%	6%
Room was usually clean	29%	28%	23%	26%	-6%	-2%
Room was sometimes or never clean	18%	17%	12%	14%	-6%	-3%
Quietness of hospital environment						
Always quiet at night	42%	42%	49%	45%	7%	3%
Usually quiet at night	33%	33%	31%	32%	-2%	-1%
Sometimes or never quiet at night	25%	25%	20%	22%	-5%	-3%
Overall rating of hospital						
Patients who gave a rating of 9 or 10 (high)	43%	49%	55%	52%	12%	3%
Patients who gave a rating of 7 or 8 (medium)	37%	33%	30%	31%	-7%	-2%
Patients who gave a rating of 6 or lower (low)	20%	17%	15%	17%	-5%	-1%
Willingness to recommend hospital						
YES, patients would definitely recommend the hospital	52%	57%	63%	58%	11%	1%
YES, patients would probably recommend the hospital	36%	34%	28%	32%	-8%	-2%
NO, patients would not recommend the hospital (they probably would not or definitely would not recommend it)	12%	9%	9%	10%	-3%	1%

NOTES: The base year comparison group data were based on a simple average of data from 7 of the 16 comparison hospitals; and the PY3 comparison group data, on 13 of the 16. The "Difference" column is the PY3 figure minus the base year figure; discrepancies are due to rounding.

SOURCE: 2008 and 2011 Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

When asked about their willingness to recommend the hospital to friends and family, 88 percent of patients treated at BIMC would definitely or probably recommend the hospital, compared with 91 percent during the performance year. Patients treated at the comparison hospitals reported that more (91 percent) would probably or definitely recommend the hospital in the base year than in the performance year (90 percent).

We present the results of a simple 2D analysis of the 10 HCAHPS measures in ***Table 7-4***. The differences presented in the table represent the difference across time and the difference between each hospital and the comparison group. For each of the 9 measures that consist of a 3-part question, we used the percentage of patients that chose the two most positive answers. For example, "Communication with nurses" refers to the percentage of patients who answered that nurses always or usually communicated well.

The 2D analysis shows that BIMC did better than the comparison hospitals in the third performance years in 8 of the 10 measures: communication with nurses, responsiveness of hospital staff, communication about medicines, discharge information, cleanliness of hospital environment, quietness of the hospital environment, overall rating of the hospital, and willingness to recommend the hospital (by 3, 5, 4, 1, 3, 2, 4, and 4 percent, respectively). There was no difference in communication with doctors, and BIMC performed 3 percent lower in pain management to patients.

With few exceptions, the double differences appear quite small (generally +/−4 percent). BIMC had larger differences in responsiveness of hospital staff. Because these estimates are based on a single observation for each hospital in each period, we cannot test the statistical significance of these estimates.

Table 7-4
BIMC difference in difference (2D) estimates of HCAHPS measures

Measure	PY3
Communication with nurses	3%
Communication with doctors	0%
Responsiveness of staff	5%
Pain management	-3%
Communication about medication	4%
Discharge information	1%
Cleanliness	3%
Quietness	2%
Overall rating	4%
Willingness to recommend	4%

SOURCE: 2008 and 2011 Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS).

7.3 Discussion

The data presented above represent baseline analyses of HCAHPS beneficiary satisfaction survey findings for CAMC, BIMC, and their comparison sites. On the basis of these results, we were not able to detect any major impacts of the Gainsharing demonstration on patient satisfaction. Almost all the patient satisfaction measures are trending in the direction of improvement. This may be cautiously considered a positive finding. However, it is difficult to assess the significance of these changes in the absence of statistical testing (which is not possible given too few observations).

This page left intentionally blank.

SECTION 8
PHYSICIAN REFERRAL PATTERNS

One potential policy concern regarding gainsharing models is that participating physicians may refer more costly patients, including those more likely to have medical complications, to non-demonstration hospitals and treat the less complex patients at gainsharing-participating hospitals. Less complex patients are easier to manage, so physicians could be less likely to treat difficult cases that would negatively affect their likelihood of receiving performance payments. Participating physicians who work in relatively competitive markets (such as BIMC) and have admitting privileges at other hospitals have more discretionary ability to selectively direct patients in response to gainsharing incentives; physicians practicing in more isolated hospital markets (such as CAMC) are less able to engage in this behavior.[19] Participating physicians with admitting privileges at multiple local acute care hospitals may also have an incentive either to transfer very costly and difficult-to-manage cases from demonstration hospitals to other non-demonstration acute care hospitals (IPPS transfers) or to discharge them to PAC providers.[20] Patients treated in a demonstration hospital's ER are expected to increase the likelihood of admission to that hospital by a participating physician because the critical status of most ER patients reduces physician discretion in the decision to admit the patient. An increase in either admitting fewer high-cost patients or transferring severe, difficult-to-manage cases may manifest itself in a reduction in unprofitable cases (especially IPPS outlier cases). This is an example of how hospital and physician incentives can align in an undesirable manner.

In the rest of this section, we present descriptive and 2D regression analyses addressing these issues. The first analysis presents a market analysis, using descriptive statistics to look at the initial distribution of patients at participating demonstration and comparison hospitals, ER admissions, outliers, and severity of admissions. Next, we look at physician referral patterns and any impact of the demonstration on those patterns. We begin our physician referral analysis by looking at the number of participating physicians with admissions at other acute care hospitals, including comparison hospitals. We then look at transfer rates for physicians at participating hospitals and the average severity of those transfers.

8.1 Data and Measures

The primary sources of data for this section are the base and performance years' inpatient claims associated with episodes of care at participating and comparison hospitals. The inpatient

---

19 Research by Cromwell and Adamache (2004a 2006) in New Jersey and Oklahoma City for CMS showed remarkably few physicians who actively admitted to two or more local hospitals.

20 Favorable patient selection among physician owners of specialty hospitals suggests similar financial incentives for physicians sharing in increased hospital profits from lower costs in not-for-profit demonstration hospitals.

data were then run through 3M APR-DRG Limited License Grouper to generate a severity score for each inpatient admission.[21]

Next, we attempted to assign each inpatient admission to a responsible physician. For surgical DRGs, the responsible physician was the operating unique physician identification number (UPIN) or National Provider Identifier (NPI) if the UPIN or NPI was valid. If the operating UPIN or NPI was not valid, then we looked at the other UPIN or NPI. If the other UPIN or NPI was not valid, then we looked at the attending UPINs and NPIs. If none of these were valid, no responsible physician was assigned to the claim. For medical DRGs, the responsible physician was identified by the attending physician's UPIN or NPI. If the attending physician's UPIN or NPI was not valid, then no responsible physician was assigned to the claim.

In the second step, we created a unique physician ID linking all the UPINs and NPIs associated with one physician (and thus linking all of that physician's admissions). This was necessary because our initial analysis of the base year data showed that hospitals were still coding UPINs rather than NPIs in more than 30 percent of the admissions. To create the unique physician ID, we first extracted from the ARC inpatient file all valid physician UPINs and NPIs. We then matched the UPINs and NPIs against the 2010 NPPES database to link the associated UPINs and NPIs. In most cases, one UPIN was associated with no more than two NPIs.[22] In these instances, if the legal name of the provider matched, then we assigned both UPINs to the same physician ID. If the legal name of the provider did not match, we dropped those UPINs. The problems in creating unique IDs during the base year were not found during the performance periods, as the transition from UPINs to NPIs had been completed.

