September 2010

Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Massachusetts General Hospital and Massachusetts General Physicians Organization (MGH)

Final Report

Prepared for

David Bott, PhD
Centers for Medicare & Medicaid Services
Office of Research, Development, and Information
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Nancy McCall, ScD
Jerry Cromwell, PhD
Carol Urato, MA
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0207964.025.000.001

EVALUATION OF MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION: MASSACHUSETTS GENERAL HOSPITAL AND MASSACHUSETTS GENERAL PHYSICIANS ORGANIZATION (MGH)

by Nancy McCall, ScD
Jerry Cromwell, PhD
Carol Urato, MA

Federal Project Officer: David Bott, Ph.D.

RTI International*

CMS Contract No. 500-00-0024 TO#25

September 2010

This project was funded by the Centers for Medicare & Medicaid Services under contract no 500-00-0024 TO#25. 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 Figures

List of Tables

EXECUTIVE SUMMARY

The purpose of this report is to present the findings from RTI International's evaluation of the Massachusetts General Hospital and the Massachusetts General Physicians Organization (MGH) Care Management Program (CMP) operated under the Center for Medicare & Medicaid Services' (CMS) Care Management for High Cost Beneficiaries (CMHCB) demonstration. Founded in 1811, MGH is the third oldest general hospital in the United States and the oldest and second largest hospital in New England. The 900-bed facility is also the original and largest teaching hospital of Harvard Medical School and one of the founding members of Partners HealthCare (Partners), an integrated health care system in Boston, Massachusetts, established in 1994. The system is composed of two academic medical centers, community hospitals, specialty hospitals, community health centers, a physician network, home health and long-term care services, and other health-related entities. MGH's mission is to provide high-quality health care; advance care through innovative research and education; and to improve the health and well-being of the diverse communities it serves.

MGH's CMHCB demonstration program involves providing practice-based care management (PBCM) services to high-cost Medicare FFS beneficiaries. Case managers, who are assigned to each MGH physician office, develop relationships with program participants to provide support across the continuum of care. The Massachusetts General Physicians Organization (MGPO), the largest multi-specialty group practice in New England, provides the overall administration and underlying structure in delivering integrated care management services under the CMP. Case managers provide patient education and connect patients with resources to address medical and psychosocial needs to help prevent acute exacerbations of disease and associated inpatient admissions and emergency room visits. The program also includes components to address mental health issues, evaluate complex pharmaceutical regimens, and support end-of-life decision making.

In addition to improving the quality of care and outcomes for Medicare beneficiaries, MGH's CMP aims to improve the quality of work life of primary care physicians and ultimately attract more physicians to the field of primary care. It is one of several initiatives in development at MGH to improve the challenging work life of primary care physicians. Ultimately, these initiatives are part of a larger vision for Partners to restructure the model for primary care practice characterized by high patient and physician satisfaction, work flow and process improvement, and the delivery of evidence-based care.

The principal objective of the CMHCB demonstration is to test a pay-for-performance contracting model and new intervention strategies for Medicare fee-for-service (FFS) beneficiaries, who are high cost and/or who have complex chronic conditions, with the goals of reducing future costs, improving quality of care and quality of life, and improving beneficiary and provider satisfaction. The desired outcomes include a reduction in unnecessary emergency room visits and hospitalizations, improvement in evidence-based care, and avoidance of acute exacerbations and complications. In addition, this demonstration provided the opportunity to evaluate the success of the "fee at risk" contracting model, a relatively new pay-for-performance model, for CMS. This model provided MGH's CMP with flexibility in its operations and strong incentives to keep evolving toward the outreach and intervention strategies that are the most effective in improving population-based outcomes.

The overall design of the CMHCB demonstration follows an intent-to-treat (ITT) model, and like the other demonstration programs, MGH's CMP was held at risk for its monthly management fees based on the performance of the full population of eligible beneficiaries assigned to its intervention group and as compared with all eligible beneficiaries assigned to its comparison group. Beneficiary participation in the CMHCB demonstration was voluntary and did not change the scope, duration, or amount of Medicare FFS benefits received. All Medicare FFS benefits continued to be covered, administered, and paid for by the traditional Medicare FFS program. Beneficiaries did not pay any charge to receive CMHCB program services.

Our evaluation focuses upon three broad domains of inquiry:

Implementation. To what extent was MGH able to implement its program?

Reach. How well did MGH's CMP engage its intended audiences?

Effectiveness. To what degree did MGH's CMP improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings?

Organizing the evaluation into these areas focuses our work on CMS's policy needs as it considers the future of population-based care management programs or other interventions in Medicare structured as pay-for-performance initiatives. We use both qualitative and quantitative research methods to address a comprehensive set of research questions within these three broad domains of inquiry.

E.1 Scope of Implementation

MGH launched its program on August 1, 2006. MGH worked with its CMS project officer and analysts from RTI and Actuarial Research Corporation (ARC) to develop a method for selecting the starting or original population for its CMP. Inclusion criteria for eligibility included:

- Medicare FFS beneficiaries with a primary residence in one of five designated counties including Boston, Massachusetts, and surrounding areas, and a high level of disease severity as indicated by Hierarchical Condition Categories (HCC) scores and high health care costs based on Medicare claims filed during calendar year 2005. Beneficiaries with HCC risk scores >=2.0 and annual costs of at least $2,000 or HCC risk scores >=3.0 and a minimum of $1,000 annual medical costs are eligible for the MGH's CMP.

- Beneficiaries who fulfilled the loyalty criteria for MGH (i.e., two visits to MGH physicians for a selected group of outpatient and emergency department procedures identified by CPT code, a majority of inpatient visits to MGH hospitals, or no inpatient visits between January 1, 2005 and December 31, 2005 were eligible). Selected CPT codes can be found in the ARC memo dated August 7, 2006.

The population was further restricted using the following exclusion criteria: resident of a skilled nursing facility or nursing home,[1] recipient of hospice care, or receiving the end-stage renal disease (ESRD) benefit, a history of dialysis treatment, enrolled in a Medicare Advantage plan or have Medicare as a secondary payer, or lack Medicare Part A or Part B coverage as of July 1, 2006.

After development of the original intervention group, it was identified that eligibility for the intervention and comparison groups was not determined at the same time. For the MGH population, the intervention group eligibility was determined on July 1st 2006 while the comparison group eligibility was determined on July 4th 2006. To correct for this 3 day difference, a historical eligibility pull was completed on September 1st 2006 for both the original intervention and comparison groups. Using this view and a new eligibility date of July 4th 2006, the intervention group was trimmed to 2,640 beneficiaries (from 2,655 - a removal of 15 beneficiaries). MGH also requested the removal of 21 intervention beneficiaries as part of the 60-day review process. There were 21 beneficiaries who met the physician visit loyalty criteria as a result of their relationship with one of two current MGH/MGPO primary care physicians who were employed by another area medical center during 2005. Using these criteria, a total of 2,619 Medicare beneficiaries were assigned to MGH's CMP intervention group.

Following the development of the original intervention group criteria, MGH worked with CMS and RTI to develop specifications to select a comparison group of beneficiaries to be used in conducting the financial reconciliation and evaluation of its demonstration program. The comparison group was selected using the following eligibility criteria:

- Medicare FFS beneficiaries with a primary residence in one of the five designated counties used to select the intervention group, an area which includes Boston, Massachusetts and surrounding areas.

- Beneficiaries who were loyal to one of 307 comparison physicians as determined by the loyalty criteria used to select the intervention group. The 307 physicians were selected because they practice at one of 18 group practices associated with 1 of 4 academic medical centers that served as comparison institutions for MGH's CMP: Beth Israel Deaconess Medical Center, Boston Medical Center, New England Medical Center, and St. Elizabeth's Medical Center. Practices were selected if at least 10% of services provided were office visits.

- Received a majority of inpatient care from the comparison facilities or any of several long-term care facilities, in-hospital rehabilitation or psychiatric units that accounted for at least 1% of all claims at one of the four comparison institutions.

The exclusion criteria that were applied to the intervention group were also used to limit the comparison group. In addition, a small number of patients were also excluded because they had qualified as members of the intervention group. The comparison group was also matched to the intervention group in terms of risk and cost, by determining the cut points that divided the

[1] Residence in a skilled nursing facility or nursing home was determined using the following CPT codes: 99301–99303, 99311–99313, and 99321–99333.

intervention group into three equal-sized tertiles based on HCC scores and three tertiles for Medicare costs. These two factors were cross-classified to form nine strata, and beneficiaries eligible for the comparison group were randomly sampled to match the number of intervention beneficiaries in each stratum. As noted above, eligibility for the intervention and comparison groups was not determined at the same time. Using a September 1st 2006 view, the comparison group was trimmed to 2,755 beneficiaries (from 2,786 – a removal of 31 beneficiaries).

The same general and specific inclusion and exclusion criteria used to select the original and intervention and comparison populations were used to identify the refresh intervention and comparison populations. Medicare claims for calendar year 2006 were used to identify the refresh intervention and comparison populations. Prior to determining loyalty, MGH requested that 9 physicians in their original group be dropped from the refresh because they were no longer MGH primary care physicians. To identify comparison physicians no longer affiliated with the selected comparison networks, RTI examined the billing patterns for all 801 comparison physicians used to select the original population. A total of 55 physicians were removed from the comparison networks.

CMS limited the size of the Mass General refresh intervention group to 30% of the original starting intervention group of 2,619 beneficiaries, or 785 beneficiaries, who were randomly selected from a pool of 1,870 beneficiaries. Using the same procedure to select the original comparison group, 785 beneficiaries from 9 cost-HCC strata were randomly selected. One sampled beneficiary who had already been selected for the intervention refresh group was subsequently identified. This beneficiary was removed from the comparison group, reducing the size of the comparison refresh population to 784 beneficiaries.

Of all MGH's CMP original intervention group beneficiaries, 88% verbally consented to participate in its demonstration at some point during the intervention period, 6% refused to participate, and 6% were not contacted or were unable to be located. Of the refresh intervention beneficiaries, 84% consented to participate at some point during the 24-month period. The percent that refused to participate was lower (4%) and the percent that were not contacted or were unable to be contacted increased to 13%. MGH's CMP ended July 31, 2009 or 36 months after initiation of the original population and 24 months after the start of the refresh population.

MGH negotiated a management fee of $120 for the original and refresh intervention groups through the duration of the demonstration. Fees were paid on a monthly basis for all beneficiaries who did not opt out during the 6-month outreach period and remained eligible for the demonstration. The net savings requirements for MGH's CMP are 5% for the original cohort and 2.5% for the refresh cohort.

E.2 Overview of MGH's CMHCB Demonstration Program

MGH's CMP was a provider-based care management (PBCM) program intended to provide an enhanced level of care to a high-risk patient population through comprehensive outpatient practice-based case management. Other goals included reducing health care costs through reductions of preventable hospitalizations and emergency room visits, improving physician work life, and generated increased understanding of delivering effective PBCM programs, including the development of a satisfying and manageable role for case managers. To

achieve all of these goals, MGH's CMP was structured to facilitate communication and leverage relationships (a) between patients and case managers, (b) between patients and physicians, (c) between case managers and physicians, and (d) among case managers.

The CMHCB demonstration provided MGH with the opportunity to develop a role for nurse case managers that allowed them to support patients effectively while maintaining job satisfaction. Further, MGH intended to use the CMHCB program to learn how to deliver the appropriate breadth and depth of CMP services for individuals with multiple chronic illnesses. According to MGH, these high-cost individuals are not well served by the current health care delivery system, which provides disease-specific care that does not sufficiently address the issues faced by patients with comorbid conditions. These patients may be better served using a case management model that addresses their needs through the continuum of care.

Approximately 2.5 years prior to the initiation of the launch of the CMP, MGH starting planning for this type of a demonstration project. As part of its efforts, it conducted a pilot study of the impact of PBCM at the MGH Revere Healthcare Center. An experienced case manager was placed within the health center to provide care management support services to patients. Physicians at the practice were asked to identify patients who were most likely to be admitted to the hospital within the next 6 to 8 weeks, and those individuals were invited to participate in the PBCM program, regardless of their insurance coverage. The case manager conducted assessments of participants to identify gaps in care and served as a physician extender helping patients deal with issues such as transportation to the physician's office and access to prescription medications. The PBCM pilot required physicians to spend time initially to discuss the organization and content of the case management assessments and services. MGH also convened a series of focus groups to obtain input from physicians and other MGH clinical staff about their priorities for PBCM so MGH could include useful interventions in its CMP. A second round of focus groups was conducted with physician groups to specifically discuss how the CMP could add value to their practices. In addition, CMP leadership identified a physician champion for the CMP within each physician practice that had at least 25 or more CMP patients at the start of the project to ease further the transitions involved in the introduction of a case manager into the practice. During program implementation physician champions provided insight about the best way to incorporate case managers into the practice and encourage colleagues to take advantage of services available from the case managers.

The core element of MGH's CMP is the one-on-one relationship between patients and their practice-based case managers, supplemented by support received from the program's mental health, pharmacist, and end-of-life components. MGH's CMP is designed so that case managers become staff members of primary care physician practices. According to MGH leadership, this association with the primary care provider engenders patient trust and willingness to discuss health care and psychosocial problems with these nurses. Case managers developed relationships with patients over time through telephone calls and in-person interactions during physician office visits or at the hospital, if they are admitted for an inpatient service. Case managers also conducted visits to patient homes on an as-needed basis. Overall, case managers assessed patient needs, collaborated with physicians to develop treatment plans, educated patients about options for medical treatment and support services, facilitated patient access to services, and supported patient self-management of medical conditions.

Case managers conducted a comprehensive assessment to evaluate the unique needs of each patient. Case managers focused the assessment on issues that were relevant to each patient and evaluated medical and psychosocial problems, the resources used to address these issues, and patient needs for additional support. The tool used to conduct these assessments was developed by MGH and includes several externally validated instruments, such as questions to evaluate challenges encountered with activities of daily living (ADL). Using information collected from the assessment, case managers developed a care plan for each patient in conjunction with the primary care provider and the practice's clinical team. Case managers implemented care plans over time by addressing urgent patient issues, conducting patient education, and providing referrals to support services. Throughout the program period, case managers continued to evaluate patients as their issues and need for support evolved over time.

Case managers educated patients about resources available and lifestyle changes that could help to prevent exacerbations of disease and to prevent or delay hospitalization. Case managers reviewed self-management activities, such as getting exercise and eating a low-salt diet, during a series of calls over a week or two to help patients adopt new behaviors. Case managers also educated patients about the purpose of their medications and other treatment interventions to help increase patient adherence to care plans.

Case managers also facilitated coordination of patient care across the continuum of health care services. Case managers received paged messages when their patients were admitted to the emergency room and an email indicating an inpatient admission. Using these real-time alerts, case managers could visit their patients in the hospital and research the cause of the hospitalization to inform refinements to the patient's care plan that may prevent future inpatient stays. Following hospital discharge, case managers contacted patients to make sure that they understood and could comply with discharge plans and coordinated with home health care providers to stay informed of patient health status.

Case managers also facilitated patient access to health care resources through patient education and referrals to other hospital or community services. For example, case managers informed patients that instead of going to the emergency room if they have a health problem, they could contact the physician's office at any time and may be able to see the doctor on the day of the call. Each week case managers received a list of patients scheduled to attend a physician office visit, and case managers contacted patients prior to their scheduled physician visits to find out if they needed assistance with transportation to the office. In addition, case managers followed up with patients via telephone if they missed their appointments to determine the issues involved and to provide support needed for patients to see their physician.

In the early stages of the CMHCB demonstration, CMP leadership learned that many high-cost, complex patients have mental health issues that were not effectively addressed by the current model of health care delivery or its pilot program. As a result, the program allocated greater resources to support mental health, hiring a social worker to assess the mental health needs of CMP participants and support them in accessing psychiatric care as needed or provide treatment if appropriate.

Since many members of the CMP population have complex medication regimens, MGH enlisted the support of a pharmacist to review the appropriateness of medication regimens and

assist patients with access to medications. The pharmacist also evaluated medication regimens to identify opportunities to reduce the number of medications and to suggest alternative therapies. Lessons learned during the early stages of the demonstration helped to motivate a change in MGH's medication delivery services—MGH authorized home delivery service for any CMP patient referred by their Care Manager.