8.2 Methods

Given the incentives created under gainsharing, we were interested in determining whether demonstration hospitals saw a decrease in admissions classified as major or extreme severity between the base year and the demonstration period. We hypothesized that gainsharing participating physicians may have had an incentive to steer more severe cases to other hospitals where they had privileges in order to increase their demonstration performance payments. We tested this hypothesis empirically using a multivariate 2D analysis to model the probability that an admission was of major or extreme severity using a logistic model following the general form specified in Equation 5-1. We then simulated the average impact of the demonstration on the likelihood of a major or extreme severity admission using the approach discussed in ***Section 6***. In addition to the age, gender, race, and DRG weight explanatory variables and the T, D, and the T*D interaction terms specified in Equation 5-1, the new explanatory variables used in the logit regressions are (1) whether the beneficiary is admitted through the emergency room (ER) and (2)

---

[21] APR-DRGs are an enhanced extension of the basic DRG concept developed by 3M's Clinical Research Group, the National Association of Children's Hospitals and Research Institutes, and several physician groups. The fundamental principle of APR-DRGs is that the severity of illness and risk of mortality are both dependent on the patient's underlying condition. High severity of illness and risk of mortality are characterized by multiple serious diseases and the interactions between the disorders.

[22] We found 48 cases in which one NPI was associated with two UPINs and 1 instance in which one UPIN was associated with four NPIs.

whether the beneficiary is transferred in from a skilled nursing facility (SNF). As in Sections 5 and 6, we also tested for year-specific 2D effects.

To test the hypothesis that some gainsharing participating physicians might have an increased incentive to transfer sicker patients to other hospitals where they have privileges and thus increase their demonstration payments, we also modeled the probability of an IPPS transfer from a participating hospital and comparison hospital using a logistic model. The dependent variable is a binary variable where 0 denotes no IPPS transfer and 1 denotes an IPPS transfer. Results for both multivariate analyses are presented in ***Section 8.4***.

8.3 Descriptive Results

Market Analysis—Emergency Rooms: ERs are an important source of inpatient and outpatient referrals. A patient who visits a hospital's ER has a higher likelihood of being admitted to that hospital because the critical nature of ER cases reduces physician discretion in the decision to admit the patient. Consequently, to the extent that a hospital operates a higher volume ER, a larger proportion of patients may be admitted through an ER rather than a physician referral. However, ERs tend to attract higher acuity patients at admission which could lead to higher costs. In BIMC's New York metropolitan market, where patients may have the option of using any of several ERs, many factors may play into which ER they visit. If a patient has a highly acute condition, such as a burn, heart attack, or stroke, the patient (or the proxy) may choose the closest ER or may be directed to a specific ER by the protocol of the ambulance service. In smaller markets, such as the CAMC marketplace, access to alternative ERs is limited, particularly for the cardiac-related events that were the focus of the CAMC intervention.

Table 8-1 shows the percentage of base year and performance year admissions from the ER. BIMC admitted 77.1 percent of its patients through its ER; this figure increased to 81.3 percent during the performance year 3 (PY3). BIMC's comparison hospitals in the same market admitted a smaller proportion of patients through their ERs: 53.3 percent in the base year and 63.9 percent during PY3. CAMC admitted 34.0 percent of patients through its ER during the base year and increased this figure to 36.4 percent by the end of its performance year. The CAMC comparison hospitals (which are not located in the same geographic area) admitted 23.8 percent of their patients through their ERs in the base year, increasing to 28 percent during the performance year. Since both BIMC and CAMC both had substantially higher admission rates through the ER than their comparison hospitals during the base year, it is not surprising the comparison hospitals had larger increases.

Market Analysis—Cost Outliers: Under the demonstration, the participating hospitals could theoretically achieve cost savings by admitting fewer potentially high-cost patients. ***Table 8-2*** shows the share of base year and performance year admissions that were Medicare IPPS cost outliers for participating and comparison hospitals. Overall, the percentage of discharges that are outliers was low in the base year, ranging from 0.74 percent for BIMC to 4.92 percent for the CAMC comparison sites. We note, however, some marked increases in share of admissions with outlier payments between the base and performance years. This trend was most apparent for CAMC and the CAMC comparison sites, where the share of admissions with outlier payments nearly doubled to 8.08 percent (CAMC) and 9.15 percent (CAMC comparison hospitals). For

Table 8-1
Percentage of admissions through the emergency room at demonstration and comparison hospitals, in base year and performance period

Demonstration site or comparison hospitals	Base year	Performance year
Beth Israel Medical Center (BIMC)	77.1	81.3
BIMC comparison hospitals	53.3	63.9
Charleston Area Medical Center (CAMC)	34.0	36.4
CAMC comparison hospitals	23.8	28.0

NOTES: For both BIMC and CAMC, the base year was calendar year 2007. BIMC's performance year was PY3 (October 1, 2010, through September 30, 2011). CAMC's performance "year" was December 1, 2008, through December 31, 2009.

SOURCE: RTI analysis of Medicare claims.

BIMC and its comparison hospitals, admissions with outlier payments also increased but remained at relatively low levels.

Market Analysis—Severity of Admission: One reason that some hospitals may have more cost outliers may relate to a higher severity of admission. To measure patient severity, we applied the classification system developed by 3M as part of its APR-DRG grouping of patients. The goal of the APR-DRG was to create a more refined measure of patient severity than existed in the initial DRG system through more accurately identifying the severity of certain medical complications and how those complications might interact to increase a patient's overall severity level.

Table 8-2
Share of base year and performance period admissions with outlier payments at demonstration and comparison hospitals

Demonstration site or comparison hospitals	Base year	Performance year
Beth Israel Medical Center (BIMC)	0.74	2.07
BIMC comparison hospitals	1.67	2.57
Charleston Area Medical Center (CAMC)	4.28	8.08
CAMC comparison hospitals	4.92	9.15

NOTES: For both BIMC and CAMC, the base year was calendar year 2007 BIMC's performance year was PY3 (October 1, 2010, through September 30, 2011). CAMC's performance "year" was December 1, 2008, through December 31, 2009.

SOURCE: RTI analysis of Medicare claims.

The APR-DRG severity of illness classification is designed to capture the extent of physiologic decompensation or organ system loss of function.[23] Each patient is first classified in an APR-DRG according to the principal diagnosis or procedure (e.g., cardiac valve procedure with catheterization, angina). All secondary diagnoses are then assigned to one of four groups: minor, moderate, major, or extreme. For example, the severity level for respiratory diagnoses progresses from bronchitis (minor), to asthma with status asthmaticus (moderate), to viral pneumonia (major), and finally to respiratory failure (extreme). Next, the algorithm adjusts upward the base classification of secondary diagnoses for more "severe" APR-DRGs (e.g., bypass) and computes the base severity level as the maximum level of any secondary diagnosis. Finally, the system reserves the major and extreme severity classes to patients with *multiple* major or extreme co-morbid diagnoses. Requiring multiple serious complications to be classified in the major and extreme categories avoids classifying all patients in a "serious" APR-DRG (e.g., bypass, hip fracture) at the top levels of severity.

One potential concern associated with the gainsharing model is the possibility that physicians who have admitting privileges in multiple hospitals might have incentives to refer their more severe cases to other non-gainsharing hospitals. If this occurred, we would expect to see a smaller increase or a decline in major or extreme severity cases at participating hospitals between the base and intervention year relative to the comparison hospitals.

Table 8-3 compares the percentage of admissions classified with an APR severity index of major or extreme at demonstration and comparison hospitals. For BIMC, the percentage of admissions classified as major or extreme severity increased by 27.71 percent between the base and performance years. The BIMC comparison sites also showed a similar, though smaller, increase of 19.29 percent. These finding suggests that BIMC physicians did not avoid major or extreme severity inpatient admissions as a result of the demonstration. At CAMC, severity of illness as measured by major or extreme severity classification decreased by 0.49 percent, whereas the comparison sites increased by 24.93 percent. From this analysis, we find no evidence that CAMC avoided more severe cases relative to its baseline patient mix.

Physician Referral Patterns: As a result of the financial incentives in gainsharing, participating sites may have had an incentive to avoid potentially high-cost admissions. Patients with APR-DRG severity scores that are major or extreme, potential cost outliers, or patients that are admitted through the ER are all potentially high cost. However, the hospitals' ability to steer high-cost patients to another facility is limited unless specialized tertiary care is available at other hospitals. Physicians are instrumental in which patients are admitted to a particular hospital. Physicians often have admitting privileges at multiple hospitals and can refer a patient to one hospital over another, regardless of the cost to the hospital or patient.