One of MGH's CMP's goals was to facilitate earlier access to end-of-life resources and create a seamless transition between hospital and home care for patients approaching the end of life. The CMP received support from a nurse who specialized in end-of-life care issues. The nurse educated case managers and physicians about how to have discussions about end-of-life issues with patients, how to support patients in developing advance directives, and provided information about the services that are available through hospice.

Once MGH had generated lists of CMP-eligible beneficiaries receiving care from each physician, the CMP medical director met with each practice to introduce the program and discuss which patients were at highest risk for acute events and should receive priority for enrollment. The medical director also met with specialty practices such as the oncology, cardiology, emergency, and orthopedics departments to explain the resources available through the program, because case managers would likely interact with these providers as they facilitated patient access to these services.

Case managers assigned to each practice met with physicians at the practices to describe the program, the skills that they bring to the physician practice, and their interest in collaborating to support patients in their efforts to manage their medical conditions. Case managers collected information from providers about how they could add value to the medical practice.

MGH invested considerable personnel resources to develop and implement its CMP. At the time of the initial site visit, the program was staffed with 11 nurse case managers who received guidance from the program leadership and support from the project manager, an administrative assistant, and a community resources specialist. Each case manager was located in a physician practice and provided case management support to a group of 180 to 220 patients who received their primary care from a provider in that group. In addition, almost all case managers also "floated" to one or two additional small physician practices, which had a relatively small panel of MGH CMP eligible patients. Responsibilities included conducting patient assessments, visiting patients who were hospitalized at MGH (when feasible), contacting patients who visited the emergency department or were recently discharged from the inpatient services, calling patients scheduled for office visits each week, following up with patients who missed office visits each week, making follow-up calls to provide case management services, and promoting the MGH CMP to physicians.

Other team members who provided support to the case managers included the CMP project manager, an administrative assistant, a community resource specialist, and a patient financial counselor who provided support for all insurance-related issues. Case managers consulted the community resource specialist in their efforts to connect patients with existing resources in the community. The community resource specialist also interacted directly with patients to address nonclinical issues such as transportation and housing needs and attended weekly meetings with the case managers to provide input in discussions about patient issues.

The program leadership provided guidance to case managers, managed the program, and marketed the program throughout MGH. The clinical team leader provided oversight and supervision of case managers, represented the case managers in discussions of program development, and coordinated orientation and staff development for case management staff. The medical director provided oversight and day-to-day management of MGH's CMP, acted as primary spokesperson and project champion among physicians, contributed to the development of program interventions, served as a consultant and clinical resource to case managers, assisted in evaluating medical problems with the case managers and making medical decisions, and saw CMP patients as a part of his practice.

MGH developed a series of clinical dashboards using data from the MGH electronic medical record (EMR), claims data, and its enrollment tracking database. The dashboards allowed MGH to examine trends in health care utilization and outcomes, overall and by enrollment status, physician practice, and/or case manager, as well as activities of its case managers. Examples of indicators that are monitored regularly are: number of assessments completed within 90 days, number of referrals or interventions conducted, number of advanced directives in place, number of participants screened for depression, number of participants with a positive screen for depression referred to mental health, and the top 10 discharge diagnoses.

MGH enlisted physician support to help ensure the success of its CMP in providing high-quality care to patients. Physicians were asked to conduct the following activities: encourage beneficiaries to participate in the program and enroll them in the program when possible; collaborate with case managers to review initial assessment findings and develop care plans for each patient; inform case managers about patient events and refinements to patient care plans during the demonstration period; and discuss advance directives with enrolled patients.

MGH provided physicians with a $150 financial incentive per patient in Year 1 and $50 in Years 2 and 3 to help cover the cost of physician time for these activities. Two hundred physicians from the 19 MGH primary care medical practices that treat adult patients served as the primary care physician to at least 1 and as many as 59 patients eligible for MGH's CMP. MGH physicians received information about the CMP from a variety of sources, including the program's medical director, the MGH electronic newsletter, and case managers assigned to each practice.

Early Experience. During the first 7 months of the demonstration period, MGH's CMP staff reportedly learned a lot about the characteristics of the intervention population and the delivery of case management services for a population of sick patients with complex medical and psychosocial needs. They expected that its program would have the greatest impact by preventing acute health care events among beneficiaries who were initially not having significant health issues; however, case managers found that they spent a lot of time dealing with urgent issues for patients who "spiraled out of control." Although initially some patients were skeptical about the MGH CMP, overall, patients quickly formed relationships with case managers, including several who requested daily contact with their case managers to help them with their numerous issues.

MGH observed that the program model appeared to work better in larger practices, where the impact of the program was more easily observed among a larger panel of patients. Further,

small physician practices were more likely to have procedures that are firmly in place, and therefore encountered more challenges in integrating the case managers into their practices.

MGH encountered several challenges in implementing its CMP, including ramping up the program in a short time frame, effectively supporting patients residing in skilled nursing facilities or other institutions, and balancing the breadth and depth of support provided to each participant.

MGH reported that it had limited time to develop the infrastructure needed to deliver its program. For example, MGH's information systems were not fully set up to manage program activities on August 1, 2006, when the demonstration period began. In addition, MGH began hiring case managers in June 2006 and did not finish the hiring process until December 2006. Consequently, beneficiary enrollment took longer and was less successful at practices that did not have a case manager during the initial months of the demonstration.

Although MGH used eligibility criteria designed to remove institutionalized beneficiaries from its intervention population, a small proportion of the intervention population resided in skilled nursing homes or other facilities. MGH speculated that these patients were included as a result of the lag between the time period represented by the claims data used to select intervention patients and the program's launch date.

Case managers were challenged in balancing the depth and breadth of support provided to each patient. Program management staff monitored the number, length, and frequency of patient interactions to evaluate the most effective and efficient way to deliver case management services to this sick population.

Program Changes. A number of changes to MGH's CMP occurred as the program matured. Noteworthy changes include the following:

- **Termination of the contract with Health Dialog and bringing assessment and data collection capabilities in-house**: Given that the contract was very expensive, communication with Health Dialog was very difficult due to HIPAA, and Health Dialog's limited experience with practice-based case management with a medically complex FFS population, the CMP leadership made a decision to terminate its contract with Health Dialog and bring enrollment and data system development capabilities in-house.

- **Revision of its assessment tool**. The program shifted from using Health Dialog applications to using the hospital-based Medical Information Data System (MIDAS) for recording patient assessment data. During the switch, the entire assessment process was reviewed by program leadership and major changes to the process were implemented. The single extensive assessment that was conducted initially was broken down into six different assessment modules that could be conducted depending on participants' needs: functional assessment, mental health, advance care planning, transportation, pharmacy, and post hospital episode.

- **Program staffing changes:**
  - addition of a designated case manager position to work specifically on post discharge assessments to enhance transitional care monitoring;
  - creation of a data analytics team to develop and strengthen program's reporting capabilities;
  - expansion of the role of the community resource specialist to provide patient support and referral for non-clinical services; and
  - expansion of the mental health component of the program by adding a mental health team director, clinical social worker, two psychiatric social workers, and a forensic clinical specialist (M.D./J.D.), who follows highly complex patients with issues such as legal issues, guardianship and substance abuse.
- **Data System Enhancements:**
  - integration of a clickable CMP icon into the MGH-wide Longitudinal Medical Record (LMR) system so every provider could immediately identify Medicare beneficiaries in the program and be able to get in touch with their patient's case manager by page or e-mail;
  - development of a data portal for primary care physicians that would be able to view their inpatient census and program panel; and
  - implementation of a dual computer monitor system so every case manager could have simultaneous access to both data systems from their desks.

E.3 Key Findings

In this section, we present key findings based upon the 36 months of MGH's CMP operations with its original population and 24 months with its refresh population. Our findings are based on the experience of approximately 6,800 ill Medicare beneficiaries split across 4 groups for analysis purposes (original and refresh intervention and comparison groups) limiting statistical power somewhat within the substantially smaller refresh population (only 30% the size of the original population) to detect differences. Eight findings on participation, intensity of engagement in MGH's program, beneficiary satisfaction and experience with care, provider satisfaction, clinical quality, acute care utilization, health outcomes, and financial outcomes have important policy implications for CMS and future disease management or care coordination efforts among Medicare fee-for-service (FFS) beneficiaries. The CMHCB demonstration program holds MGH financially responsible for financial savings but does not hold MGH financially responsible for quality of care improvements.

Key Finding #1: MGH's CMP achieved a high participation level that reached broadly across its intervention population in terms of beneficiary demographic characteristics, prior health status and health care costs, and health status measured during the early months of its demonstration.

The MGH CMP had an ambitious goal of gaining participation from 100% of its original population beneficiaries. It was successful in recruiting 88% of its original population beneficiaries and 84% of its refresh population beneficiaries. We found few statistically significant differences between participants and nonparticipants in either the original or refresh intervention populations. In multivariate modeling of factors that predict likelihood of participation, we had low explanatory power suggesting that MGH's CMP was able to recruit broadly across its intervention population as no particular set of factors that we tested strongly predicted participation. The substantially smaller sample size for the refresh population also limited our ability to detect participation factors.

Key Finding #2: MGH's CMP successfully targeted beneficiaries with high rates of acute care utilization.

A cornerstone of MGH's CMP was the one-on-one relationship between participants and their practice-based case managers. Telephone calls were the most dominant form of contact. In our multivariate regression modeling of likelihood of being in a high contact versus low contact group for both the original and refresh populations, we found hospitalizations during the demonstration period to be very strong predictors of contact. A major focus of MGH's CMP was to prevent hospitalizations or re-hospitalizations. These findings reveal that MGH's CMP was successful in their effort to contact beneficiaries who had been hospitalized and at high risk of re-hospitalization albeit MGH was unsuccessful at reducing rates of readmission.

Key Finding #3: MGH's CMP improved beneficiary reported satisfaction with helpfulness of discussions and communication with their health care team. MGH's CMP also improved physical functioning. MGH's CMP did not improve beneficiary reported ability to cope with their chronic condition nor improve self-efficacy or self-care activities or mental health functioning.

The beneficiary survey was designed to obtain assessments directly from beneficiaries about key outcomes of beneficiary experience of care, self-management, and physical and mental function. We asked beneficiaries about the extent to which their health care providers helped them to cope with their chronic condition. We supplemented this item with questions related to two key components of MGH's CMP intervention: helpfulness of discussions and quality of communication with their health care team. In addition, the survey instrument collected information about beneficiary self-care frequency and self-efficacy related to medications, diet, and exercise and Clinician and Group Adult Primary Care Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

Among the 19 outcomes covered by the survey, we found three statistically significant ANCOVA-adjusted intervention effects, two in the experience and satisfaction with care domain and one in the physical and mental health function domain. Survey results indicate that members

of MGH's CMP intervention group were more satisfied with the discussion of their treatment choices and rated their communication with heath providers higher than MGH's CMP beneficiaries in the comparison group. These achievements, however, were not translated into any improvements in self-efficacy or in self-care activities. In addition, MGH's CMP beneficiaries in the intervention group reported significantly higher Veterans RAND-12 Physical Health Composite (PHC) scores than those in the comparison group, suggesting that the intervention improved physical functioning of participating Medicare beneficiaries. No other statistically significant outcomes were found in the physical and mental health function domain.

Key Finding #4: MGH's CMP improved primary care provider (PCP) assessment of the quality of medical practice and quality of care for their patients.

In addition to improving the quality of care and outcomes for Medicare beneficiaries, MGH's CMP aimed to improve the quality of primary care physicians' work lifeand ultimately attract more physicians to the field of primary care. It is one of several initiatives in development at MGH to improve the challenging work life of primary care physicians. Ultimately, these initiatives are part of a larger vision for Partners HealthCare to restructure the practice model for primary care practice characterized by high patient and physician satisfaction, work flow and process improvement, and the delivery of evidence-based care.

RTI conducted two site visits to MGH's CMP and spoke with a small number of primary care physicians during each site visit to gauge their assessment of satisfaction with the demonstration program. At the time of the first site visit, a small number of physicians expressed concerns about the program. For example, they had questions about whether CMP patients would divert services from other patients in their practices. And, some physicians did not have a full understanding of the role of the case managers. However, as physicians gained experience working with the case managers, the most common concern they voiced was frustration about their inability to include additional patients in the program. One provider noted that for each patient eligible for the program, there are two additional patients in the practice who could benefit from such case management support.

At the time of the second site visit, physicians gathered for the focus group reported great overall satisfaction with the CMP. The following first three quotes highlight the essence of their satisfaction with MGH's CMP with the fourth quote expressing a widely held view among the interviewed physicians:

- *"The program "wraps its arms" around the most difficult and complex patients."*
- *"The program signifies a move towards a true medical home model-it is a team of providers. The program does what every PCP needs to be doing but cannot do anymore because of the medicine practice and reimbursement realities and primary care provider shortages."*
- *"The program has done a remarkable job in training and cultivating case managers who are very good at breaking barriers and making it work for the most difficult patients."*

- *"We do not want the program to end—it is very valuable! Once the program is gone, participants will become "frequent flyers" in the emergency department and hospital."*

Key Finding #5: MGH's CMP did not improve the rate of compliance in four quality-of-care process measures.

We have defined quality improvement for this evaluation as an increase in the rate of receipt of claims-derived, evidence-based quality-of-care measures. We selected three measures appropriate for different populations of elderly beneficiaries: influenza vaccine for all beneficiaries; low-density lipoprotein cholesterol (LDL-C) testing for beneficiaries with diabetes or ischemic vascular disease (IVD); and rate of annual HbA1c testing for beneficiaries with diabetes. During the demonstration, we find no evidence of systematic improvement in quality of care among the intervention beneficiaries.

For the original and refresh populations and within both the intervention and comparison groups, there were high rates of "always being compliant" in receipt of three of the process-of-care measures with the noted exception of influenza vaccination. Thus, there was limited room for improvement in either population. During the last year of the demonstration, only 10% of intervention beneficiaries with diabetes were not compliant in receipt of annual HbA1c testing, and 20% and 24% of intervention beneficiaries with diabetes or IVD, respectively, were not compliant in receipt of annual LDL-C testing. For influenza vaccination, the original and refresh intervention groups' rate increased during the demonstration but increased less than the rates for the comparison groups. However, baseline rates for the original and refresh intervention groups were far higher than the comparison groups' rate.

This finding is not unexpected. MGH's CMP leadership felt that there was a very good reason standard quality measures are not part of this demonstration's outcomes. Program leaders reported that such measures are not good quality of care indicators for the program's population. For their group of patients, something like testing for HbA_{1c} levels is not a relevant measure of how well the program is managing the care of their very sick and complex patients. The CMP leadership and MGH leadership believe that ER use and acute hospitalizations are in essence the measures that need to be used. In addition to these outcomes, other types of measures related to care coordination that they believe are highly relevant to this population include how fast case managers follow up on patient-initiated calls, can appointments be consolidated so frail beneficiaries do not have to drive to the hospital 3 times a week, and so forth.

Key Finding #6: MGH's CMP was successful reducing the rate of increase in acute care hospitalizations and ER visits but not 90-day readmissions. MGH's CMP did not impact use of the Medicare hospice benefit.

During the course of MGH's CMP, we generally observe increasing rates of all-cause and ambulatory care sensitive condition (ACSC) hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations. However, MGH's CMP was successful at substantially reducing the rate of increase in all-cause and ACSC hospitalizations and ER visits among the original and refresh intervention beneficiaries. We observe no statistically significant difference in the rate of readmission

between the intervention and comparison original and refresh populations. During the last 12 months of the demonstration, rates of growth in acute care utilization narrowed between the original intervention and comparison beneficiaries; yet the lower rates of growth among intervention beneficiaries remain statistically significant.

One component of MGH's CMP was end-of-life planning including advance directives and use of hospice. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their use of the Medicare hospice benefit or in mean or median number of days of hospice.

Key Finding #7: MGH's CMP was successful at reducing the mortality rate within the intervention group of beneficiaries.

Another key outcome metric is mortality. Over the 36-month demonstration period for the original population, 28% of beneficiaries in the intervention group died and 30% of beneficiaries in the comparison group died ($p=0.19$). MGH CMP's mortality rate was statistically lower than in the comparison group after adjusting for differences in baseline characteristics using Cox proportional hazard modeling ($p<0.05$). For the refresh population, we do observe a lower rate of mortality in the intervention group. During the 24-month refresh period, 16% of beneficiaries in the intervention group died while 20% of beneficiaries in the comparison group died ($p=0.04$). Controlling for baseline characteristics in the multivariate modeling narrows the statistical significance of this difference, yet the mortality difference remains statistically significant at the $p=0.05$ level for the refresh intervention group.