[23] Risk of mortality, the other dimension, captures differential risks of dying. Because so few hospital patients die during the inpatient admission, we believe that 3M's severity of illness classification is better suited to measure severity differences in referral patterns. 3M also has developed relative cost weights for all the APR-DRGs and their four severity levels. We investigated their use but found that they overstate the severity of patients who are undergoing expensive, but not necessarily risky or severe procedures (e.g., PTCA).

Table 8-3
Percentage of inpatient admissions with major or extreme severity in demonstration and comparison hospitals, base year and performance period

Demonstration site or comparison hospitals	Base year	Performance year	Percent change
Beth Israel Medical Center (BIMC)	20.6	31.6	53.4
BIMC comparison hospitals	28.8	39.3	36.4
Charleston Area Medical Center (CAMC)	42.6	42.4	−0.5
CAMC comparison hospitals	35.9	44.9	24.9

NOTES: For both BIMC and CAMC, the base year was calendar year 2007. BIMC's performance year was PY3 (October 1, 2010, through September 30, 2011). CAMC's performance "year" was December 1, 2008, through December 31, 2009. The percentage of major or extreme severity was based on claims with an all patient refined diagnosis-related group (APR-DRG) and severity.

SOURCE: RTI analysis of Medicare claims.

In this demonstration, an internal facility cost-saving component is shared between the physicians and hospitals. Consequently, physician financial incentives may be more aligned with hospital financial incentives to avoid potentially high-cost patients or to transfer high-cost patients once they are admitted. This section examines physician referral patterns in participating hospitals.

Physician Referral—Admission Patterns: To gauge the number of acute care hospitals to which participating physicians may have admitting privileges, we used all eligible inpatient claims in New York for BIMC. For CAMC, claims from hospitals located in West Virginia were used as well as claims from its comparison hospitals. From this analysis we were able to determine the number of physicians at each hospital who, on the basis of Medicare claims, were also admitting physicians at another market area hospital. ***Table 8-4*** shows, for each time period, the number of physicians with admissions at both a participating demonstration hospital and another market area acute care hospital. The first set of columns in Table 8-4 shows the demonstration participating hospitals. The second set of columns in Table 8-4 shows the number of responsible physicians identified with gainsharing-eligible admissions at each demonstration hospital. The third set of columns shows the number of gainsharing admitting physicians at a particular demonstration hospital who also admitted at a comparison hospital. The last set of columns shows the number of participating physicians at the demonstration hospital who also admitted patients at another acute care hospital that was neither a participating nor a comparison hospital.

The number of participating physicians at BIMC increased from 433 during the base year to 478 in PY3. However, the increase peaked at 500 in PY2. The number at CAMC decreased from 33 during the base year to 22 in PY1. The total number of participating physicians at CAMC is much lower than at BIMC because the CAMC demonstration was focused on cardiac procedures, whereas BIMC included most DRGs. Identifying the base year number of

Table 8-4
Number of responsible physicians who admitted at a demonstration hospital and also at a comparison or other acute care hospital

	Number of gainsharing physicians at demonstration hospital				Number of gainsharing physicians at demonstration hospital who also admit at a comparison hospital				Number of gainsharing physicians at demonstration hospital who also admit at another market area acute care hospital			
Demonstration Hospital	Base year	PY 1	PY 2	PY 3	Base year	PY 1	PY 2	PY 3	Base year	PY 1	PY 2	PY 3
Beth Israel Medical Center (BIMC)	433	456	500	478	77	78	103	60	96	81	78	69
Charleston Area Medical Center (CAMC)	33	22	—	—	n/a	n/a	—	—	9	2	—	—

NOTES:

1. For both BIMC and CAMC the base year is calendar year 2007
2. BIMC's first performance year (PY1) is October 1, 2008 through September 30, 2009
3. BIMC's PY2 is October 1, 2009 through September 20, 2010
3. BIMC's PY3 is October 1, 2010 through September 30, 2011
4. CAMC's only performance period is December 1, 2008 through January 31, 2010
5. n/a: CAMC does not have any comparison hospitals located within 100 miles of CAMC.

SOURCE: RTI Analysis of Medicare claims

physicians was difficult.[24] These physician identification problems might be responsible for the apparent decrease in the number of participating physicians at CAMC.

The number and share of BIMC participating physicians admitting patients to comparison hospitals increased slightly from the base year to PY2, but then fell in PY3. Some of the increase might be due to these physicians increasing their admissions at comparison hospitals. Some of the increase, however, might be due to changes in hospital affiliation by physicians over time. CAMC did not have any comparison hospitals within reasonable proximity (100 miles) of Charleston.

The number and share of BIMC participating physicians admitting patients to other (non-comparison) hospitals located in New York City decreased from the base year to PY3. This result stands in contrast to the experience with the comparison hospitals. The number of CAMC participating physicians admitting patients to other Charleston hospitals fell from nine to two. As with the total number of CAMC's participating physicians in the base year, these figures need to be interpreted cautiously. In any event, CAMC performs most of the open heart procedures in West Virginia. With or without the Gainsharing demonstration, individual CAMC physicians have little incentive to move cardiac patients from CAMC to any other hospitals.

Physician Referral—Transfer Analysis: Participating physicians may have an incentive to transfer very costly patients, once admitted to a demonstration hospital, to other acute care hospitals (IPPS transfers). They might also have an incentive to discharge patients sooner than otherwise to PAC facilities such as long-term care hospitals, rehabilitation hospitals and units, and SNFs. Should these occur, these would be of significant concerns to CMS.

To test this hypothesis, we calculated the share of demonstration admissions that were transferred to another facility as well as the share of transfers that were major or extreme severity cases. We then compared the changes in transfers between base year and the performance period for both demonstration and comparison hospitals. We defined a transfer from one facility to another facility as any two claims for the same patient in which (1) the admission date of the second claim was within one day of the discharge of the first claim and (2) the two hospital provider IDs did not match (i.e., not a readmission to the same hospital).[25]

Table 8-5 shows transfer rates of discharges from participating and comparison hospitals in the base year and the performance period. Transfer rates (to all sources, including other acute facilities and PAC facilities) increased from about 14 percent to 15.7 percent for BIMC and from 15.9 percent to 17.8 percent for its comparison hospitals. For CAMC, transfers to all facilities fell from 6.8 percent to about 6.1 percent, whereas they increased from 11.3 percent to 14.7 percent for its comparison hospitals.

---

[24] Identifying the base year number of physicians was difficult because 2007 was a transition year from the UPINs to the NPIs. Because the National Plan and Provider Enumeration System (NPPES) does not contain a complete "crosswalk" of UPINs to NPIs, the coding of the physician identifiers on the claims was poor. Also, some claims had neither a UPIN nor an NPI.

[25] We chose this definition because it does not rely on the sometimes inaccurate discharge destination reported on the claim but rather on dates of service for matched claims for the same patient.

Table 8-5
Transfer rates of discharges from demonstration and comparison hospitals, total and to other acute care, by base year and performance period

Demonstration site or comparison hospitals	Total discharges transferred, base year (%)	Total discharges transferred, performance year (%)	Transferred to other acute care, base year (%)	Transferred to other acute care, performance year (%)
Beth Israel Medical Center (BIMC)	14.0	15.7	1.7	1.6
BIMC comparison hospitals	15.9	17.8	2.2	1.9
Charleston Area Medical Center (CAMC)	6.8	6.1	0.2	0.0
CAMC comparison hospitals	11.3	14.7	0.25	0.27

NOTES: For both BIMC and CAMC, the base year was calendar year 2007. BIMC's performance year was PY3 (October 1, 2010, through September 30, 2011). CAMC's performance period is December 1, 2008, through December 31, 2009. Sample sizes for long-term care and acute care hospitals for CAMC and its comparison hospitals are small ($n \leq 6$). Total transfers include transfers to another acute care hospital, skilled nursing facility, long-term care facility, and other hospitals not in the inpatient prospective payment system. Discharges to home health are excluded.