Key Finding #8: MGH's CMP achieved substantial, statistically significant savings. The Medicare program's return on investment (ROI) was 2.65 for MGH's original intervention group and 3.35 for MGH's refresh intervention group.

PBPM costs showed considerable variability because of the nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. Cost trends in MGH's market area also showed strong positive growth in both groups. Nevertheless, substantial, statistically significant, savings were found for the intervention in the original population. Relative costs (or gross savings) rose -$288 slower in the original intervention group (12.1% of monthly comparison costs); yet needed to exceed just 6.8% to be considered statistically significant at the 95% confidence level. For every dollar invested in MGH's CMP management fees, Medicare received $2.65 in savings on beneficiary health care services.

If anything, MGH's CMP performed even better with its refresh population. Gross savings averaged -$355 (15.8% of comparison monthly costs). Based on an average monthly management fee of $120 paid on 84% of participating intervention eligibles, Medicare's return on investment was 3.35. For every management dollar spent, Medicare received $3.35 in return in the form of lower cost increases.

A few material imbalances were found between the intervention and comparison groups in the base period. However, controlling for imbalances had little effect on our overall final conclusion of statistically significant savings.

The CMHCB demonstration program at MGH exhibited strong regression-to-the-mean effects while average beneficiary costs also were increasing rapidly in the greater Boston area. Intervention group costs continued to rise because minor reductions in costs in the very high cost group were more than offset by larger increases among the greater majority of beneficiaries. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. Regression-to-the-mean effects make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater is the potential for regression-to-the-mean, the greater is the challenge to identify lower cost, lower utilizing beneficiaries initially to avoid expensive hospitalizations in the near future. Nevertheless, it appears that MGH's CMP staff was able to work successfully across the full spectrum of low-to-high cost beneficiaries, resulting in a financially successful outcome.

The cost analyses presented in this report differ from those conducted for financial reconciliation by ARC under contract to CMS. ARC determined savings based on the demonstration's terms and conditions negotiated between CMS and MGH. RTI's estimation of savings, detailed subsequently, differs in that

- differences in savings rates between intervention and comparison groups are first determined at the beneficiary level and are then tested using statistical confidence intervals,
- beneficiary PBPM costs are not trimmed using a 1% outlier dollar threshold, and
- both base year and demonstration period PBPM costs are weighted by each beneficiary's fraction of eligible days during the demonstration period.

A more detailed explanation and justification for these differences is provided in ***Chapter 7, Section 7.3***.

E.4 Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that MGH's CMP had success at improving primary care providers' satisfaction with their quality of work life and improving some measures of beneficiary experience with care and functional status. We also find that MGH's CMP had substantial success reducing acute care hospitalizations and ER visits and mortality, and achieving substantial cost savings. We find these latter successes within both the original and refresh intervention groups. The financial savings is particularly noteworthy given the relatively small sample sizes and regression to the mean effects. PBPM costs showed considerable variability because of the nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. With only roughly 2,600 original and 800 refresh intervention beneficiaries and 2,700 original and 800 refresh comparison beneficiaries, we had limited our power to detect significant savings in the refresh population in particular. Gross savings had to be at least 6.8% in the original intervention population and 13.7% or more in the refresh intervention population to be considered significant at the 95% confidence level.

What might explain the observed success in MGH's demonstration program? Two explanations may be the depth of institutional support to (1) develop an MGH-specific program,

and (2) to fully integrate the CMP into MGH's health care system. Based upon interviews with senior MGH and CMP leadership, it was noted that from the beginning the CMP had the complete backing from the Board of Trustees and hospital and physician leadership. As one example, MGH physicians received communications about the program directly from the hospital's and the physician organization's leadership. And, MGH invested considerable time and resources in their CMHCB program development before launching the program in Eastern Massachusetts. Pre-launch activities included conducting a pilot study of practice-based care management (PBCM), conducting focus groups to inform the design of the Care Management Program, and hiring and training staff to implement the program.

Approximately 2.5 years prior to the initiation of the launch of the CMP, MGH began the planning for this type of demonstration. As one example, MGH conducted a pilot study of the impact of PBCM at the MGH Revere Healthcare Center. An experienced case manager was placed within the health center to provide care management support services to patients. Physicians at the practice were asked to identify patients who were most likely to be admitted to the hospital within the next 6 to 8 weeks, and those individuals were invited to participate in the PBCM program, regardless of their insurance coverage. The case manager conducted assessments of participants to identify gaps in care and served as a physician extender helping patients deal with issues such as transportation to the physician's office and access to prescription medications. The PBCM pilot required physicians to spend time initially to discuss the organization and content of the case management assessments and services. An evaluation of the program showed that physicians were very satisfied, referring to the case manager as a "fairy godmother."

MGH also convened a series of focus groups, referred to as capstone groups, to obtain input from physicians and other MGH clinical staff about their priorities for PBCM so MGH could include useful interventions in its CMP. Initially, MGH conducted multidisciplinary capstone group sessions with representatives from social work, mental health, and the MGH case management department in addition to leadership from primary care practices. A second round of focus groups was conducted with physician groups to specifically discuss how the CMP could add value to their practices.

In addition to providing input about the design of the CMP, the capstone groups provided an opportunity to obtain physician buy-in to the PBCM program. Despite the fact that some physician practices already had case managers, CMP management observed that most physician practices were apprehensive about changes such as the introduction of new staff into their practice. CMP leadership used a tailored approach to discuss the project with each practice, offering positive anecdotes from the PBCM pilot project as appropriate. In addition, CMP leadership identified a physician champion for the CMP within each physician practice that had at least 25 or more CMP patients at the start if the project to further ease the transitions involved in the introduction of a case manager into the practice. During program implementation physician champions provided insight about the best way to incorporate case managers into the practice and encourage colleagues to take advantage of services available from the case managers.

At the time of the program launch, strong integration support from MGH leadership afforded the case managers physical entry into the primary care practice settings whereby the

case managers were co-located with the primary care physicians ultimately becoming a part of the beneficiaries' primary health care teams. At the time of the first site visit, a small number of physicians expressed concerns about the program. However, as physicians gained experience working with the case managers, the most common concern they voiced was frustration about their inability to include additional patients in the program. At the time of the second site visit, physicians gathered for the focus group reported great overall satisfaction with the CMP. Acquiring buy-in from participating physician practices was viewed as very important. However, it was recognized early on that buy-in was needed on all levels. There was some concern among practice-based nurses, particularly at smaller practices, that there would be a duplication of effort. To obtain buy-in from the nurses, the CMP case managers spent time working with the practice-based nurses to educate them that the goal of the program was to augment and not to replicate their efforts.

With leadership support for integration within the MGH health system, the CMP was able to marshal a wide range of MGH internal resources to more fully develop particular aspects of their program that were tailored to the needs of the MGH patient population. Specific examples included the development of a CMP-specific mental health team comprised of MGH psychiatrists and a CMP social workers to screen for and treat depression among its participants; development of a shared planning protocol with MGH discharge planning case managers; and enlistment of an MGH pharmacist to review the appropriateness of medication regimens and assist patients with access to medications.

Another critical element of integration was the use of MGH's information technology (IT) system to support CMP operations. By gaining access to MGH's existing IT system and MGH internal resources to make necessary modifications during early stages of implementation, the CMP was able to draw upon existing infrastructure and augment it to provide immediate decision management support for its case managers.

Further, MGH's IT systems span all care settings at MGH, including all MGH physician practice settings. And, according to CMP leadership, MGH patients are very loyal to MGH and receive the vast majority of their health care from the large network of MGH-affiliated providers. Thus, CMP case managers had access to real-time patient information across virtually their patients' entire continuum of care. Yet, the CMP went through several iterations of data system enhancements at considerable expense as it sought to increase usefulness of its systems for managing patient care and reducing documentation burden.

During our site visits, CMP leadership opined that creating a similar program may require a large setting like a teaching hospital where the information technology component and the related underlying infrastructure are in place before program implementation. CMP leadership felt that a care management program such as theirs may not work well in individual practices because of resource constraints.

As one specific example of the value of integration of the CMP with the MGH electronic medical record (EMR), MGH's EMR was modified to include a CMP icon to alert providers that a particular patient was a participant in the CMP. Because of the leadership effort to make the program both visible and integrated, the CMP icon generated the type of response from MGH providers that has eluded other Medicare chronic care management programs that we have

evaluated. According to CMP leadership, the day the CMP launched the MGH ER notified the CMP Project Manager that dozens of Medicare beneficiaries with CMP icons were in their emergency room. A key focus of the MGH CMP and many other chronic care management programs is to prevent acute care hospitalizations yet many other programs that we have evaluated have been unable establish real-time notification systems with emergency rooms allowing case managers an opportunity to intervene prior to admission.

CMP integration with MGH's EMR also allowed case managers to receive weekly reports that showed which patients were scheduled for a physician visit so they could meet with patients to conduct patient interventions such as providing educational information. Case managers could also use this information to identify patients who missed their appointments so they could follow up and determine if the beneficiary needed additional support. Case managers also received timely notifications of patient admission to the emergency department or to an inpatient unit so that they could assist with transitions out of acute care and help patients avoid future exacerbations that would require acute care.

A third explanation may be elements of the management of the CMP itself. CMP leadership recognized prior to launch of its program that its population would require case managers with substantial experience in dealing with frail and medically complex patients. The CMP selected nurses with strong clinical skills, critical thinking abilities, and the ability to work independently. In addition, CMP leadership felt strongly that to be successful case managers had to learn quickly how to fit into their assigned practice setting in a way that would be helpful and valued by the physicians and their clinical and administrative staff members.

The CMP leadership organized a comprehensive orientation program to introduce the case managers to MGH, including patient resources available across the continuum of care at the hospital. Case managers met with various MGH staff members and spent time observing preceptor case managers to learn how to perform their jobs. Case managers also received training to conduct patient assessments, create comprehensive care plans, arrange for referrals to various services (e.g., transportation), and use the information systems available to support the CMP.

Due to the complexity of the CMP demonstration population, CMP leadership felt that constant and good communication between all staff within the program was essential. The CMP leadership implemented *Virtual Rounds*, regular e-mail reports that went to all staff, as a mechanism of providing feedback on a weekly basis. Case managers used *Virtual Rounds* to report on difficult patients and unnecessary admissions, and to describe both positive and negative events. *Virtual Rounds* were also used for case reviews with forms that staff filled out at the end of the week. These case reviews were then discussed with physicians in weekly face-to-face meetings. Common themes and issues from the *Virtual Rounds* were also presented at bi-monthly management meetings. The bi-monthly management meetings were used to review protocols, present resources, provide training, and identify issues and brainstorm solutions.

CMP leadership also emphasized team support and peer counseling by developing infrastructure that provided opportunities for mutual support among CMP case managers and peer counseling from the members of the mental health team as the emotional toll on staff of working with a highly frail and sick population are substantial. Such challenges include high mortality rates among program enrollees and challenges making a meaningful clinical or social

impact in very advanced cases. In focus groups conducted during both of our site visits, case managers expressed strong support for CMP leadership and a strong sense of job satisfaction.

In evaluations of other Medicare chronic care management programs, we have observed other programs that exhibited *strong program leadership* and a strong sense of job satisfaction among the case managers, yet we have not observed the same degree of *integration* of the care management program into an integrated health system and its *IT system* as we do with MGH's CMP. And, MGH's CMP beneficiaries were sufficiently concentrated in the primary care practices making placement of *full-time case managers, in general, in the practices* economically feasible. It may be that all four elements are necessary to be successful reducing acute care utilization and the cost of care for chronically ill Medicare beneficiaries.

Yet, even with the level of observed practice and IT integration MGH's CMP was not successful reducing 90-day readmissions. This is surprising given the emphasis of MGH's CMP on managing care transitions across settings. In the CMP, case managers followed a protocol for every transition between care settings. The protocols include step by step daily workflow instructions for the following transitions of care: emergency department admissions, inpatient MGH admissions, post hospital discharges to home from MGH, post hospital discharges to other facilities, and post discharge from post acute facility to home. In addition, a post-episode assessment was completed within 24-72 hours of the patient's discharge from the MGH emergency department or inpatient unit, and other acute or post-acute care facilities, if known. In addition, the patient's primary care-based case manager interfaced with the MGH inpatient case manager during the admission and prior to discharge.

This level of effort to prevent readmissions and the disappointing results suggests that broadly reducing readmissions among chronically ill Medicare FFS beneficiaries may be a far bigger challenge than has been envisioned by MGH's CMP leadership or Federal policy makers. In Phase II, MGH's CMP has been granted a demonstration waiver to allow for direct admission to a skilled nursing facility (SNF) beneficiaries who meet specific clinical criteria. It will be important to examine if the SNF waiver is a tool that can be used to reduce readmissions of patients who become clinically unstable after discharge. In Phase II, it will also be important to explore in greater detail the reasons for readmission and the degree to which the readmission is clinically related to the prior admission. Our analyses showed that ambulatory care sensitive conditions (ACSCs) account for about one-quarter of reasons for admission. It is plausible that there is a sizeable subset of beneficiaries being admitted in short spans of time for acute care services not readily amenable to reduction through case management of care transitions (e.g., hip replacement, cataract extraction, etc.).

CHAPTER 1
INTRODUCTION TO THE MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION AND THE MASSACHUSETTS GENERAL HOSPITAL (MGH) AND THE MASSACHUSETTS GENERAL PHYSICIANS ORGANIZATION (MGPO) CARE MANAGEMENT PROGRAM (CMP)

1.1 Background on the CMHCB Demonstration and Evaluation

The purpose of this report is to present the findings from RTI International's evaluation of the Massachusetts General Hospital and the Massachusetts General Physicians Organization Care Management for High Cost Beneficiaries (CMHCB) demonstration program. On July 6, 2005, the Centers for Medicare & Medicaid Services (CMS) announced the selection of six care management organizations (CMOs) to operate programs in the CMHCB demonstration:

1. Care Level Management (CLM)

2. VillageHealth (formerly known as RMS) and its Key to Better Health program (KTBH)

3. The Health Buddy® Consortium (HBC), composed of Robert Bosch Health Care (RBHC, formerly known as the Health Hero Network), the American Medical Group Association (AMGA), Bend Memorial Clinic, and Wenatchee Valley Medical Center

4. Texas Tech University Health Sciences Center (TTUHSC) and its Texas Senior Trails (TST) program

5. Montefiore Medical Center (MMC)

6. Massachusetts General Hospital and Massachusetts General Physicians Organization (MGH) and its Care Management Program (CMP)

These programs offer a variety of models, including "support programs for healthcare coordination, physician and nurse home visits, use of in-home monitoring devices, provider office electronic medical records, self-care and caregiver support, education and outreach, behavioral health care management, and transportation services" (CMS, 2005).

The principal objective of this demonstration is to test a pay-for-performance contracting model and new intervention strategies for Medicare fee-for-service (FFS) beneficiaries, who are high cost and/or who have complex chronic conditions, with the goals of reducing future costs, improving quality of care and quality of life, and improving beneficiary and provider satisfaction. The desired outcomes include a reduction in unnecessary emergency room visits and hospitalizations, improvement in evidence-based care, and avoidance of acute exacerbations and complications. In addition, this demonstration provides the opportunity to evaluate the success of the "fee at risk" contracting model, a relatively new pay-for-performance model, for CMS. This model provides the CMOs with flexibility in their operations and strong incentives to keep evolving toward the outreach and intervention strategies that are the most effective in improving population outcomes.

The overall design of the CMHCB demonstration follows an intent-to-treat (ITT) model, and the CMOs are held at risk for their monthly management fees based on the performance of the full population of eligible beneficiaries assigned to their intervention group and as compared with all eligible beneficiaries assigned to their comparison group. Beneficiary participation in the CMHCB demonstration is voluntary and does not change the scope, duration, or amount of Medicare FFS benefits received. All Medicare FFS benefits continue to be covered, administered, and paid for by the traditional Medicare FFS program. Beneficiaries do not pay any charge to receive CMHCB program services.