SOURCE: RTI analysis of Medicare claims.

Transfers to IPPS acute care hospitals decreased for both BIMC and its comparison hospitals: from 1.7 to 1.6 percent for BIMC and from 2.2 to 1.9 percent for its comparison hospitals. The IPPS transfer rates for CAMC and its comparison hospitals are extremely low because these hospitals are the hospitals, by definition, where complicated open heart procedures are performed. The most IPPS transfers CAMC and its comparison hospitals had, during any given year, was six (at a comparison hospital). In fact, CAMC's IPPS transfers fell to zero during the demonstration, while the IPPS transfer rate for its comparison hospitals increased very slightly, from 0.25 percent to 0.27 percent.

These results can be difficult to interpret. These descriptive findings alone are insufficient to determine whether increased transfers among physicians practicing at demonstration hospitals were specifically in response to incentive payments. Many factors influence physicians' decisions to transfer patients, including the severity of the admission, the hospital's ability to care for very severe admissions, and the availability of other hospitals to which patients can be transferred.

Another variant on our transfer analysis relates specifically to incentives to use transfers to avoid sicker, more acute care patients. Under the gainsharing methodology, participating physicians may have an increased incentive to transfer sicker patients. To test this hypothesis,

we examined the share of IPPS transfers from demonstration hospitals for patients with an APR severity index of major or extreme. ***Table 8-6*** shows the percentage of base year and performance period IPPS transfers that were classified as major or extreme severity cases. The share of such transfers by BIMC increased from 26.1 percent to 33.1 percent by PY2 and then fell back to 26.0 percent in PY3. For its comparison hospitals, it increased from 28.4 percent to 37.0 percent.

Table 8-6
Percentage of patients transferred to acute care hospitals that have major or extreme severity, from demonstration and comparison hospitals

Demonstration site or comparison hospitals	Base year	Performance year
Beth Israel Medical Center (BIMC)	26.1	26.0
BIMC comparison hospitals	28.4	37.0
Charleston Area Medical Center (CAMC)	100.0	0.0
CAMC comparison hospitals	50.0	71.4

NOTES: For both BIMC and CAMC, the base year was calendar year 2007. BIMC's performance year was PY3. CAMC's performance year was December 1, 2008, through December 31, 2009. The percentage of major or extreme severity cases was based on claims with an all patient refined diagnosis-related group (APR-DRG) and severity. Sample sizes for CAMC and its comparison hospitals are small. In both the base year and the performance period, the number of admissions with any severity marker was at most five.

SOURCE: RTI analysis of Medicare claims.

CAMC had only 2 IPPS transfers during the base period, whereas its comparison hospitals had only 15 (one of which was not grouped). And CAMC did not have any IPPS transfers during the demonstration. The major/extreme shares of IPPS transfers are high in Table 8-6, not only for both CAMC and its comparison hospitals, but for BIMC and its comparison hospitals as well. An important reason why these shares are high is because the number of any IPPS transfers is low at these hospitals.

Physician Referral—Overlap Between Participating and Comparison Hospitals: Given the geographic proximity of BIMC to its comparison hospitals, there is at least a practical possibility that BIMC's participating physicians are also admitting at comparison hospitals. In Table 8-4, we saw that the same participating physicians were admitting at both participating and comparison hospitals. This may create a potential bias in our analyses, making the cost savings appear greater if participating physicians admit their sicker, more costly, patients to comparison hospitals.[26] In this section, we more closely investigate the overlap in physicians between

[26] Alternatively, there could also be a negative bias, underestimating cost savings, if participating physicians' new, more efficient behavior spills over to the comparison hospitals.

participating and comparison hospitals for BIMC.[27] ***Table 8-7*** shows the number of physicians with admissions at both BIMC and comparison hospitals. ***Table 8-8*** shows the number of admissions by BIMC participating physicians at a comparison hospital. Table 8-7 shows no discernable pattern in the number of BIMC physicians admitting to comparison hospitals. The number is rising at for some comparison hospitals (e.g., Lenox Hill Hospital), steady at others (e.g., Maimondies), and falling at others (e.g., New York Methodist). The changes in the number seem low as well.

Table 8-7
Number of Beth Israel Medical Center physicians admitting at comparison hospitals, by base or performance year

	Admissions during			
Comparison Hospital	Base year	PY1	PY2	PY3
Mount Sinai Hospital	7	7	4	9
New York Hospital Medical Center of Queens	0	2	3	2
Brooklyn Hospital Center at Downtown Campus	2	2	1	2
Lenox Hill Hospital	11	13	13	17
Staten Island University Hospital	1	2	4	4
Flushing Hospital Medical Center	1	1	1	2
Maimonides Medical Center	14	16	15	8
Long Island Jewish Medical Center	2	1	2	1
NYU Hospitals Center	2	4	5	9
Wyckoff Heights Medical Center	0	1	1	1
Brookdale Hospital Medical Center	12	9	7	9
New York Methodist Hospital	24	20	17	19
Lutheran Medical Center	3	6	7	7

SOURCE: RTI analysis of Medicare claims.

[27] Because CAMC's comparison hospitals are too distant from Charleston, this analysis was performed only for BIMC physicians.

Table 8-8
Number of admissions by Beth Israel Medical Center participating physicians from a comparison hospital in a base or performance year

Comparison Hospital	Admissions during			
	Base year	PY1	PY2	PY3
Mount Sinai Hospital	32	26	36	53
New York Hospital Medical Center of Queens	0	10	19	11
Brooklyn Hospital Center at Downtown Campus	7	20	3	16
Lenox Hill Hospital	188	180	153	214
Staten Island University Hospital	8	76	73	35
Flushing Hospital Medical Center	1	1	3	3
Maimonides Medical Center	298	191	176	78
Long Island Jewish Medical Center	22	11	13	5
NYU Hospitals Center	2	33	17	76
Wyckoff Heights Medical Center	0	2	2	5
Brookdale Hospital Medical Center	75	112	94	83
New York Methodist Hospital	564	406	389	502
Lutheran Medical Center	56	56	90	42
Total admissions by BIMC physicians at comparison hospitals	1,253	1,124	1,068	1,123
As a share of total admissions at comparison hospitals	1.87%	1.71%	1.59%	1.66%

SOURCE: RTI analysis of Medicare claims.

Although BIMC participating physicians admitted patients to all of its comparison hospitals, four comparison hospitals had more admissions from these physicians than other comparison hospitals: Lenox Hill, Maimonides, Brookdale, and New York Methodist (Table 8-8). Lenox Hill is located in Manhattan and the other three are located in Brooklyn, where BIMC's Kings Highway campus is located. As discussed earlier, the physicians at Kings Highway are community physicians who admit to many hospitals. By contrast, the physicians at BIMC's Petrie campus are often teaching physicians and thus less inclined to admit patients at other hospitals.

Overall, of the average 66,827 eligible admissions at comparison hospitals, 1,142 or 1.7 percent are attributed to physicians with eligible admissions at BIMC. This rate fell from 1.8 to 1.59 percent in PY2 but then rose slightly to 1.66 percent in PY3 (Table 8-8). This level of overlapping admitting privileges indicates that any bias resulting from admissions by BIMC physicians at comparison hospitals is likely to be small.

8.4 Multivariate Results

Severity of Admission: In Table 8-3, we presented descriptive results that showed that BIMC had an increase in admissions classified as major or extreme severity from the base year to the third demonstration year. CAMC, on the other hand, had a slight decrease in the share of admissions classified as major or extreme severity. We had hypothesized that physicians have an incentive to direct more severe cases to other hospitals where they have privileges in order to increase their demonstration payments. We tested this hypothesis more fully by controlling for factors such as change in case mix that might affect the share of high-severity admissions. ***Table 8-9*** shows the multivariate results. Because the logistic coefficients cannot be used directly to measure the effects of each explanatory variable, they need to be transformed. Two types of transformation were performed: (1) odds ratios and (2) the simulated changes for the 2Ds estimates as described in ***Section 6***.

The results for BIMC and CAMC differ, possibly because CAMC was focused only on cardiac procedures. Both logit regressions show an increasing likelihood of an admission being major or severe severity as the age of the beneficiary increases. For instance, patients 80 years old or more at BIMC (and its comparison hospitals) are about twice as likely as patients 65 to 69 years old to be classified major or extreme severity. Admissions for females are less likely than those for males to be major or extreme at BIMC and its comparison hospitals but more likely to be major or extreme at CAMC and its comparison hospitals. Nonwhites are more likely than whites to be classified as major or extreme severity. For BIMC and its comparison hospitals, the strongest predictor of severity is admission from a SNF, with admission through the ER the second highest. For CAMC and its comparison hospitals, the strongest predictor of severity is admission through the ER.[28] BIMC patients with higher risk scores increased the likelihood of an admission being major or severe severity.[29]

To test our hypothesis that the demonstration could give participating physicians incentives to decrease their admissions of severe cases, we needed to look at the parameter estimates for the participating hospital indicator, performance period indicator, and their interaction term (2D)—all three are statistically significant in both the BIMC and CAMC logit regressions. The parameter estimate for the participating hospital indicator is negative for BIMC and positive for CAMC. BIMC was less likely than its comparison hospitals to have major or extreme severity admissions during the base year. CAMC, on the other hand, was more likely than its comparison hospitals to have major or extreme severity admissions during the base year. The performance period indicator shows about a 40 percent increase in the likelihood of admission during the demonstration at the comparison hospitals for BIMC and CAMC.

---

[28] The effect of admissions from SNFs could not be tested for CAMC.

[29] The HCC risk score was not included in the final CAMC logistic regression since it was not statistically significant during preliminary work. The likely reason it was not statistically significant is because most of the cardiac procedures are not undertaken unless there is a good prospect of recovery. And, the health of heart patients are more homogeneous than the general patient population of hospitals.

Table 8-9
Probability that an admission has major or extreme severity

Explanatory variable	Beth Israel Medical Center				Charleston Area Medical Center			
	Logistic coefficient	Robust standard error	P > | t |	Odds ratio	Logistic coefficient	Robust standard error	P > | t |	Odds ratio
Patient age 0–64	0.186	0.055	0.001	1.204	0.122	0.109	0.261	1.1297
Patient age 70–74	0.077	0.021	0.000	1.080	0.112	0.058	0.052	1.1185
Patient age 75–79	0.235	0.035	0.000	1.265	0.254	0.076	0.001	1.2894
Patient age 80 or more	0.631	0.057	0.000	1.880	0.422	0.049	0.000	1.5244
Female	-0.107	0.015	0.000	0.898	0.088	0.044	0.047	1.0916
Nonwhite	0.151	0.071	0.032	1.163	0.481	0.085	0.000	1.6171
Admission from a skilled nursing facility	1.181	0.101	0.000	3.258	omitted	n/a	n/a	n/a
Admission through the emergency room	0.939	0.200	0.000	2.557	0.553	0.066	0.000	1.7388
Diagnosis-related group weight	0.742	0.054	0.000	2.100	0.461	0.037	0.000	1.5849
HCC risk score	0.215	0.009	0.000	1.240	omitted	n/a	n/a	n/a
Participating hospital indicator	-0.574	0.063	0.000	0.563	0.293	0.111	0.008	1.3408
Performance period indicator	0.337	0.037	0.000	1.400	0.397	0.082	0.000	1.4871
Difference-in-difference estimator	0.152	0.038	0.000	1.165	−0.411	0.082	0.000	0.6632
Constant term	-3.455	0.206	0.000	0.032	−2.955	0.079	0.000	0.0521
Pseudo R^2	0.1562	n/a	n/a	n/a	0.1285	n/a	n/a	n/a
Number of observations	304,196	n/a	n/a	n/a	11,883	n/a	n/a	n/a

NOTE: Claims used in the regressions for both demonstration sites includes claims from their respective comparison hospitals as well as from the demonstration hospitals themselves.

SOURCE: RTI analysis of Medicare claims.

The simulated probabilities needed for estimating the 2D effect are shown in ***Table 8-10***. The probability of a major or extreme severity admission at BIMC was 21.5 percent during the base period and increased to 28.8 percent during the demonstration, an increase (difference) of 7.3 percentage points. The probability of a major or extreme severity admission at BIMC's comparison hospitals was 30.2 percent during the base period and increased to 36.3 percent during the demonstration, an increase of 6.1 percentage points. The 2D effect then is 7.3 percentage points minus 6.1 percentage points, which equals 1.2 percentage point. That is, BIMC's probability of a major or extreme admission increased 1.2 percentage point more than did those of its comparison hospitals. This result provides evidence that the Gainsharing demonstration did not deter BIMC from admitting major or extreme severity cases any more than its competitors.

Table 8-10
Estimates of difference in differences for major/extreme admissions
(based on regression results from Table 8-9)

Demonstration site or comparison hospitals	Mean probability that index admission is classified as major/extreme		Performance minus base period	Difference in differences
	Base period	Performance period		
Beth Israel Medical Center (BIMC)	21.5%	28.8%	7.3%	1.2%
BIMC comparison hospitals	30.2%	36.3%	6.1%	—
Charleston Area Medical Center (CAMC)	42.3%	42.0%	−0.3%	−8.3%
CAMC comparison hospitals	36.4%	44.4%	8.0%	—

SOURCE: RTI processing of Medicare claims.

The probability of a major or extreme severity admission at CAMC was 42.3 percent during the base period and fell slightly to 42.0 percent during the demonstration, a decrease (difference) of 0.3 percentage points. The probability of a major or extreme severity admission at CAMC's comparison hospitals was 36.4 percent during the base period and increased to 44.4 percent during the demonstration, an increase of 8 percentage points. The 2D effect then is −0.3 percentage points minus 8 percentage points, which equals −8.3 percentage points. That is, CAMC's probability of a major or extreme severity admission was 8.3 percentage points less than that of its comparison hospitals. This result could be due to upcoding at the comparison hospitals, as the probability of major or extreme severity admissions barely changed at CAMC whereas, for its comparison hospitals, the probability was lower than CAMC's in the base year and increased to a level above CAMC's during the demonstration.

Since the BIMC demonstration lasted three years, we were able to test for year-specific 2D effects. The logistic regression containing year-specific 2D coefficients for BIMC are shown in ***Table 8-11***. Aside from the performance year and 2D year-specific variables the values of the

Table 8-11
Probability that a BIMC or BMIC-comparison hospital admission has major or extreme severity, year specific

Explanatory variable	Logistic coefficient	Robust standard error	P > | t |	Odds ratio
Patient age 0 to 64	0.185	0.055	0.001	1.204
Patient age 70 to 74	0.077	0.021	0.000	1.080
Patient age 75 to 79	0.236	0.034	0.000	1.267
Patient age 80 plus	0.633	0.057	0.000	1.884
Female	-0.106	0.015	0.000	0.899
Non-white	0.150	0.071	0.033	1.162
Admission from a skilled nursing facility	1.167	0.100	0.000	3.214
Admission through the emergency room	0.933	0.202	0.000	2.542
DRG weight	0.741	0.055	0.000	2.099
Risk score	0.216	0.009	0.000	1.241
Participating hospital indicator	-0.573	0.062	0.000	0.564
Performance year indicator				
Performance year 1	0.250	0.050	0.000	1.284
Performance year 2	0.342	0.028	0.000	1.408
Performance year 3	0.413	0.066	0.000	1.512
2D estimator				
Performance year 1	0.121	0.051	0.018	1.129
Performance year 2	0.111	0.031	0.000	1.118
Performance year 3	0.205	0.070	0.003	1.227
Constant term	-3.453	0.209	0.000	0.032
Pseudo R^2	0.1567	—	—	—
Number of observations	304,196	—	—	—

SOURCE: RTI processing of Medicare claims

logistic coefficients for the other variables are nearly the same as in Table 8-9. The coefficients for performance year indicators progressively increase from 0.25 in PY1 to 0.41 in PY3. The 2D coefficients for the first two performance years are about the same while there is a large increase to 0.205 in PY3.