The CMOs receive from CMS a monthly administrative fee per participant, contingent on intervention group savings in Medicare payments being equal to fees paid to the CMO plus an additional 5% savings safety margin calculated as a percentage of its comparison group's Medicare payments. CMS developed the CMHCB initiative with considerable administrative risk as an incentive to reach assigned beneficiaries and their providers and to improve care management. To retain all of their accrued fees, the CMOs have to reduce average monthly payments by the proportion of their comparison groups' Medicare program payments that the fee comprises. In addition, to insure that savings estimates were not simply the result of random variation in estimates of claims costs, CMS required an additional 5% in savings (net savings). If the CMOs are able to achieve net savings beyond the 5% requirement, it is RTI's understanding that the shared savings provision is as follows:

1. Savings in the 0%-5% range will be paid 100% to CMS.

2. Savings in the >5%-10% range will be paid 100% to MGH.

3. Savings in the >10%-20% range will be shared equally between MGH (50%) and CMS (50%).

4. Savings of >20% will be shared between MGH (70%) and CMS (30%).

One year after the launch of each demonstration program, CMS offered all CMOs the option of supplementing their intervention and comparison populations with additional beneficiaries to offset the impact of attrition primarily due to death. This group of beneficiaries is referred to as the "refresh" population. The CMOs are at financial risk for fees received for their refresh populations plus an additional 2.5% savings.

We use the chronic care model developed by Wagner (1998) as the conceptual foundation for our evaluation because the CMHCB programs are generally provider-based care models. This chronic care model is designed to address systematic deficiencies and provides a standard framework that the area of chronic care management lacks. The model identifies six elements of a delivery system that lead to improved care for individuals with chronic conditions: the community, the health system, self-management support, delivery system design, decision support, and clinical information systems (Glasgow et al., 2001; Wagner, 2002; Wagner et al., 2001). According to the model, patients are better able to actively take part in their own care and interact productively with providers when these components are developed, leading to improved functional and clinical outcomes. Our evaluation focuses upon three broad domains of inquiry:

1. *Implementation.* To what extent were the CMOs able to implement their programs?

2. *Reach.* How well did the CMOs engage their intended audiences?

3. *Effectiveness.* To what degree were the CMOs able to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings?

Organizing the evaluation into these areas focuses our work on CMS's policy needs as it considers the future of population-based care management programs or other interventions in Medicare structured as pay-for-performance initiatives. We use both qualitative and quantitative research methods to address a comprehensive set of research questions within these three broad domains of inquiry.

RTI International was hired by CMS to be the evaluator of the CMHCB demonstration and has previously conducted and reported to CMS findings from site visits to each CMO and a beneficiary survey of each CMO's intervention and comparison populations. In general, we made two rounds of site visits to each CMO to observe program start-up and to assess CMO implementation over time. The first round of site visits was conducted at the close of the outreach period for each program, and the second round of site visits was conducted approximately 2 years later. For each site visit, data were collected through telephone interviews, in-person interviews, and secondary sources, including program monitoring reports. Two RTI evaluation team members participated in 1- to 2-day on-site visits at each CMO location.

The first site visit focused on learning about CMHCB program start-up; examining the elements of the CMHCB programs; determining the nature of the CMOs' relationship with physicians in each community; learning about ways the CMOs manage costs, quality, and beneficiary utilization of care; and obtaining information on the types of services that comprise the intervention offered. The second site visit focused on engagement of the refresh population, program evolution, program monitoring/outcomes, and implementation experience/lessons learned. During the site visits, RTI met with a small number of physicians to develop an overall impression of satisfaction and experiences with the CMHCB programs. The primary objectives of the interviews were to (1) assess physicians' awareness of the CMHCB program and (2) gauge their perceptions of the effectiveness of these programs.

RTI also conducted an assessment of beneficiary satisfaction with the CMHCB program and whether the program improved knowledge and self-management skills that led to behavioral change and improved health status among intervention beneficiaries. Program success for each of four beneficiary survey domains, satisfaction, care experience, self-management, and physical and mental health functioning, was evaluated by surveying intervention and comparison beneficiaries once at Month 18 of the intervention period. The MGH CMP's survey was conducted between January 7, 2008 and May 4, 2008. Surveying was conducted with beneficiaries from the original populations. No surveying was conducted with beneficiaries from any of the refresh populations. The findings from the beneficiary surveys were reported to CMS in RTI's Third Annual Report (Smith et al., 2008).

This final report presents evaluation findings based on the full 36 months of MGH's CMP operations with its original population and 24 months with its refresh population. We start by reporting on the degree to which MGH's CMP was able to engage its intervention population.

We measure degree of engagement in two ways: (1) participation rates and characteristics of participants; and (2) number and nature of contacts between MGH's CMP staff and participating beneficiaries from encounter data provided to RTI from MGH's CMP. We then report findings related to the effectiveness of MGH's CMP to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings.

1.2 MGH's CMHCB Demonstration Program Design Features

1.2.1 MGH Organizational Characteristics

Founded in 1811, MGH is the third oldest general hospital in the United States and the oldest and second largest hospital in New England. The 900-bed facility is also the original and largest teaching hospital of Harvard Medical School and one of the founding members of Partners HealthCare (Partners), an integrated health care system in Boston, Massachusetts, established in 1994. The system is composed of two academic medical centers, community hospitals, specialty hospitals, community health centers, a physician network, home health and long-term care services, and other health-related entities. MGH's mission is to provide high-quality health care; advance care through innovative research and education; and to improve the health and well-being of the diverse communities it serves.

MGH's CMHCB demonstration program involves providing practice-based care management (PBCM) services to high-cost Medicare FFS beneficiaries. Case managers, who are assigned to each MGH physician office, develop relationships with program participants to provide support across the continuum of care. The Massachusetts General Physicians Organization (MGPO), the largest multi-specialty group practice in New England, provides the overall administration and underlying structure in delivering integrated care management services under the CMP. Case managers provide patient education and connect patients with resources to address medical and psychosocial needs to help prevent acute exacerbations of disease and associated inpatient admissions and emergency room visits. The program also includes components to address mental health issues, evaluate complex pharmaceutical regimens, and support end-of-life decision making.

In addition to improving the quality of care and outcomes for Medicare beneficiaries, MGH's CMP aims to improve the quality of work life of primary care physicians and ultimately attract more physicians to the field of primary care. It is one of several initiatives in development at MGH to improve the challenging work life of primary care physicians. Ultimately, these initiatives are part of a larger vision for Partners to restructure the practice model for primary care practice characterized by high patient and physician satisfaction, work flow and process improvement, and the delivery of evidence-based care.

1.2.2 Market Characteristics

MGH's CMP is targeted to patients who are loyal to MGH (i.e., receive most of their care at MGH and its affiliated physician practices). MGH serves a diverse population in the city of Boston and its surrounding suburban communities. Although the majority of the population is Caucasian, there are substantial populations of African American, Asian, and Hispanic residents.

Boston also has significant socioeconomic diversity that encompasses highly affluent as well as low-income individuals.

Congruent with national trends, MGH has observed a sharp decline in the number of primary care physicians in the Boston area. For example, more than 5 years ago MGH received an average of 8 applications for each open primary care physician position, whereas more recently, only 1.5 physicians, on average, currently apply for such positions.

Since MGH has a very high inpatient census and an emergency room that is often overloaded, MGH's CMP is not viewed as a threat to the overall revenue of the hospital, despite the fact that the program could decrease the utilization of inpatient and emergency services. Rather, the program is viewed as potentially freeing up needed inpatient beds.

1.2.3 MGH Original and Refresh Intervention and Comparison Populations

MGH worked with its CMS project officer and analysts from RTI and ARC to develop a method for selecting the starting or original population for its CMP. Inclusion criteria for eligibility included:

- Medicare FFS beneficiaries with a primary residence in one of five designated Massachusetts counties (Norfolk, Suffolk, Middlesex, Essex, and Plymouth) including Boston, and a high level of disease severity as indicated by Hierarchical Condition Categories (HCC) scores and high health care costs based on Medicare claims filed during calendar year 2005. Beneficiaries with HCC risk scores >=2.0 and annual costs of at least $2,000 or HCC risk scores >=3.0 and a minimum of $1,000 annual medical costs are eligible for the MGH's CMP.

- Beneficiaries who fulfilled the loyalty criteria for MGH (i.e., two visits to MGH physicians for a selected group of outpatient and emergency department procedures identified by CPT code, a majority of inpatient visits to MGH hospitals, or no inpatient visits between January 1, 2005 and December 31, 2005 were eligible). Selected CPT codes can be found in the ARC memo dated August 7, 2006.

The population was further restricted using the following exclusion criteria: resident of a skilled nursing facility or nursing home,[2] recipient of hospice care, or receiving the end-stage renal disease (ESRD) benefit, a history of dialysis treatment, enrolled in a Medicare Advantage plan or have Medicare as a secondary payer, or lack Medicare Part A or Part B coverage as of July 1, 2006.

After development of the original intervention group, it was identified that eligibility for the intervention and comparison groups was not determined at the same time. For the MGH population, the intervention group eligibility was determined on July 1st 2006 while the comparison group eligibility was determined on July 4th 2006. To correct for this 3 day difference, a historical eligibility pull was completed on September 1st 2006 for both the original

[2] Residence in a skilled nursing facility or nursing home was determined using the following CPT codes: 99301–99303, 99311–99313, and 99321–99333.

intervention and comparison groups. Using this view and a new eligibility date of July 4th 2006, the intervention group was trimmed to 2,640 beneficiaries (from 2,655 - a removal of 15 beneficiaries). MGH also requested the removal of 21 intervention beneficiaries as part of the 60-day review process. There were 21 beneficiaries who met the physician visit loyalty criteria as a result of their relationship with one of two current MGH/MGPO primary care physicians who were employed by another area medical center during 2005. Using these criteria, a total of 2,619 Medicare beneficiaries were assigned to MGH's CMP intervention group.

Following the development of the original intervention group criteria, MGH worked with CMS and RTI to develop specifications to select a comparison group of beneficiaries to be used in conducting the financial reconciliation and evaluation of its demonstration program. The comparison group was selected using the following eligibility criteria:

- Medicare FFS beneficiaries with a primary residence in one of the five designated counties used to select the intervention group.

- Beneficiaries who were loyal to one of 307 comparison physicians as determined by the loyalty criteria used to select the intervention group. The 307 physicians were selected because they practice at one of 18 group practices associated with 1 of 4 academic medical centers that served as comparison institutions for MGH's CMP: Beth Israel Deaconess Medical Center, Boston Medical Center, New England Medical Center, and St. Elizabeth's Medical Center. Practices were selected if at least 10% of services provided were office visits.

- Received a majority of inpatient care from the comparison facilities or any of several long-term care facilities, in-hospital rehabilitation or psychiatric units that accounted for at least 1% of all claims at one of the four comparison institutions.

The exclusion criteria that were applied to the intervention group were also used to limit the comparison group. In addition, a small number of patients were also excluded because they had qualified as members of the intervention group. The comparison group was also matched to the intervention group in terms of risk and cost, by determining the cut points that divided the intervention group into three equal-sized groups, that is tertiles, based on HCC scores and three tertiles for Medicare costs. These two factors were cross-classified to form nine strata, and beneficiaries eligible for the comparison group were randomly sampled to match the number of intervention beneficiaries in each stratum. As noted above, eligibility for the intervention and comparison groups was not determined at the same time. Using a September 1st 2006 view, the comparison group was trimmed to 2,755 beneficiaries (from 2,786 – a removal of 31 beneficiaries).

The same general and specific inclusion and exclusion criteria used to select the original and intervention and comparison populations were used to identify the refresh intervention and comparison populations. Medicare claims for calendar year 2006 were used to identify the refresh intervention and comparison populations. Prior to determining loyalty, MGH requested that 9 physicians in their original group be dropped from the refresh because they were no longer MGH primary care physicians. To identify comparison physicians no longer affiliated with the selected comparison networks, RTI examined the billing patterns for all 801 comparison

physicians used to select the original population. A total of 55 physicians were removed from the comparison networks.

CMS limited the size of the Mass General refresh intervention group to 30% of the original starting intervention group of 2,619 beneficiaries, or 785 beneficiaries, who were randomly selected from a pool of 1,870 beneficiaries. Using the same procedure to select the original comparison group, 785 beneficiaries from 9 cost-HCC strata were randomly selected.

One sampled beneficiary who had already been selected for the intervention refresh group was subsequently identified. This beneficiary was removed from the comparison group, reducing the size of the comparison refresh population to 784 beneficiaries.

1.2.4 CMP Operations

MGH's CMP was launched on August 1, 2006. MGH negotiated a per-beneficiary-per-month (PBPM) payment of $120 for the duration of the demonstration. At the end of the 3-year period, MGH was contractually obligated to achieve a 5% savings in Medicare payments among the intervention group (regardless of participation in its Care Management Program) compared to the comparison group, and to cover program fees collected. In addition, MGH has the opportunity to share a portion of any savings beyond 5% that is achieved.

MGH's CMP was intended to provide an enhanced level of care to a high-risk patient population through comprehensive outpatient practice-based case management. Other goals included the following:

- reduce health care costs through reductions of preventable hospitalizations and emergency room visits,
- improve physician work life, and
- generate increased understanding of delivering effective PBCM programs, including the development of a satisfying and manageable role for case managers.

To achieve all of these goals, MGH's CMP was structured to facilitate communication and improve relationships (a) between patients and case managers, (b) between patients and physicians, (c) between case managers and physicians, and (d) among case managers.

The CMHCB demonstration provides MGH with the opportunity to develop a role for nurse case managers that allows them to effectively support patients while maintaining job satisfaction. Further, MGH intends to use the CMHCB program to learn how to deliver the appropriate breadth and depth of CMP services for individuals with multiple chronic illnesses. According to MGH, these high-cost individuals are not well served by the current health care delivery system, which provides disease-specific care that does not sufficiently address the issues faced by patients with comorbid conditions. These patients may be better served using a case management model that addresses their needs across all care settings.

1.2.5 Overview of the MGH CMHCB Demonstration Program

RTI conducted two site visits to the MGH'S CMP in Boston, MA. The initial site visit was conducted in March 2007, approximately 7.5 months after the launch of the program. The site visit was designed to focus on implementation – understanding the services offered by MGH's CMP and reporting early experiences with program implementation and engagement of eligible beneficiaries, providers, and CMS. The second site visit occurred approximately 25 months into the demonstration and focused on the program staff's impressions and interpretation of their two year experience working on the demonstration program. During the follow-up visit, RTI staff met with the senior management of the MGH Physicians Organization, MGH's CMP leadership, and key CMP staff. The protocol to conduct the follow-up interviews included a wide range of questions related to program implementation, program monitoring/outcomes to date, and implementation experience/lessons learned.

The description of MGH's CMP and its activities contained in this report reflects the CMP staff's impressions and interpretation of its experience, and does not necessarily reflect RTI's or CMS's perspective on these issues. First, we describe the continuum of services provided to CMP participants and physicians, as well as the clinical protocols/analytic tools to support the CMP nurse case managers and other health professionals who deliver these services. Second, we discuss program changes and enhancement activities that occurred as the program evolved.

The core element of MGH's CMP is the one-on-one relationship between patients and their practice-based case managers, supplemented by support received from the program's mental health, pharmacist, and end-of-life components. During the first 6 months of the demonstration period, case managers focused their efforts almost exclusively on beneficiary outreach and enrollment tasks. Therefore, case management activities began in earnest on February 1, 2007.

Case management. MGH's CMP is designed so that case managers become staff members of each physician practice. According to MGH leadership, this association with the primary care provider engenders patient trust and willingness to discuss health care and psychosocial problems with these nurses. Case managers developed relationships with patients over time through telephone calls and in-person interactions during physician office visits or at the hospital, if they are admitted for an inpatient service. Case managers also conducted visits to patient homes on an as-needed basis. Overall, case managers assessed patient needs, collaborated with physicians to develop treatment plans, educated patients about options for medical treatment and support services, facilitated patient access to services, and supported patient self-management of medical conditions.

Case managers conducted a comprehensive assessment to evaluate the unique needs of each patient. Case managers focused the assessment on issues that were relevant to each patient and evaluated medical and psychosocial problems, the resources used to address these issues, and patient needs for additional support. The tool used to conduct these assessments was developed by MGH and includes several externally validated instruments, such as questions to evaluate challenges encountered with activities of daily living (ADL). Using information collected from the assessment, case managers developed a care plan for each patient in

conjunction with the primary care provider and the practice's clinical team. Case managers implemented care plans over time by addressing urgent patient issues, conducting patient education, and providing referrals to support services. Throughout the program period, case managers continued to evaluate patients as their issues and need for support evolved over time.

Case managers educated patients about resources available and lifestyle changes that could help to prevent exacerbations of disease, to prevent or delay hospitalization. Case managers reviewed self-management activities, such as getting exercise and eating a low-salt diet, during a series of calls over a week or two to help patients adopt new behaviors. Case managers also educated patients about the purpose of their medications and other treatment interventions to help increase patient adherence to care plans.

Case managers also facilitated coordination of patient care across the continuum of health care services. Case managers received paged messages when their patients were admitted to the emergency room and an email indicating an inpatient admission. Using these real-time alerts, case managers could visit their patients in the hospital and research the cause of the hospitalization to inform refinements to the patient's care plan that may prevent future inpatient stays. Following hospital discharge, case managers contacted patients to make sure that they understood and could comply with discharge plans and coordinated with home health care providers to stay informed of patient health status.

Case managers also facilitated patient access to health care resources through patient education and referrals to other hospital or community services. For example, case managers informed patients that instead of going to the emergency room if they have a health problem, they could contact the physician's office at any time and may be able to see the doctor on the day of the call. Each week case managers received a list of patients scheduled to attend a physician office visit, and case managers contacted patients prior to their scheduled physician visits to find out if they needed assistance with transportation to the office. In addition, case managers followed up with patients via telephone if they missed their appointments to determine the issues involved and to provide support needed for patients to see their physician.

Mental health program. During the early stages of the CMHCB demonstration program, MGH learned that many high-cost, complex patients have mental health issues that were not effectively addressed by the current model of health care delivery or its pilot program. As a result, the program allocated greater resources to support mental health, hiring a social worker to assess the mental health needs of CMP participants and support them in accessing psychiatric care as needed or provide treatment if appropriate.

CMP pharmacist. Since many members of the CMP population have complex medication regimens, MGH enlisted the support of a pharmacist to review the appropriateness of medication regimens and assist patients with access to medications. The pharmacist also evaluated medication regimens to identify opportunities to reduce the number of medications and to suggest alternative therapies. In addition, the pharmacist worked with a PharmD resident to conduct a 6-week pilot program which involved visiting program participants in the hospital on the day of their discharge to discuss access to medications once they return home. Lessons learned from this pilot helped to motivate a change in MGH's medication delivery services—MGH authorized home delivery service for any CMP patient referred by their Care Manager.

End-of-life care. MGH's goal is to facilitate earlier access to end-of-life resources and create a seamless transition between hospital and home care for patients approaching the end of life. The CMP received support from a nurse who specializes in end-of-life care issues. The nurse educated case managers and physicians about how to have discussions about end-of-life issues with patients, how to support patients in developing advance directives, and provided information about the services that are available through hospice.

Physician Support Services. Once MGH had generated lists of CMP-eligible beneficiaries receiving care from each physician, the CMP medical director met with each practice to introduce the program and discuss which patients were at highest risk for acute events and should receive priority for enrollment. The medical director also met with specialty practices such as the oncology, cardiology, emergency, and orthopedics departments to explain the resources available through the program, because case managers would likely interact with these providers as they facilitated patient access to these services.

Case managers assigned to each practice met with physicians at the practices to describe the program, the skills that they bring to the physician practice, and their interest in collaborating to support patients in their efforts to manage their medical conditions. Case managers collected information from providers about how they could add value to the medical practice.

Staffing. MGH invested considerable personnel resources to develop and implement its CMP. At the time of the initial site visit, the program was staffed with 11 nurse case managers who received guidance from the program leadership and support from the project manager, an administrative assistant, and a community resources specialist. Each case manager was located in a physician practice and provided case management support to a group of 180 to 220 patients who received their primary care from a provider in that group. In addition, almost all case managers also "floated" to one or two additional small physician practices, which had a relatively small panel of MGH CMP eligible patients. Responsibilities included conducting patient assessments, visiting patients who were hospitalized at MGH (when feasible), contacting patients who visited the emergency department or were recently discharged from the inpatient services, calling patients scheduled for office visits each week, following up with patients who missed office visits each week, making follow-up calls to provide case management services, and promoting the MGH CMP to physicians.

Other team members who provided support to the case managers included the CMP project manager, an administrative assistant, a community resource specialist, and a patient financial counselor who provided support for all insurance-related issues. Case managers consulted the community resource specialist in their efforts to connect patients with existing resources in the community. The community resource specialist also interacted directly with patients to address nonclinical issues such as transportation and housing needs and attended weekly meetings with the case managers to provide input in discussions about patient issues.

The program leadership provided guidance to case managers, managed the program, and marketed the program throughout MGH. The clinical team leader provided oversight and supervision of case managers, represented the case managers in discussions of program development, and coordinated orientation and staff development for case management staff. The medical director provided oversight and day-to-day management of MGH's CMP, acted as

primary spokesperson and project champion among physicians, contributed to the development of program interventions, served as a consultant and clinical resource to case managers, and assisted in evaluating medical problems and making medical decisions.

Clinical Outcomes Monitoring. MGH developed a series of clinical dashboards using data from the MGH electronic medical record, claims data, and its enrollment tracking database. The dashboards allowed MGH to examine trends in health care utilization and outcomes, overall and by enrollment status, physician practice, and/or case manager as well as activities of its case managers. Examples of indicators included in the dashboard are: number of assessments completed within 90 days, number of referrals or interventions conducted, number of advanced directives in place, number of participants screened for depression, number of participants with a positive screen for depression referred to mental health, and the top 10 discharge diagnoses.

Beneficiary Outreach/Engagement

Beneficiary Outreach/Engagement. As required by CMS, MGH initially sent a letter from the Medicare program to introduce the intervention cohort to MGH's CMP. MGH then sent a second letter on its own letterhead signed by the MGH CMP Medical Director, Eric Weil, with information about the PBCM program. In response to these initial letters, approximately 10% to 15% of the intervention population called MGH to enroll in the program. MGH contracted with Health Dialog, a disease management vendor, to conduct a telephone outreach campaign to solicit additional participation for MGH's CMP so that case managers could focus their efforts on delivering care management services. After observing lower than expected enrollment rates during the first 2 months of program operations, MGH changed its outreach process such that case managers solicited beneficiary participation as representatives from their primary care providers' offices, by calling beneficiaries, talking with them during physician visits, and visiting with patients admitted to the hospital. Simultaneously, many of the physician practices developed the infrastructure required to send letters encouraging participation in the program, which were signed by each patient's primary care physician. In addition, physicians were able to enroll patients in the program during office visits. To manage the outreach process, case managers prioritized patients admitted to the hospital and those scheduled for a physician office visit.

Provider Outreach/Participation. MGH enlisted physician support to help ensure the success of its CMP in providing high-quality care to patients. Physicians were asked to conduct the following activities: encourage beneficiaries to participate in the program and enroll them in the program when possible; collaborate with case managers to review initial assessment findings and develop care plans for each patient; inform case managers about patient events and refinements to patient care plans during the demonstration period; and discuss advance directives with enrolled patients.

MGH provided physicians with a $150 financial incentive per patient in Year 1 and $50 in Years 2 and 3 to help cover the cost of physician time for these activities. Two hundred physicians from the 19 MGH primary care medical practices that treat adult patients served as the primary care physician to at least 1 and as many as 59 patients eligible for MGH's CMP. MGH physicians received information about the CMP from a variety of sources, including the

program's medical director, the MGH electronic newsletter, and case managers assigned to each practice.

1.2.6 Early Experience

During the first 7 months of the demonstration period, MGH's CMP staff reportedly learned a lot about the characteristics of the intervention population and the delivery of case management services for a population of sick patients with complex medical and psychosocial needs. They expected that its program would have the greatest impact by preventing acute health care events among beneficiaries who were initially not having significant health issues; however, case managers found that they spent a lot of time dealing with urgent issues for patients who "spiraled out of control." Although initially some patients were skeptical about the MGH CMP, overall, patients quickly formed relationships with case managers, including several who requested daily contact with their case managers to help them with their numerous issues.

MGH observed that the program model appeared to work better in larger practices, where the impact of the program was more easily observed among a larger panel of patients and it is economically more feasible to embed a full-time case manager. Further, small physician practices were more likely to have procedures firmly in place, and therefore encountered more challenges in integrating the case managers into their practices.

MGH encountered several challenges in implementing its CMP, including ramping up the program in a short time frame, effectively supporting patients residing in skilled nursing facilities or other institutions, and balancing the breadth and depth of support provided to each participant. MGH reported that it had limited time to develop the infrastructure needed to deliver its program. For example, MGH's information systems were not fully set up to manage program activities on August 1, 2006, when the demonstration period began. In addition, MGH began hiring case managers in June 2006 and did not finish the hiring process until December 2006. Consequently, beneficiary enrollment took longer and was less successful at practices that did not have a case manager during the initial months of the demonstration.

Although MGH tried to remove institutionalized beneficiaries from its intervention population, a small proportion of the intervention population resided in skilled nursing homes or other facilities. MGH speculated that these patients were included as a result of the lag between the time period represented by the claims data used to select intervention patients and the program's launch date. Case managers were challenged in balancing the depth and breadth of support provided to each patient. Program management staff monitored the number, length, and frequency of patient interactions to evaluate the most effective and efficient way to deliver case management services to this sick population.

1.2.7 Program Changes

A number of changes to MGH's CMP occurred after the initial site visit. Noteworthy changes include the following:

- **Termination of the contract with Health Dialog and bringing assessment and data collection capabilities in-house**: Given that the contract was very expensive, communication with Health Dialog was very difficult due to HIPAA, and Health

Dialog's limited experience with practice-based case management with a medically complex FFS population, the CMP leadership made a decision to terminate its contract with Health Dialog and bring enrollment and data system development capabilities in-house.

- **Revision of its assessment tool**. The program shifted from using Health Dialog applications to using the hospital-based Medical Information Data System (MIDAS) for recording patient assessment data. During the switch, the entire assessment process was reviewed by program leadership and major changes to the process were implemented. The single extensive assessment that was conducted initially was broken down into six different assessment modules that could be conducted depending on participants' needs: functional assessment, mental health, advance care planning, transportation, pharmacy, and post hospital episode.

- **Program staffing changes:**

  - addition of a designated case manager position to work specifically on post discharge assessments to enhance transitional care monitoring;

  - creation of a data analytics team to develop and strengthen the program's reporting capabilities;

  - expansion of the role of the community resource specialist to provide patient support and referral for non-clinical services; and

  - expansion of the mental health component of the program by adding a mental health team director, clinical social worker, two psychiatric social workers, and a forensic clinical specialist (M.D./J.D.), who follows highly complex patients with issues such as legal issues, guardianship and substance abuse.

- **Data System Enhancements:**

  - integration of a clickable CMP icon into the MGH-wide Longitudinal Medical Record (LMR) system so every provider may immediately identify Medicare beneficiaries in the program and be able to get in touch with their patient's case manager by page or e-mail;

  - development of a data portal for primary care physicians that would be able to view their inpatient census and program panel; and

  - implementation of a dual computer monitor system so every case manager can have simultaneous access to both data systems from their desks.

1.3 Organization of Report

In ***Chapter 2***, we provide an overview of our evaluation design and a description of the data and methods used to conduct our analyses. ***Chapter 3*** contains a summary of our previously reported assessment of beneficiary satisfaction, self-management, and functioning at the

midpoint of MGH's CMP period and provider satisfaction with MGH's CMP culled from interviews with physicians during the second site visit. In ***Chapter 4***, we provide the results of our analyses of participation levels in MGH's CMP and level of intervention with participating beneficiaries (i.e., the number of in-person visits and/or telephonic contacts). In ***Chapters 5 and 6***, we provide the results of our analyses of changes in clinical quality of care and health outcomes, respectively. ***Chapter 7*** presents our analyses of financial outcomes. We conclude with an overall summary of key findings and a discussion of the policy implications of these findings for future Medicare care management initiatives. ***Supplemental Tables for Chapters 2, 4, and 7*** are available from the CMS Project Officer upon request.

CHAPTER 2
EVALUATION DESIGN AND DATA

2.1 Overview of Evaluation Design

2.1.1 Gaps in Quality of Care for Chronically Ill

Medicare beneficiaries with multiple progressive chronic diseases are a large and costly subgroup of the Medicare population. The Congressional Budget Office (CBO) estimated that in 2001 high-cost beneficiaries (i.e., those in the top 25% of spending) accounted for 85% of annual Medicare expenditures (CBO, 2005). Three categories of high-cost users—beneficiaries who had multiple chronic conditions, were hospitalized, or had high total costs—were identified by CBO for study of persistence of Medicare expenditures over time. Beneficiaries that were selected based upon hospitalization or being in the high total cost groups had baseline expenditures that were four times as high as expenditures for a reference group. Beneficiaries selected based upon presence of multiple comorbid conditions had baseline expenditures that were roughly twice as high as expenditures for a reference group. Subsequent years of costs remained higher for all three cohorts than the reference group; however, total expenditures declined the most for those beneficiaries who were identified as high cost due to a hospitalization followed by beneficiaries who had had high total costs in the base year. Subsequent costs were virtually unchanged for beneficiaries with multiple chronic conditions.

Further, these beneficiaries currently must navigate a health care system that has been structured and financed to manage their acute, rather than chronic, health problems. When older patients seek medical care, their problems are typically treated in discrete settings rather than managed in a holistic fashion (Anderson, 2002; Todd and Nash, 2001). Because Medicare beneficiaries have multiple conditions, see a variety of providers, and often receive conflicting advice from them, there is concern that there is a significant gap between what is appropriate care for these patients and the care that they actually receive (Jencks, Huff, and Cuerdon, 2003; McGlynn et al., 2003). The MGH CMP demonstration has been designed to address current failings of the health care system for chronically ill Medicare FFS beneficiaries.

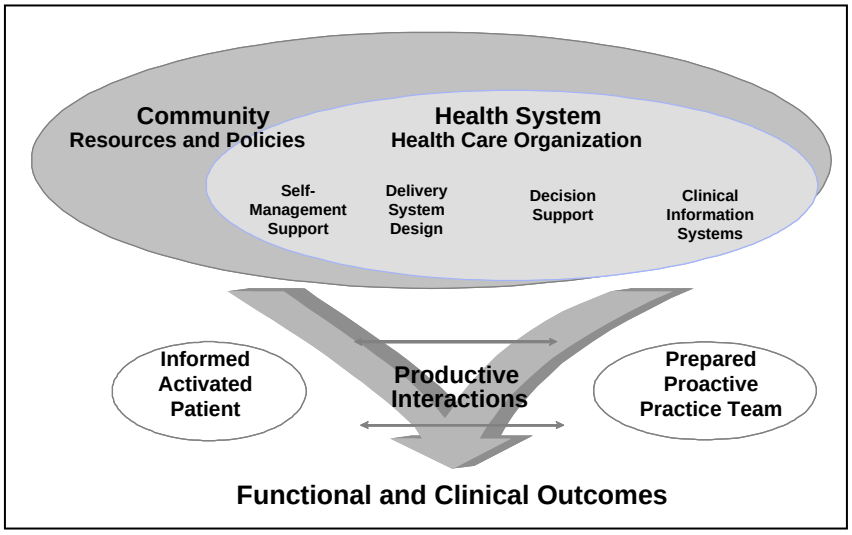
2.1.2 Emerging Approaches to Chronic Care

The Chronic Care Model—The concept of chronic care management as a patient-centered and cost-effective approach to managing chronic illness has been evolving for years. The Chronic Care Model (CCM), developed by Wagner (1998), has become a familiar approach to chronic illness care (***Figure 2-1***). This model is designed to address systematic deficiencies and offers a conceptual foundation for improving chronic illness care. The model identifies six elements of a delivery system that lead to improved care for individuals with chronic conditions (Glasgow et al., 2001; Wagner, 2002; Wagner et al., 2001):

- the community,
- the health system,
- self-management support,

- delivery system design,
- decision support, and
- clinical information systems.

Figure 2-1
Chronic Care Model



SOURCE: Wagner (1998). Reprinted with permission.

According to the model, patients are better able to actively take part in their own care and interact productively with providers when these components are developed, leading to improved functional and clinical outcomes.

Disease management and case management—The two most common approaches to coordinating care for people with chronic conditions are disease management and intensive case management programs (Medicare Payment Advisory Commission [MedPAC], 2004). Disease management programs teach patients to manage their chronic conditions and are often provided on a broader scale than case management programs. Services provided under a disease management program may include health promotion activities, patient education, use of clinical practice guidelines, telephone monitoring, use of home monitoring equipment, registries for providers, and access to drugs and treatments. Most disease management programs target persons with specific medical conditions but then take the responsibility for managing all of their additional chronic conditions. Case management programs typically involve fewer people than disease management programs (Vladek, 2001). Case management programs also tend to be more intensive and individualized, requiring the coordination of both medical and social support services for high-risk individuals. Typically, disease management programs are used with intensive case management for high-risk individuals who have multiple chronic conditions and complex medical management situations.