The simulated probabilities needed for estimating the year-specific 2D effects are shown in ***Table 8-12***. The 2D effect for PY1 indicates BIMC's likelihood of a major/extreme admission was 1.0 percentage point higher than for the comparison group. The 2D effect fell in PY2 to 0.6 percentage point but then increased to 2.0 percentage points in PY3. So, the 2D effect varied by year with a small decrease during PY2 but rebounding to its highest level in PY3. The year-specific 2D effects are consistent with the overall 2D effect shown in Table 8-10.

Table 8-12
Year-specific estimates of difference in differences for major/extreme admissions (based on regression results from Table 8-11)

BIMC and comparison hospitals	Mean probability that index admission is classified as major/extreme		Performance minus base period	Difference in differences
	Base period	Performance period		
Performance Year 1				
Beth Israel Medical Center	21.5%	26.9%	5.4%	1.0%
BIMC Comparison Hospitals	30.2%	34.6%	4.4%	—
Performance Year 2				
Beth Israel Medical Center	21.5%	28.2%	6.7%	0.6%
BIMC Comparison Hospitals	30.2%	36.4%	6.2%	—
Performance Year 3				
Beth Israel Medical Center	21.5%	31.0%	9.5%	2.0%
BIMC Comparison Hospitals	30.2%	37.7%	7.5%	—

NOTE: all estimates of the difference in differences statistically significant at the 5% level or better

SOURCE: RTI processing of Medicare claims

Transfers: We hypothesized that physicians have an incentive to transfer sicker patients to other acute-care hospitals where they have privileges in order to increase their demonstration payments. To test this hypothesis, we modeled the probability of an IPPS transfer from a participating hospital and a comparison hospital using a logistic model. ***Table 8-13*** shows coefficients for BIMC. A logit regression could not be estimated for CAMC because CAMC did not have any IPPS transfers during the demonstration (Table 8-6).

Table 8-13
Probability of a BIMC or a BIMC-comparison hospitalization triggering an IPPS transfer

Explanatory variable	Logistic coefficient	Robust standard error	P > | t |	Odds ratio
Patient age 0–64	-0.067	0.069	0.332	0.935
Patient age 70–74	-0.002	0.045	0.958	0.998
Patient age 75–79	0.059	0.069	0.392	1.061
Patient age 80 plus	-0.277	0.099	0.005	0.758
Female	-0.013	0.108	0.902	0.987
Nonwhite	-0.147	0.163	0.366	0.863
Admission from a skilled nursing facility	-0.810	0.278	0.004	0.445
Admission through the emergency room	-0.316	0.208	0.129	0.729
DRG weight	0.068	0.020	0.001	1.071
HCC risk score	-0.068	0.046	0.141	0.934
Participating hospital indicator	-0.144	0.211	0.495	0.866
Performance period indicator	-0.059	0.046	0.193	0.942
2D estimator	0.013	0.044	0.765	1.013
Constant term	-3.467	0.325	0.000	0.031
Pseudo R^2	0.0119	—	—	—
Number of observations	303,645	—	—	—

SOURCE: RTI processing of Medicare claims.

The regression results are very weak. The pseudo R^2 is only 0.0119—a very low value. Only three of the explanatory variables were statistically significant: patients 80 years old or older, admission from a SNF, and the DRG weight. Patients aged 80 years old or older were less likely to have an IPPS transfer than patients 65–69 years old. Patients admitted from a SNF were also less likely to have an IPPS transfer. The higher the DRG weight, the higher the likelihood of a transfer.

The 2D estimator, as well as its components, is not statistically significant. Because the 2D estimator and its components are not statistically significant, we did not simulate the probabilities of IPPS transfers using the method shown in Table 8-10. The results provide evidence that the demonstration did not change the likelihood of an IPPS transfer from BIMC.

With three years of demonstration experience, we were able to test for year-specific 2D effects for BIMC IPPS transfers. The logistic regression containing year-specific 2D coefficients for BIMC are shown in ***Table 8-14***. Aside from the performance year and 2D year-specific variables the values of the logistic coefficients for the other variables are nearly the same as in Table 8-9. The coefficients for performance year indicators progressively decrease from 0.012

Table 8-14
Probability of a BIMC or a BIMC-comparison hospitalization triggering an IPPS transfer, year specific

Explanatory variable	Logistic coefficient	Robust standard error	P > | t |	Odds ratio
Patient age 0–64	-0.067	0.069	0.334	0.936
Patient age 70–74	-0.002	0.045	0.961	0.998
Patient age 75–79	0.058	0.069	0.400	1.060
Patient age 80 plus	-0.278	0.100	0.005	0.757
Female	-0.014	0.108	0.898	0.986
Nonwhite	-0.146	0.163	0.369	0.864
Admission from a skilled nursing facility	-0.799	0.283	0.005	0.450
Admission through the emergency room	-0.313	0.210	0.137	0.731
DRG weight	0.068	0.020	0.001	1.071
HCC risk score	-0.068	0.046	0.140	0.934
Participating hospital indicator	-0.144	0.211	0.493	0.866
Performance year indicator				
Performance year 1	0.012	0.053	0.828	1.012
Performance year 2	-0.074	0.058	0.204	0.929
Performance year 3	-0.121	0.068	0.077	0.886
2D estimator				
Performance year 1	-0.121	0.053	0.023	0.886
Performance year 2	0.072	0.056	0.199	1.074
Performance year 3	0.087	0.062	0.158	1.091
Constant term	-3.467	0.326	0.000	0.031
Pseudo R^2	0.0122	—	—	—
Number of observations	303,645	—	—	—

SOURCE: RTI processing of Medicare claims.

in PY1 to -0.121 in PY3, none of them were statistically significant. The 2D coefficients progressively increase from -0.121 in PY1 to 0.087 in PY3. Only the 2D coefficient for PY1 is statistically significant. This indicates that the 2D effect was transitory.

The simulated probabilities needed for estimating the year-specific 2D effects are shown in ***Table 8-15***. The probability of an IPPS transfer from BIMC was 1.84 percent during the base period and fell slightly to 1.65 percent during PY1, a decrease (difference) of 0.19 percentage points. The probability of an IPPS transfer from BIMC's comparison hospitals was 2.12 percent during the base period and increased slightly to 2.14 percent during PY1, an increase of

Table 8-15
Year-specific estimates of difference in differences for IPPS transfers
(based on regression results from Table 8-14)

BIMC and comparison hospitals	Mean probability that index admission triggered an IPPS transfer		Performance minus base period	Difference in differences
	Base period	Performance period		
Performance Year 1				
Beth Israel Medical Center	1.84%	1.65%	-0.19%	-0.21%
BIMC Comparison Hospitals	2.12%	2.14%	0.02%	—
Performance Year 2				
Beth Israel Medical Center	1.84%	1.83%	0.00%	0.14%
BIMC Comparison Hospitals	2.12%	1.97%	-0.15%	—
Performance Year 3				
Beth Israel Medical Center	1.84%	1.78%	-0.06%	0.18%
BIMC Comparison Hospitals	2.12%	1.88%	-0.24%	—

NOTE: Only Performance Year 1 difference in differences statistically significant at the 5% level.