The empirical research on the effectiveness of disease management and case management approaches is mixed. Some studies have shown support for the clinical improvements and cost-effectiveness of disease management programs (Lorig, 1999; Norris et al., 2002; Plocher and Wilson, 2002; Centers for Disease Control and Prevention [CDC], 2002). Other programs, such as the CMS case management demonstration programs in the early 1990s, which required physician consent for patient participation, resulted in increased beneficiary satisfaction but failed to achieve any improvement in health outcomes, patient self-care management, or cost savings (Schore, Brown, and Cheh, 1999). In 2002, CMS selected 15 demonstration programs of varying sizes and intervention strategies as part of the Medicare Coordinated Care Demonstration (MCCD). None of the 15 programs produced any statistical savings in Medicare outlays on services relative to the comparison group, and two had higher costs (Peikes et al., 2009).[3] There were a few, scattered quality of care improvement effects. Two programs did show some promise in reducing hospitalizations and costs, suggesting that care coordination might at least be cost neutral. A major reason given for the lack of success in both Medicare savings and better health outcomes is attributed to the absence of a true transitional care model in which patients were enrolled during their hospitalizations. Studies have shown that approach to significantly reduce admissions within 30/60 days post-discharge, when patients are at high risk of being readmitted (Coleman et al., 2006; Naylor et al., 1999; Rich et al., 1995).

2.1.3 Conceptual Framework and CMHCB Demonstration Approaches

The care management organizations awarded contracts under this CMS initiative offered approaches that blend features of the chronic care management, disease management, and case management models. Their approaches relied, albeit to varying degrees, on engaging both physicians and beneficiaries and supporting the care processes with additional systems and staff. They proposed to improve chronic illness care by providing the resources and support directly to beneficiaries through their relationships with insurers, physicians, and communities in their efforts. The CMOs also planned to use all available information about beneficiaries to tailor their interventions across the spectrum of diseases that the participants exhibited.

Although each of the CMOs has unique program characteristics, all have some common features. These features include educating beneficiaries and their families on improving self-management skills, teaching beneficiaries how to respond to adverse symptoms and problems, providing care plans and goals, ongoing monitoring of beneficiary health status and progress, and providing a range of resources and support for self-management. Features of the CMHCB programs include:

- *Individualized assessment*. Several CMOs use proprietary algorithms to calculate a risk score or risk scores, while others depend on judgment of clinical staff. The scores are used to customize interventions to the participants' needs.

- *Education and skills*. A key step in improving self-management is educating beneficiaries and their families about their illnesses, how to react to symptoms, and

[3] These findings were based on regressions controlling for age, gender, race, disabled/aged entitlement, Medicaid coverage, and whether beneficiaries used skilled nursing facility (SNF) or hospital services prior to the demonstration.

what lifestyle changes to make. All of the CMOs provide a range of educational resources.

- *Medication management and support.* All of the CMO programs include efforts to optimize the medication regimens of participating beneficiaries. Some monitor compliance, some facilitate access to low-cost pharmaceuticals, and others offer face-to-face meetings with pharmacists.

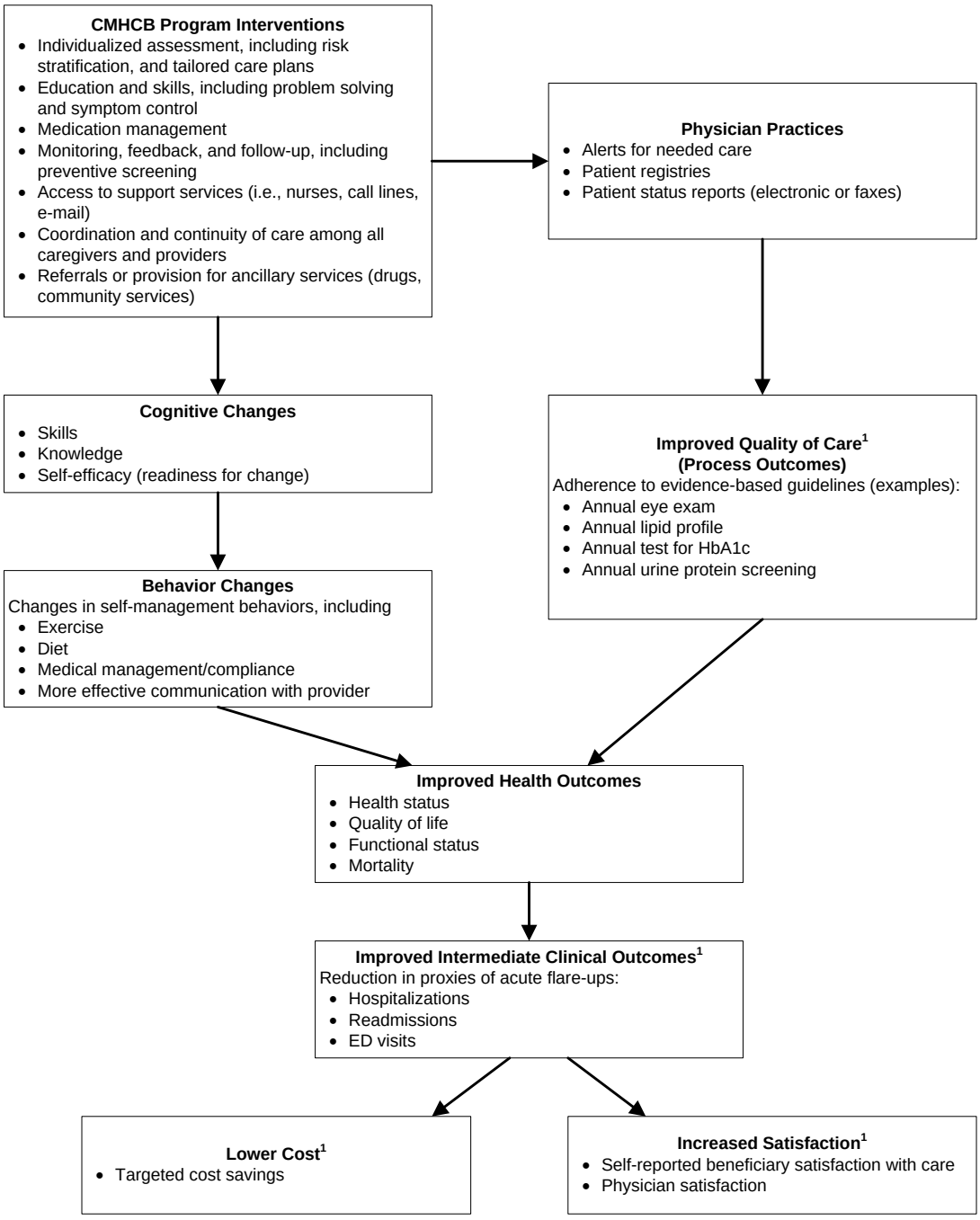
- *Monitoring, feedback, and follow-up.* Activities in this domain include ongoing biomonitoring of beneficiaries by placing scales or other equipment in their homes or by having the beneficiaries self-report their weights, blood sugars, or other measures. When data on preventive services, screenings, or recommended tests are available, the programs remind beneficiaries and/or their doctors to have them done. Flu shots are just one example.

- *Coordination and continuity of care.* One hallmark of the care management model is that it uses data from all available sources to disseminate information to providers and caregivers involved with a beneficiary's care. A limited number of the CMOs have case managers directly embedded in the physician practices, allowing for day-to-day and face-to-face interactions. Several CMOs also have direct communication with physicians via a shared electronic medical record. However, the majority of CMOs must engage physicians or physician practices more indirectly through telephone and fax communication.

- *Referrals or provision for community-based ancillary services.* Not all of a participant's needs are provided directly by the CMOs. All CMOs have recognized the need for transportation, low-cost prescriptions, or other services typically provided by community service organizations (e.g., social workers, dieticians). The CMOs developed relationships with other service providers and programs and helped selected beneficiaries receive these services through their participation in the CMHCB program.

Figure 2-2 presents RTI's conceptual framework for the overall CMHCB demonstration evaluation. It synthesizes the common features of the CMHCB demonstration implemented interventions and the broad areas of assessment within our evaluation design. The CMHCB demonstration programs employ strategies to improve quality of care while reducing costs by empowering Medicare beneficiaries to better manage their care. The programs do so in three ways: (1) by enhancing beneficiaries' knowledge of their chronic condition through educational and coaching interventions, (2) by improving beneficiaries' communication with their care providers, and (3) by improving beneficiaries' self-management skills. Successful interventions should alter beneficiaries' use of medications, eating habits, and exercise and should allow beneficiaries to interact more effectively with their primary health care providers. All of the CMHCB demonstration programs hypothesized that lifestyle changes and better communication with providers as well as improved adherence to evidence-based quality of care should improve health and functional status, which will mitigate acute flare-ups in chronic conditions, thereby reducing hospital admissions and readmissions and the use of other costly health services such as emergency rooms and visits to specialists. Experiencing better health and less acute care

Figure 2-2
Conceptual framework for the CMHCB programs



NOTE: CMHCB = Care Management for High Cost Beneficiaries; CMO = Care Management Organization; ED = emergency department.

SOURCE: RTI conceptual framework for the Medicare Care Management for High Cost Beneficiaries evaluation. Portions of this model are adapted from other sources, including the Chronic Care Model and the disease management model described in CBO (2004).

utilization, beneficiaries should also be more satisfied that their health care providers are effectively helping them cope with their chronic medical conditions, and providers should be more satisfied with the outcomes of care for their chronically ill Medicare FFS beneficiaries.

In this report, we present our findings with respect to the degree to which the MGH's CMP was able to engage its intervention population and achieve four outcomes. ***Table 2-1*** presents a summary of research questions and data sources, organized by three evaluation domains: Reach, Implementation, and Effectiveness. MGH's CMP implementation experience is reported in Chapter 1.

Table 2-1
Evaluation research questions and data sources

Research questions	Site visits	CMO data	Claims	Survey
IMPLEMENTATION: To what extent was MGH able to implement its CMP?				
1. To what extent were specific program features implemented as planned? What changes were made to make implementation more effective? How was implementation related to organizational characteristics of MGH's CMP?	Yes	Yes	No	No
2. What were the roles of physicians, the community, the family, and other clinical caregivers? What was learned about how to provide this support effectively?	Yes	No	No	No
3. To what extent did MGH's CMP engage physicians and physician practices in their programs?	Yes	No	No	No
REACH: How well did MGH's CMP engage its intended audiences?				
1. Were there systematic baseline differences in demographic characteristics and disease burden between the intervention and comparison group beneficiaries at the start of the demonstration?	No	No	Yes	No
2. How many individuals did MGH's CMP engage, and what were the characteristics of the participants versus nonparticipants (in terms of baseline clinical measures, demographics, and health status)?	No	Yes	Yes	No
3. What beneficiary characteristics predict participation in MGH's CMP?	No	Yes	Yes	No
4. To what extent were the intended audiences exposed to MGH's CMP interventions? To what extent did participants engage in the various features of the program?	No	Yes	No	Yes
5. What beneficiary characteristics predict a high level of MGH CMP intervention versus a low level of intervention?	No	Yes	Yes	No
EFFECTIVENESS: To what degree was MGH's CMP able to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings?				
Satisfaction outcomes				
1. Did MGH's CMP lead beneficiaries to be more satisfied with their ability to cope with their chronic conditions than beneficiaries in the comparison group?	No	No	No	Yes
2. How satisfied were physicians with MGH's CMP intervention?	Yes	No	No	No

(continued)

Table 2-1 (continued)
Evaluation research questions and data sources

Research questions	Site visits	CMO data	Claims	Survey
Functioning and health behaviors				
1. Did the program improve knowledge and self-management skills?	No	No	No	Yes
2. Did MGH's CMP result in greater engagement in health behaviors?	No	No	No	Yes
3. Did MGH's CMP result in better physical and mental functioning and quality of life than would otherwise be expected?	No	No	No	Yes
Quality of care and health outcomes				
1. Did MGH's CMP improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?	No	No	Yes	No
2. Did MGH's CMP improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?	No	No	Yes	No
3. Did MGH's CMP improve health outcomes by decreasing mortality?	No	No	Yes	No
Financial and utilization outcomes				
1. What were the Medicare costs per beneficiary per month (PBPM) in the base year versus the first 36 or 24 months of the demonstration for the intervention and the comparison groups?	No	No	Yes	No
2. What were the levels and trends in PBPM costs for intervention group participants and nonparticipants? Did nonparticipation, alone, materially reduce the intervention's overall cost savings?	No	No	Yes	No
3. How variable were PBPM costs in this high cost, high risk, population? What was the minimal detectable savings rate given the variability in beneficiary PBPM costs?	No	No	Yes	No
4. How did Medicare savings for the 36- or 24-month period compare with the fees that were paid out? How close was MGH's CMP in meeting budget neutrality?	No	No	Yes	No
5. How balanced were the intervention and comparison group samples prior to the demonstration's start date? How important were any imbalances to the estimate of savings?	No	No	Yes	No
6. Did the intervention have a differential effect on high-cost and high-risk beneficiaries?	No	No	Yes	No
7. What evidence exists for regression-to-the-mean in Medicare costs for beneficiaries in the intervention and comparison groups?	No	No	Yes	No

NOTE: MGH's CMP = Massachusetts General Hospital's Care Management Program; ER = emergency room; PBPM = per beneficiary per month.

2.1.4 General Analytic Approach

The CMHCB initiative is what is commonly called a "community intervention trial" (Piantadosi, 1997). It is a "community" in the sense of being population based for a pre-specified geographic area. It is "experimental" because it tests different CMHCB program interventions in different areas. It is a "trial" that employs randomization (or selection of a comparison population) following an "intent-to-treat" (ITT) model. The initiative is unusual because it employs a "pre-randomized" scheme, wherein CMS assigns eligible beneficiaries to an intervention or comparison stratum before gaining their consent to participate. In fact, comparison beneficiaries are not contacted at all. Further, beneficiaries opting out of the intervention are assigned to the intervention group, even though they will receive no CMO services. These refusals are included in the same stratum as those receiving care coordination services on an ITT basis.

Beneficiaries who become ineligible during the demonstration program are removed from the intervention and comparison groups for the total number of days following loss of eligibility for purposes of assessing cost savings and quality, outcomes, and satisfaction improvement. A beneficiary's eligibility status for the CMHCB program may change multiple times during the 3-year demonstration. For example, an eligible beneficiary may switch to a Medicare Advantage program during the second year and switch back to FFS during the third year. Our evaluation includes all months in which a beneficiary is eligible for the initiative, and we accounted for differential periods of eligibility in the analysis.

Further, the CMOs differentially engaged and interacted more with beneficiaries for whom they believe their programs will result in the greatest benefit, either in terms of health outcomes or cost savings. Thus, not all intervention beneficiaries participated nor did all beneficiaries receive the same level of intervention. In fact, some participants received very few services.

The CMHCB programs reflect a dynamic process of system change leading to behavioral change leading to improved clinical outcomes, and the type of experimental design within this demonstration calls for a pre/post, intervention/comparison analytic approach—sometimes referred to as a difference-in-differences approach—to provide maximum analytic flexibility. The strategy will be used to construct estimates of all performance outcomes of each demonstration program.

Our proposed model specification to explain any particular outcome variable, Y_{t+1}, measured during the intervention program follow-up period:

$$Y_{t+1} = \alpha + \beta_1 I + \beta_2 Y_t + \beta_3 I \bullet Y_t + \beta_4 X + \varepsilon \qquad (2.1)$$

where

α = the intercept term, or reference group;

I = 0,1 intervention indicator;

Y_t= the outcome measured during a base or predemonstration period;

X = a vector of beneficiary covariates; and

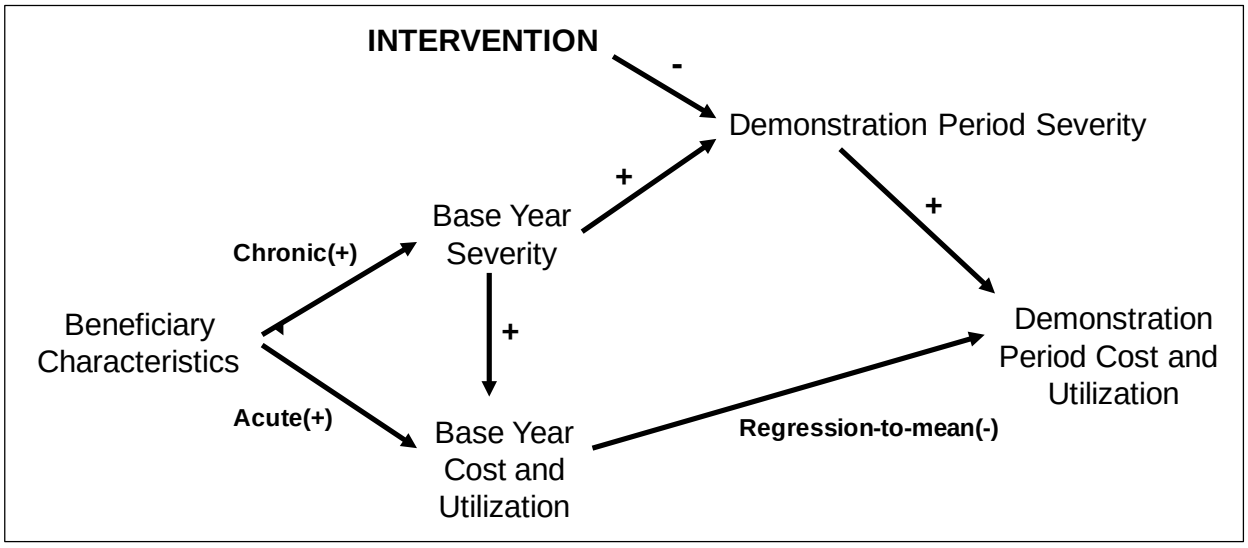
ε = a regression error term.

This model uses three sets of variables in analysis of covariance (ANCOVA) format to capture differences between intervention and comparison beneficiaries. The coefficient provides a test of the difference between the intervention group and comparison group in the base period for a particular outcome variable. (The reference comparison group mean value is in the α intercept.) If preprogram assignment is successful, will be approximately zero before controlling for beneficiary-specific (X) factors. The β_2 coefficient tests for temporal changes between pre- and post-demonstration outcomes, while the β_3 interaction coefficient tests whether the intervention group's performance profile differs over time from the comparison group's performance. The vector of β_4 coefficients controls for beneficiary-specific covariates influencing individual differences in the dependent variable of interest. Including covariates should set the estimated equal to 0, if selection of a comparable comparison population is contravened in some way. Program effects during the demonstration are reflected in the interaction coefficients. The null hypothesis is that the coefficient for β_3 is zero, implying no CMHCB program impact. Estimates that are significant at the 95% confidence level imply distinct program effects. The model may also be expanded to conduct analyses across beneficiary subpopulations and CMHCB intervention characteristics.

Because we will be analyzing change over time, it is important to consider the likely trajectory in our outcome measures as a function of beneficiary characteristics at baseline. ***Figure 2-3*** displays an alternative conceptualization of how the CMHCB intervention could alter the expected demonstration period outcomes of interest. At baseline, beneficiaries were selected for the demonstration because of higher baseline risk scores as well as high baseline expenditures as a proxy for clinical severity. These beneficiaries also have a multiplicity of other health care issues—chronic and acute—leading to high baseline costs and acute care utilization. The bottom half of ***Figure 2-3*** displays the statistical phenomenon observed in cohort studies of regression-to-the-mean. Beneficiaries with high costs and utilization are likely to regress toward average levels in a subsequent period and vice versa. Because we start with beneficiaries with high costs and utilization, our expectation is that there would be significant negative regression to the mean; thus, we would observe lower costs and utilization in the demonstration period absent an intervention effect.

Prior research has shown that physical health status declines rather substantially over time for elderly populations, and in particular, for chronically ill elderly populations (Ware 1996). The top half of ***Figure 2-3*** displays the expected positive relationship between base year and demonstration period severity and the positive relationship between increasing severity of illness and medical costs and utilization during the demonstration period absent an intervention effect. The CMHCB demonstration is aimed at improving or preventing further deterioration in health and functional status. Thus, our expectation is that the CMHCB program intervention would have a negative or moderating influence on growing patient severity during the demonstration period, thereby reducing the expected positive relationship between demonstration period severity and costs and utilization.

Figure 2-3
Conceptualization of influence of beneficiary baseline health status and cost and utilization patterns on CMHCB demonstration period acute care utilization and costs



2.2 Participation, Clinical Quality and Health Outcomes, and Financial Outcomes Data and Analytic Variables

This section provides a description of the data used to evaluate participation in and the effectiveness of MGH's CMP. As noted in Chapter 1, we also conducted a survey of MGH's CMP beneficiaries to assess their satisfaction with the CMHCB program and semi-structured interviews with a small number of physicians to assess their awareness of and satisfaction with the CMHCB program. The data used to make those assessments are described in ***Chapter 3***.

2.2.1 Data

We used six types of data for our evaluation analyses related to participation, clinical quality and health outcomes, and financial outcomes. Specifically, we used the following data sources:

- *Participant status files*. We received participant status files from ARC. The participant status information originates from MGH's CMP and was submitted to ARC. This file was updated quarterly and logged status changes among the intervention groups by MGH's CMP. Participation status was able to be determined on a monthly basis using three monthly indicators on a given quarterly file, and we used these indicators to determine the participation decision of the original and refresh intervention beneficiaries during each month of the demonstration.

- *Finder file*. RTI used this file, produced by ARC, to identify the group into which each MGH CMP beneficiary was assigned—intervention or comparison—for both the original and refresh populations.

- *Enrollment Data Base (EDB) daily eligibility files.*

  - ARC provided RTI with an EDB file for MGH's CMP comprised of all original and refresh beneficiaries. RTI used this file to determine daily eligibility based on MGH's CMP eligibility criteria (***Table 2-2***). The EDB file, in conjunction with the eligibility criteria, allowed us to identify beneficiaries as eligible or ineligible for each day of the intervention period and retrospectively for each day one-year prior to MGH's CMP launch date. We used the files to identify days of eligibility during the 12-month baseline period and the intervention periods of the demonstration and to select claims data during periods of eligibility in both the baseline and intervention periods. *Only beneficiaries who had at least 1 day of eligibility in the baseline and the demonstration periods are included in our evaluation.*

  - RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (July 4, 2006) for MGH's CMP original population.

  - RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (July 18, 2007) for MGH's CMP refresh population.

- *Medicare claims data produced by ARC*. In keeping with the financial reconciliation, CMS requested that RTI use the ARC claims files for all analyses. Monthly, ARC receives claims data from a CMS prospective claims tap, and on a quarterly basis creates netted claims files. As of each quarter's processing, ARC updates prior quarterly netted claims files with claims data processed after the prior cutoff dates. These files contain the claims experience for original and refresh intervention and comparison beneficiaries during the 12 months prior to MGH's CMP start date and claims with processing dates that span the full intervention period and 9 months thereafter (or claims run out).

- *CMO beneficiary intervention data files*. Quarterly, MGH's CMP sent RTI beneficiary-level intervention files that contained summary counts of telephonic contact, in-person and home visits, referrals with the participant subsequently seen by a care provider, and indicators of participants on telemonitoring devices. More detailed information on the contents of these files is in ***Chapter 4***.

- *FU Long Term Indicator (LTI) file*. Information in this file is obtained from the Minimum Data Set (MDS) of nursing home assessments and contains data on which Medicare beneficiaries are residents of nursing homes. We use this file to determine institutionalization status during the original and refresh intervention periods for the participation analysis.

Table 2-2
Criteria used for determining daily eligibility during MGH's CMP

Ineligibility reasons	Description
Death	Ineligible beginning on day following date of death.
ESRD	Ineligible beginning on day of ESRD enrollment. Eligible on day following ESRD disenrollment.
MA plan	Ineligible on day of MA plan enrollment when GHO contract number does not equal the contract number for MGH's CMP. Eligible on day following MA plan disenrollment.
Medicare secondary payer	Ineligible on day Medicare becomes secondary payer for working-aged beneficiary with an employer group health plan (primary payer code A) or for working disabled beneficiary (primary payer code G). Eligible on day following Medicare secondary payer end date.
Residence	Ineligible on residence change date indicating that a beneficiary has moved out of the service area determined by state code or state and county codes. Eligible on subsequent residence change date indicating that a beneficiary has moved into the service area determined by state code or state and county codes.
Part A/Part B enrollment	Eligible on day Part A/Part B coverage begins/resumes. Ineligible on day after Part A/Part B coverage ends.

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; ESRD = end-stage renal disease; MA = Medicare Advantage; GHO = Group Health Organization.

Table 2-3 contains MGH's CMP evaluation start and end dates, both baseline and intervention periods, for the original and refresh populations.

Table 2-3
Analysis periods used in MGH's CMP analysis of performance

Intervention period start date	Intervention period final end date	Intervention period months of intervention data	Baseline period start date	Baseline period end date
Original Population				
8/1/06	7/31/09	36	8/1/05	7/31/06
Refresh Population				
8/1/07	7/31/09	24	8/1/06	7/31/07

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

2.2.2 Analytic Variables

To conduct our participation, clinical quality and health outcomes, and financial analyses, we constructed nine sets of analytic variables from the aforementioned files.

1) ***Demographic Characteristics and Eligibility***. Age, gender, race, Medicare status (aged-in versus disabled), and urban residence were obtained from the EDB and determined as of the date of assignment, July 4, 2006 for the original population and July 18, 2007 for the refresh population. Medicaid enrollment was determined at any time during the baseline period and was also determined using the EDB.

   Daily eligibility variables were used to create analytic variables representing the fraction of the baseline and demonstration period that the intervention and comparison beneficiaries were CMHCB program eligible. These eligibility fractions were created based on the time period of the analysis. For example, the baseline eligibility fraction is constructed using the number of eligible days divided by 365. For the full intervention period, the denominator is adjusted based on the number of days that MGH's CMP was active in the demonstration. The numerator is the number of days the beneficiary is eligible during that time period. MGH participated in the demonstration for the full 36 months, so the number of days in the denominator for each original population beneficiary in MGH's CMP is 1,096 (MGH's CMP end date minus MGH's CMP start date + 1). If a beneficiary died 420 days into the intervention period, the eligibility fraction would be 420 divided by 1,096, or 0.383.

2) ***Institutionalized Status.*** Four binary indicators of institutionalization were created for both the original and refresh populations:

- Whether a beneficiary was in a nursing home for any one or more months of the initial 6 months of the demonstration period using the FU LTI file. This measure of institutionalization is used in all but the financial analyses.

- Whether a beneficiary had any baseline long-term-care (LTC) hospital costs in the baseline year. LTC hospitals are identified if the last four digits of the provider ID ranged from 2000 to 2299.

- Whether a beneficiary had any baseline skilled nursing facility (SNF) costs.

- Whether a beneficiary had any baseline nursing home services. These claims were identified if the Current Procedural Terminology (CPT) codes ranged from 99304 to 99340 or the location of service ranged from 31 to 33. An indicator for nursing home services was only created if there were two or more encounters during 2 consecutive months 3 months prior to the intervention period.

3) ***Hierarchical Condition Category (HCC) Risk Scores.*** Two HCC scores are used in this evaluation:

- A *prospective HCC score* calculated by RTI for a 12-month period prior to the start of the demonstration program using the 2006 CMS-HCC risk-adjustment payment model for both the original and refresh populations.

- A *concurrent HCC score* calculated by RTI for the first 6 months of the intervention period for both the original and refresh populations. In contrast to the predictive model, which uses a prior year's worth of claims data to generate a predicted HCC score, the concurrent model produces an HCC score based upon the current period's claims experience. Furthermore, we restrict the model to only 6 months of data. In RTI's experience, 80% of the HCC score is determined by 6 months of claims. Thus, we inflated the concurrent HCC score by 1.25 to approximate a score that otherwise would be calculated on a full year's data. The concurrent model used in this project is a 2004 model that was calibrated to the CMS Physician Group Practice (PGP) demonstration population. This is a FFS population that used services, rather than the entire FFS population used for payment purposes. This is a reasonable reference population because all CMHCB demonstration populations were also required to have used services to be selected for randomization.

4) ***Health Status***. We constructed three sets of analytic variables to reflect health status prior to and during the demonstration:

- *Charlson index.* We constructed the Charlson comorbidity index using claims data from the inpatient, outpatient, physician, and home health claims files. We created an index for the year prior to the start of the demonstration program. ***Supplement 2A*** contains the SAS code used to create this index.

- *Comorbid conditions.* RTI reviewed the frequency of diagnoses associated with evaluation and management (E&M) visits for the full study population in the year prior to the demonstration program to identify frequently occurring comorbid conditions: heart failure; coronary artery disease; other respiratory disease; diabetes without complications; diabetes with complications; essential hypertension; valve disorders; cardiomyopathy; acute and chronic renal disease; renal failure; peripheral vascular disease; lipid metabolism disorders; cardiac dysrhythmias and conduction disorders; dementias; strokes; chest pain; urinary tract infection; anemia; malaise and fatigue (including chronic fatigue syndrome); dizziness, syncope, and convulsions; disorders of joint; and hypothyroidism. This list is also inclusive of the top 11 groups of comorbidities that were provided to RTI by MGH's CMP. Beneficiaries were identified as having a comorbid condition if they had one inpatient claim with the clinical condition as the principal diagnosis or had two or more physician or outpatient department (OPD) claims for an E&M service (CPT codes 99201-99429) with an appropriate principal or secondary diagnosis. The physician and/or OPD claims had to have occurred on different days. The diagnosis codes used to identify these clinical conditions are in ***Supplement 2A***.

- *Ambulatory Care Sensitive Conditions (ACSCs).* We constructed variables to indicate the presence of an ACSC in the year prior to the demonstration and during the demonstration, using the primary diagnosis on a claim. ACSCs include heart failure, diabetes, asthma, cellulitis, chronic obstructive pulmonary disease and chronic bronchitis, dehydration, bacterial pneumonia, septicemia, ischemic stroke, and urinary tract infection. The diagnosis codes used to identify these conditions are found in ***Supplement 2A***.

5) Utilization. We constructed three sets of utilization variables for this evaluation as proxies for intermediate clinical outcomes. These sets of variables were also constructed for the following principal diagnoses: all-cause and the 10 ACSCs, using the primary diagnosis (from the header portion of the claim) for claim types inpatient and outpatient:

- the number of acute hospitalizations,

- 90-day readmissions, and

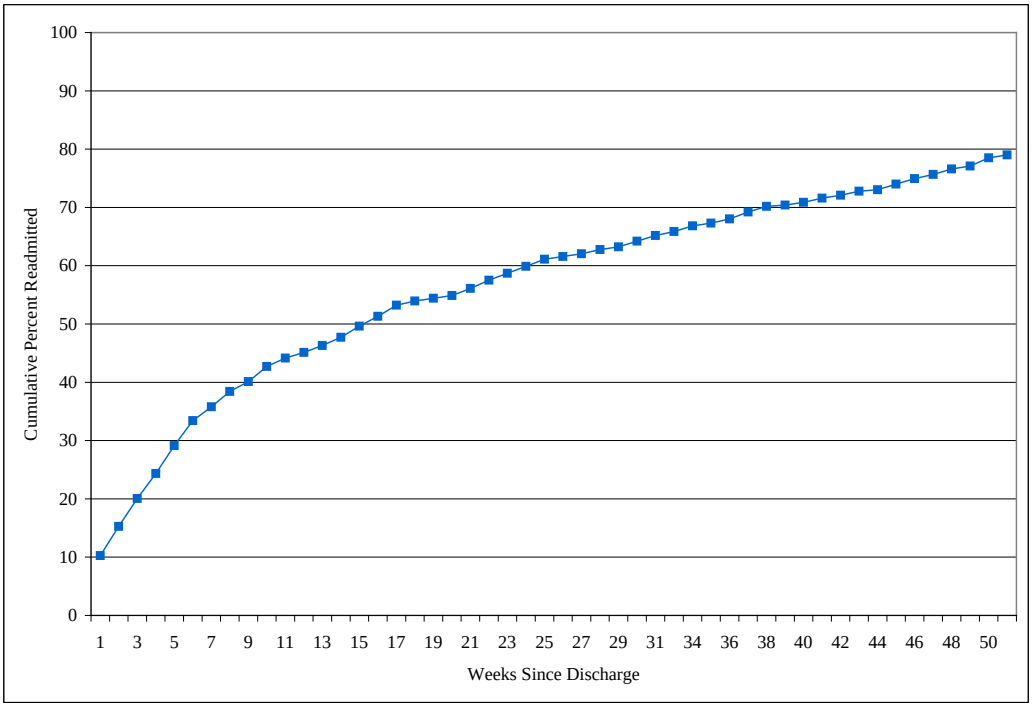
- emergency room visits, including observation bed stays.