SOURCE: RTI processing of Medicare claims.

0.02 percentage points. The 2D effect for PY1 then is -0.19 percentage points minus 0.02 percentage points, which equals -0.21 percentage point. That is, BIMC's probability of an IPPS transfer decreased 0.21 percentage point relative to its comparison hospitals. The year-specific 2D effects for PY2 and PY3, had positive values, but they were not statistically significant—they are shown in the table for reporting completeness.

8.5 Discussion

As part of the demonstration, physicians practicing at demonstration participating hospitals may have had an increased financial incentive to avoid potentially high-cost admissions. In this section, we conducted a market analysis of demonstration and comparison hospitals. We also analyzed admitting patterns for physicians admitting at both demonstration and comparison hospitals.

The descriptive results indicate little, if any, demonstration impact on most referral performance measures. In some cases, changes in a performance measure at BIMC and CAMC were about the same as at their comparison hospitals. For instance, the share of admissions with an outlier payment increased about the same amount for both the participating and the comparison hospitals. For BIMC and its comparison hospitals, admissions increased as some New York City hospitals closed. BIMC attributed its ability to increase admission to lower

LOSs due to the demonstration. But the same could have happened at its comparison hospitals, as LOS fell for them as well. Some of the observed changes need to be interpreted cautiously, as the number of cases used in analyses was low. For instance, the number of IPPS transfers was low for both demonstration and comparison hospitals. Although many of BIMC's Kings Highway campus physicians had admitting privileges at both comparison hospitals and other local competitors, there does not seem any large change in admitting behavior between the base period and the demonstration period.

In our multivariate analyses, we found that the direction of the demonstration impact on severity of admissions was different for BIMC and CAMC. Even though the direction of the demonstration impact differed, the results indicate that neither BIMC nor CAMC avoided major or extreme severity admissions. The results also indicate that the Gainsharing demonstration had a transitory impact on BIMC IPPS transfers.

This page left intentionally blank.

SECTION 9
SUMMARY OF RESULTS

The primary goal of the Gainsharing demonstration was to evaluate gainsharing strategies aimed at improving the quality of care and efficiency in health delivery systems. The demonstration sites implemented approaches that may better align physician and hospital financial incentives and ultimately lead to reductions in the overall internal hospital costs of care. The gainsharing sites were required to maintain quality of care. Participating sites were not required to generate Medicare program savings, but they were required to maintain budget neutrality within the inpatient stay and up to 30 days beyond in any post-acute settings. In this evaluation we analyzed the impact of the Gainsharing demonstration gainsharing models on hospital efficiency, physician practice patterns, Medicare expenditures, quality, and beneficiary satisfaction. This final report is focused on the available performance years for both sites. For CAMC, we evaluated impacts during the single performance year in which this site participated. For BIMC, we presented findings from the full three performance years.

The findings presented in this report have a few limitations. The CAMC site operated for only 1 year and was limited to a subset of cardiac-related DRGs. Both of these factors limit the generalizability of CAMC's gainsharing experience and performance and our analysis of them. The findings reported for BIMC are more robust because they represent 3 performance years for a wide range of DRGs. Finally, our analytic approach for both sites focused on a 2D methodology. This approach is useful for the purposes of estimating the impacts of a complex intervention while controlling for secular trends in key outcomes that may have occurred even in the absence of the demonstration. However, results based on this methodology should be interpreted with some consideration of the relative baseline values of the intervention and control groups. Particularly with regard to quality of care, groups whose performance at baseline approaches high performance have less ability to improve over time.

Operational Experiences: One element of our evaluation focused on the performance and operational experiences of the participating sites. We gathered this information through a series of site visits with hospital leadership and staff, supplemented by focus group discussions with physicians participating in gainsharing. The two Medicare Gainsharing Demonstration sites, CAMC and BIMC, implemented different gainsharing methodologies, each with a different clinical focus. For these reasons, direct comparisons between the findings of these two different sites should be made with caution. Still, some common themes that emerged from our site visits and physician focus groups may point to lessons learned about the overall gainsharing model. First, both sites agreed that gainsharing is a promising model for health care reform but that in practice the model may work better for some hospitals than others. Overall, individuals we spoke with at both CAMC and BIMC felt that the gainsharing model was theoretically a promising way to improve health care delivery by improving physicians' awareness of cost and by better aligning hospital and physician financial incentives. Both sites also reported that the gainsharing model improved communication between physicians and hospital administration on issues related to lowering costs and maintaining or improving quality of care.

Second, sites found that it was challenging to fully educate participating physicians on the gainsharing reporting metrics, their underlying data, and the overall purpose of the project. Although the sites had different clinical focuses, both provided similar detailed quality-of-care

and cost performance reports to participating physicians as a way to substantiate the payment (or nonpayment) of gainsharing incentives. However, our site visits and focus groups found that many physicians didn't understand the reports, although this situation appeared to improve somewhat during the second site visit to BIMC. Across both sites we heard comments that the physician reporting was, for example, "incomprehensible" or "overly complex," despite the efforts of hospital administration to explain the basis for gainsharing incentive payments. Physician understanding did improve somewhat over time but this increased level of understanding came at what hospital leadership reported as considerable effort and cost. One-on-one meetings between hospital leadership and physicians, as implemented by BIMC, seemed to be a promising way to improve physician understanding of and buy-in to the metrics. But because physicians generally struggled with understanding and accepting the performance information presented, the perceived direct link between the actual amount of gainsharing payments and improvements in cost and quality of care was not always clear to the participating physicians. Also, physicians generally commented that the performance data were often "too old" (performance data presented to physicians was generally lagged between 6 and 9 months) and they felt that more timely information would be more powerful. Leadership in both sites felt that the overall gainsharing model was successful, but mostly in improving the communication between hospitals and physicians and in improving awareness about the need to lower costs and maintain and improve quality of care. We asked hospital leadership to describe specific examples of changes in the way participating physicians practiced care in their hospital. BIMC hospital leadership had difficulty in providing specific examples beyond feedback from physicians that they approved of additional resources (such as improved access to discharge planners and social workers) and a marked improvement in timely completion of medical records and documentation (metrics that were used by BIMC as qualifiers for incentive payments). Beyond these, we were not able to identify additional specific examples of instances where physician behavior changed in response to gainsharing metrics.

Finally, we found that ongoing operational success of a gainsharing model depended greatly on a strong organizational champion and on significant investment, in time and other resources, by the participating site. Both demonstration sites discussed the substantial level of effort required to implement and maintain a gainsharing model. Both sites also stressed that making gainsharing work in practice requires an internal champion within the implementing organization.

Medicare Expenditures and Savings: One goal of the Gainsharing demonstration was to reduce the hospital's internal costs enough to generate savings that could be shared with participating physicians. Both CAMC and BIMC determined that internal savings were generated and incentive payments were therefore made to participating physicians. Savings to the Medicare program through negotiated discounts were not required in this demonstration, and since Medicare did not directly change the DRG-based payment methodology for participating hospitals, no direct savings were anticipated. However, there was a possibility for indirect Medicare savings because the behavioral and other incentive changes inherent in the gainsharing model had the potential to change physician behavior in ways that might reduce Medicare's outlays per episode through either reduced physician Part B charges, reduction in postacute care, reductions in testing, and potentially other sources. Our findings found little evidence that this occurred.

For the same DRGs, CAMC physicians' Medicare payments grew $171.69 less than those of physicians at the comparison sites. These savings, however, did not spill over into any of CAMC's four total episode payment measures (which did not decrease). Although neither of these two initiatives would be expected to have a direct effect on physician Part B billing, some impact on overall episode cost might have resulted, perhaps if the CAMC intervention had operated for more than a single year.

We did not find a significant impact of the Gainsharing demonstration on BIMC's per-episode expenditures. We found what turned out to be a temporary significant impact of the first year of the Gainsharing demonstration on BIMC's per-episode inpatient physician payments. The 2D coefficients for BIMC's four episode payment measures all had negative signs, but only one was statistically significant (total episode payments during PY 1). Despite BIMC's self-reported emphasis on reducing LOS as a source for reducing internal costs, BIMC's average LOS were actually slightly higher in PY 3 (possibly due to a slightly increased complex casemix) whereas the average LOS for the comparison hospitals continually fell during the three years of the demonstration. This suggests that the internal cost savings were either driven by factors other than changes in the LOS or changes in physician billing behavior. At BIMC, the sources of internal cost savings did not translate to detectable savings for the Medicare program (which had been projected but not guaranteed by BIMC in its initial demonstration application).

Quality of Care Metrics: For CAMC, our results indicate only a small and statistically insignificant impact of the demonstration on the quality indicators measured. This is likely the result of convergence in quality across hospitals given the national emphasis on quality improvement over the past two decades, as well as the difficulty in detecting demonstration impacts with only 1 year of performance.

For BIMC, we found no evidence that the Gainsharing demonstration had any statistically significantly negative effects on the quality of care that was received at BIMC over the period of the demonstration. We did note however that there were some trends in selected BIMC measures that appeared in some performance years. For example, BIMC had an increased mortality rate of 23 percent between the base year and year 3; performance for the comparison group remained nearly constant. We also noted some trends in the descriptive findings (particularly increases in 30 and 90 day mortality rates in performance year 2) that indicated the Gainsharing Demonstration at BIMC did not result in improved quality of care.

Patient Satisfaction: The analyses of HCAHPS beneficiary satisfaction survey findings for the two Gainsharing demonstration hospitals and comparison sites show that all these hospitals performed at similar levels during the performance periods. On the basis of these results, we were not able to detect any major impacts of the Gainsharing demonstration on patient satisfaction. In the case of BIMC, the demonstration site with 3 years of performance data, almost all the patient satisfaction measures are trending in the direction of improvement.

Referral Patterns: As part of the demonstration, physicians practicing at demonstration hospitals have an increased financial incentive to avoid potentially high-cost admissions. We analyzed admitting patterns for physicians admitting at both demonstration and comparison hospitals to determine whether physician referral patterns changed substantially as a result of the Gainsharing demonstration. The descriptive results indicate little, if any, demonstration impact

on most referral performance measures. In our multivariate analyses, we found that the direction of the demonstration impact on severity of admissions was different for BIMC and CAMC. Even though the direction of the demonstration impact differed, the results indicate that neither BIMC nor CAMC avoided major or extreme admissions. The results also indicate that the demonstration had no impact on IPPS transfers. On the basis of these analyses, there is no indication of adverse impacts of the demonstration related to changes in physician referral patterns.

REFERENCES

Coffey, R., Milenkovic, M., and Andrews, R.M.: The case for the present-on-admission (POA) indicator. HCUP Methods Series. Report No. 2006-01. Rockville, Md. Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services, June 2006. Available at http://www.hcup-us.ahrq.gov/reports/methods.jsp.

Cromwell, J., and Adamache, W.O.: Rates and Savings Report—Final Design Report. CMS Contract No. 500-92-0013. Research Triangle Park, N.C. Research Triangle Institute, 2004.

Cromwell, J., Dayhoff, D.A., McCall, N.T., et al.: Medicare Participating Heart Bypass Center Demonstration. Health Care Financing Administration Contract No. 500-92-0013. Research Triangle Park, N.C. Research Triangle Institute, 1998.

Cromwell, J., Dayhoff, D.A., and Thoumaian, A.H.: Cost savings and physician responses to global bundled payments for Medicare heart bypass surgery. Health Care Financ. Rev. 19(1):41-57, Fall 1997.

Kautter, J., Pope, G. C., Trisolini, M., et al.: Medicare Physician Group Practice Demonstration design: Quality and efficiency pay-for-performance. Health Care Financ. Rev. 29(1):15-29, Fall 2007.

Leitman, I.M., Levin, R., Lipp, M., et al.: Quality and financial outcomes from gainsharing for inpatient admissions: A three year experience. J. Hosp. Med. 5(9):501-507, Nov.-Dec. 2010.

Marcoux, V.M.M.: Why healthcare fraud and abuse laws should allow appropriate hospital gainsharing. Ala. Law Rev. 59(1):539-558, Winter 2007.

This page left intentionally blank.

APPENDIX TABLES

Appendix Table 6-1
Base year observed mortality and population at risk for selected conditions: BIMC and its comparison group

Hospital	AMI	CHF	Stroke	GI Hemorrhage	Hip Fracture	Pneumonia
Base Year						
Beth Israel Medical Center						
Number of occurrences	14	19	19	5	2	21
Population at risk	190	474	162	191	131	359
Rate per 1,000	73.68	40.08	117.28	26.18	15.27	58.50
Comparison Hospitals						
Number of occurrences	141	149	168	53	33	157
Population at risk	1,282	3,469	1,408	1,728	1,000	2,291
Rate per 1,000	109.98	42.95	119.32	30.67	33.00	68.53
Year 3						
Beth Israel Medical Center						
Number of occurrences	24	15	22	11	8	14
Population at risk	166	385	192	210	127	288
Rate per 1,000	144.58	38.96	114.58	52.38	62.99	48.61
Comparison Hospitals						
Number of occurrences	124	111	201	55	31	114
Population at risk	1,213	3,143	1,425	1,745	951	1,677
Rate per 1,000	102.23	35.32	141.05	31.52	32.60	67.98

SOURCE: 2007–2011 Medicare inpatient prospective payment system (IPPS) claims.

Appendix Table 6-2
Base year observed mortality and population at risk for selected conditions: BIMC and its comparison group

Hospital	Death in low-mortality DRGs	Pressure Ulcer	Death among surgical patients	Iatrogenic pneumothorax	Central venous catheter-related bloodstream infections	Postoperative hip fracture	Postoperative hemorrhage or hematoma	Postoperative physiologic and metabolic derangements	Postoperative respiratory failure	Postoperative pulmonary embolism or deep vein thrombosis	Postoperative sepsis	Postoperative wound dehiscence	Accidental puncture or laceration
Base Year													
Beth Israel Medical Center													
Number of occurrences	3	23	6	4	7	0	2	0	7	12	1	1	11
Population at risk	859	4,513	55	8,414	7,219	1,339	2,021	923	695	2,006	266	282	8,716
Rater per 10,000	34.92	50.96	1090.91	4.75	9.70	0.00	9.90	0.00	100.72	59.82	37.59	35.46	12.62
Comparison Hospitals													
Number of occurrences	7	313	117	37	38	1	50	3	74	192	30	4	132
Population at risk	5,429	31,882	710	62,384	51,034	14,414	19,977	10,911	6,784	19,823	3,288	2,872	65,609
Rater per 10,000	12.89	98.17	1647.89	5.93	7.45	0.69	25.03	2.75	109.08	96.86	91.24	13.93	20.12
Year 3													
Beth Israel Medical Center													
Number of occurrences	1	15	15	2	3	0	5	0	2	13	1	0	15
Population at risk	886	4,709	88	9,313	7,413	1,405	2,160	940	668	2,142	320	270	9,616
Rater per 10,000	11.29	31.85	1704.55	2.15	4.05	0.00	23.15	0.00	29.94	60.69	31.25	0.00	15.60
Comparison Hospitals													
Number of occurrences	1	61	113	41	17	0	45	6	97	208	48	3	156
Population at risk	5,291	30,268	835	62,811	50,060	13,466	20,021	11,620	7,665	19,868	3,811	2,521	66,296
Rater per 10,000	1.89	20.15	1353.29	6.53	3.40	0.00	22.48	5.16	126.55	104.69	125.95	11.90	23.53

SOURCE: 2007–2011 Medicare inpatient prospective payment system (IPPS) claims.