Only claims that occurred during periods of eligibility were included in the utilization measures. For both the demonstration and baseline periods, claims were included if services were started during days that the beneficiary met MGH's CMP eligibility criteria, as determined from the ARC daily eligibility file. We flagged claims for services that occurred during a period of eligibility by comparing the eligibility period with a specific date on the claim, following the decision rules that were applied for the financial reconciliation. The exact date fields used are based on the claim type, as follows:

- inpatient and skilled nursing facility claims: *admission date;*
- all other types of services: *from date.*

Prior to conducting our final set of analyses, we critically examined the timing of readmissions using data from the year prior to the start of the demonstration. ***Figure 2-4*** displays a graphic representation of time from discharge to next admission for original population comparison beneficiaries who had a subsequent admission. In this figure, we display all-cause readmission; thus, beneficiaries were not required to have the same reason for both the initial and subsequent admission for the hospitalization to be considered a readmission. The graphic shows that there is a steep trajectory of readmissions during the first 90-day period following discharge, with a gradual tapering off of number of readmissions thereafter. Thus, we constructed 90-day readmission rates to capture close to 50% of subsequent admissions in our analyses[4].

Figure 2-4
Percent with readmission for any diagnosis: MGH's CMP original baseline comparison population



[4] We evaluated time to readmission based upon days post sentinel hospitalization discharge; however, the graph displays time to readmission in increments of weeks for visual presentation purpose.

We examined readmissions following admissions that occurred during two 12-month periods for the original population and one 12-month period for the refresh population. In order to capture readmissions following admissions that occurred late in the baseline and demonstration periods, we used a total of 15 months of data for each period to identify readmissions. For the baseline period, we identified admissions during the 12 months preceding the start of the demonstration and also included readmissions through the first 3 months of the intervention period for those admissions that occurred within 3 months of the start of the demonstration. The intervention periods for the original populations examined admissions during the periods of months 7 through 18 and months 22 through 33 and included readmissions through months 21 and 36, respectively. The intervention period for the refresh population examined admissions during months 10 through 21 and readmissions through month 24. A readmission was defined as an admission up to 90 days after an index hospitalization discharge date. We constructed all-cause readmission rates for all hospitalizations and same-cause readmission rates for the 10 ACSCs.

6) ***Expenditures***. RTI constructed a set of Medicare payment variables to reflect payments during periods of baseline and demonstration eligibility using the claims selection decision rules discussed previously. Total Medicare payments—exclusive of beneficiary deductibles, coinsurance payments, and third-party payments—were summarized for the annual period prior to the start date of the demonstration and also for the full intervention period and placed on a PBPM basis by dividing total payments by the total number of eligible days divided by 30.42. We defined a month as 30.42 days (365 days in a year divided by 12 months, rounded to two decimal places). This standardizes the definition of a month. For the demonstration period, total Medicare payments were summarized for the 36-month original intervention period and the 24-month refresh intervention period.

7) ***Guideline Concordant Care***. We define quality of care as adherence to evidence-based guideline-concordant care and have selected measures from the National Quality Forum (NQF)-endorsed National Voluntary Consensus Standards for Physician-Focused Ambulatory Care (February 2008). The selected measures are also used by other CMS pay-for-performance initiatives, such as the PQRI, or in evaluations of other pay-for-performance demonstrations (physician group practice demonstration) or pilot programs (Medicare Health Support). Thus, these measures have been extensively tested and are widely accepted as clinically important measures and appropriate for use in pay-for-performance initiatives. Further, we restrict the selection of measures to those that do not require the use of CPT II codes.

   First, we selected a measure that is broadly applicable to the Medicare fee-for-service population, influenza vaccination. Second, we selected several measures that are specific to beneficiaries with diabetes and heart failure as these populations are prevalent in MGH's CMP population and we subset the study populations to the appropriate clinical cohorts when constructing these measures.

The selected measures and relevant disease population are as follows:

- Rate of influenza shots for adults > 50 years (for patients with ESRD, the age is 18 years and older) – all beneficiaries
- Rate of annual HbA1c testing – beneficiaries with diabetes
- Rate of low-density lipoprotein cholesterol (LDL-C) testing – beneficiaries with diabetes
- Rate of low-density lipoprotein cholesterol (LDL-C) testing – beneficiaries with ischemic vascular disease

The method used to create these measures can be found in ***Supplement 2A***. CMS requested that we use existing, widely adopted specifications for evidence-based measures of care. Based on that request, RTI selected the National Quality Forum (NQF)–endorsed National Voluntary Consensus Standards for Physician-Focused Ambulatory Care. While the NQF-endorsed specifications restrict the diabetes quality-of-care measures to beneficiaries ages 18 to 75, we did not use this age restriction because no such restriction is used by MGH's CMP. The specifications used for the final set of analyses are from NQF-Endorsed™ National Voluntary Consensus Standards for Physician-Focused Ambulatory Care, Appendix A—National Committee for Quality Assurance (NCQA) Measure Technical Specifications, April 2008, V.7.

Claims for these process-of-care measures were included regardless of CMHCB demonstration eligibility in order to ensure that we fully captured the behavior of intervention and comparison populations that was not subject to Medicare eligibility or payment rules and to provide credit to MGH's CMP in case the services occurred after exposure to the CMHCB demonstration intervention and during the intervention period. One could envision that MGH's CMP encouraged the receipt of the process-of-care measures; however, the actual service was provided during a brief period of ineligibility (e.g., nonpayment of the Part B premium for a month). To the extent that the service was included in the Medicare claims files during a period of ineligibility as a denied claim, it reflects actual receipt of the service and was therefore included in our analyses.

8) Mortality. Date of death during the demonstration period was obtained from the Medicare EDB and was used to create a binary mortality variable.

9) Measures of CMHCB Program Intervention. Using the encounter data submitted by MGH's CMP, we constructed contact counts of the number of in-person visits/home visits, telephonic contact, referrals with subsequent visits, as well as total contacts.

CHAPTER 3
BENEFICIARY AND PHYSICIAN SATISFACTION

3.1 Beneficiary Satisfaction

The CMHCB demonstration programs' principal strategy to improve quality of care while reducing costs is by empowering Medicare beneficiaries to better cope with their chronic disease(s) and manage their care. The programs do this in three ways: (1) by enhancing beneficiary knowledge of their chronic condition through educational and coaching interventions, (2) by improving beneficiary communication with their care providers, and (3) by improving beneficiary self-management skills. Successful interventions should alter beneficiaries' use of medications, eating habits, and exercise, as well as promote more effective interaction with their primary health care providers. The CMHCB programs hypothesized that lifestyle changes and better communication with providers would mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions[5].

Specifically, MGH's disease management demonstration provides practice-based care management (PBCM) services using emerging information technology solutions to improve the quality of care delivered to high-cost Medicare FFS beneficiaries. This demonstration program follows the "High Performing Medicine Initiative" started in 2003, which involves practice-based care management services to high-cost Medicare beneficiaries. Case managers, who are assigned to each physician office, develop relationships with program participants to provide support across the continuum of care. Case managers provide patient education and connect patients with resources to address medical and psychosocial needs to help prevent acute exacerbations of disease and associated inpatient admissions and emergency room visits. The PBCM program also includes components to address mental health issues, evaluate complex pharmaceutical regimens, and support end-of-life decision making. In addition to improving the quality of care and outcomes for Medicare beneficiaries, the PBCM program aims to improve the quality of work life for primary care physicians and thereby ultimately attracting more physicians to the field of primary care.

The primary outcomes examined in the beneficiary survey were experience of care, self-management, and physical and mental function. We anticipated that the intervention's more intensive disease management activities would lead to greater levels of service helpfulness and greater self-efficacy. This in turn would increase the frequency with which intervention beneficiaries would engage in self-care activities, resulting in better functioning and higher satisfaction levels than in the comparison group. The same survey method and instrument was used across all six CMHCB demonstration programs for budgetary reasons. To isolate the

[5] In our survey, we examine satisfaction more broadly than satisfaction with a particular member of their health care team or a particular member of the MGH demonstration program team. We do so for the primary reason that we are asking the comparison population the same question and we desire to isolate the effect of the MGH intervention on the beneficiaries' assessment of satisfaction that their full health care team is helping them to cope with their chronic conditions.

intervention effects, the same survey instrument was administered to samples of beneficiaries from both the intervention and comparison groups. The findings from all six CMHCB beneficiary surveys have been reported to CMS previously (Smith et al., 2008).

3.1.1 Survey Instrument Design

The beneficiary survey was designed to obtain assessments directly from beneficiaries about key outcomes of beneficiaries' ***experience of care, self-management***, and ***physical and mental function.*** We asked beneficiaries about the extent to which their health care providers helped them to cope with their chronic conditions. We supplemented this item with questions related to two key components of the CMHCB interventions: helpfulness of discussions with their health care teams and quality of communication with their health care teams. In addition, the survey instrument collected information about beneficiary ***self-care*** frequency and ***self-efficacy*** related to medications, diet, and exercise and Clinician and Group Adult Primary Care Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

3.1.1.1 Measures of Experience and Satisfaction with Care

The impact of the interventions is critically dependent on the relationships between beneficiaries and their "health care teams" (defined as nurses, case managers, doctors, and/or pharmacists with whom they interacted, either in person or telephonically). The first set of survey measures assesses several dimensions of the interactions between beneficiaries and providers. These items were worded to be applicable to all beneficiaries, regardless of their intervention or participation status. As a result, questions referred to beneficiaries' health care teams rather than to the names of the CMOs.

Helping to cope with a chronic condition—The single item "How would you rate your experience with your health care providers in helping you cope with your condition?" provides an overall satisfaction rating. Ratings are made on a five-point scale (1 = poor, 2 = fair, 3 = good, 4 = very good, 5 = excellent).

Helpfulness of discussions with the health care team—This section addresses services received during the previous 6 months. Five types of services are addressed: (1) one-on-one educational or counseling sessions, (2) discussions about when and how to take medicine, (3) discussions about dealing with stress or feeling sad, (4) discussions about diet, and (5) discussions about exercise. The services could be provided through in-person visits, telephone calls, or mailings. Each service is rated on a four-point scale ranging from "very helpful" to "not helpful." A fifth response option identifies services that had not been discussed. Responses are summarized by counting the number of discussion topics rated as "very" or "somewhat" helpful so that the score for this item ranges from 0 (for no items helpful) to 5 (for all items helpful).

Discussing treatment choices—This item assesses a specific aspect of communication with providers by asking beneficiaries whether their health care team talks to them about pros and cons of their medical treatment or health care in general. Ratings are made on a four-point scale (1 = definitely no, 2 = somewhat no, 3 = somewhat yes, 4 = definitely yes).

Communication with health care team—Beneficiary communication is an important dimension of experience and satisfaction. Six communication items from the CAHPS® Survey were included in the questionnaire. These items assess how often the team (1) explained things in a way that was easy to understand, (2) listened carefully, (3) spent enough time with the beneficiary, (4) gave easy-to-understand instructions about what to do to take care of health problems, (5) seemed informed about up-to-date health issues, and (6) showed respect. Six frequency options (always, almost always, usually, sometimes, almost never, and never) are converted into CAHPS® composite scores ranging from 0 (never to all items) to a maximum of 100 (always to all items).

Getting answers to questions quickly—This measure includes two survey items that assess how quickly the health care team gets back to beneficiaries with answers to their medical questions. The questions ask how often beneficiaries received answers the same day during office hours or if they called after regular office hours, how often their questions were answered. Six frequency options (always, almost always, usually, sometimes, almost never, and never) are converted into composite scores ranging from 0 (never to all items) to a maximum of 100 (always to all items).

Medication support and information about treatment options—The Multimorbidity Hassles scale is designed to measure frustrating problems that patients experience in getting comprehensive care for chronic illnesses (Parchman, Noel, and Lee, 2005). Unlike disease-specific or physician-specific measures, this instrument was developed to apply broadly to patients with single or multiple conditions. Of the 16 items in the full scale, we selected the first six questions, which focus on problems with medications and treatment options. Example items are "lack of information about treatment options" and "side effects from my medications." Each item is rated on a five-point scale ranging from 0 = "no problem" to 4 = "a very big problem." The total Hassles score is the sum of the scores for the individual items and can range from 0 to 24. A higher score indicates more problems. Cronbach's alpha was 0.94 for the full scale. In the original development sample, the mean Hassles score for these six items was 5.86 (Parchman, Noel, and Lee, 2005).

3.1.1.2 Self-Management Measures

Patient self-management has been shown to be critical to health outcomes, particularly in chronic disease management (Hibbard et al., 2007). Chronic disease self-management interventions begin by helping patients set goals and make plans to address those goals and by helping patients manage their illnesses by practicing behaviors that may affect their health and well-being.

Setting health care goals—The question asks whether someone from the team had "helped you SET GOALS to take care of your health problems in the past 6 months." This item is answered either yes or no.

Making health care plans—A second yes or no item asks whether someone had "helped you MAKE A PLAN to take care of your health problems."

Self-efficacy—Self-efficacy refers to the confidence that one can perform health promotion activities. Previous research has shown that self-efficacy is a key determinant of

adherence to recommended behaviors, and self-efficacy expectations are a key target of many health care interventions. To assess self-efficacy, respondents were asked how sure they were that they could perform each of three specific behaviors: taking medications, planning meals according to dietary guidelines, and engaging in physical exercise. These items were drawn in part from the Confidence in Diabetes Self-Care Scale (Van Der Ven et al., 2003). Ratings are made on a five-point scale ranging from 1 = very unsure to 5 = very sure.

Self-care activities—A goal of chronic disease management is to promote patient compliance with self-care behaviors that may help to maintain or improve health status. Health-promoting behavior is assessed by the frequency with which beneficiaries engage in the same three self-care activities that are used to evaluate self-efficacy. These items were adapted from the Summary of Diabetes Self-Care Activities instrument (Toobert, Hampson, and Glasgow, 2000). Respondents indicate the number of days (0-7) in the past week that they performed each self-care activity.

3.1.1.3 Physical and Mental Health Function

Self-reported health status and function are important outcome measures that are not available through claims data. To assess the impact of the CMHCB demonstration on beneficiary function, the survey included two broad constructs: (1) physical and mental functioning and (2) activities of daily living. Here, we describe in detail how these constructs are measured.

Physical and mental function—Functioning levels were tracked by the responses to the Veterans RAND-12 (VR-12) instrument (Kazis, 2004). The VR-12 consists of 12 items, half of which reflect physical function and half of which are indicators of mental function. We used the RAND-12 scoring algorithm (Hays, 1998) to compute summary Physical Health Composite (PHC) and Mental Health Composite (MHC) scores. These scores are normalized so that the mean composite score is 50 (SD = 10) in the general U.S. adult population. Higher scores indicate higher levels of functioning. The scoring algorithm is based on Item Response Theory scaling yielding composite scores that may be correlated with one another. The algorithm also imputes scores for no more than one missing item in each composite.

Mental health status was also measured by the Patient Health Questionnaire-2 (PHQ-2), a widely used depression screening tool (Kroenke, Spitzer, and Williams, 2003). The PHQ-2 consists of two items: one for anhedonia ("How often have you been bothered by little interest or pleasure in doing things?") and one tapping depressed mood ("How often have you been bothered by feeling down, depressed, or hopeless?"). Each item is assessed in terms of weekly frequency (0 = not at all, 3 = nearly every day). The total PHQ-2 score is the sum of these values, which may range from 0 to 6 points. Higher scores indicate greater depressive symptoms. Scores of three points or more are commonly used in screening to identify cases that require further clinical evaluation.

Activities of daily living—A related measure of beneficiary functioning is the ability to perform basic activities of daily living (ADLs). The questionnaire collected information about six standard activities—bathing, dressing, eating, getting in and out of chairs, walking, and using the toilet. Respondents were first asked if they had any difficulty performing each activity. Possible responses were that they were unable to perform, had difficulty, or did not have difficulty doing the activity. They were then asked, with responses of yes or no, if they needed

help from another person to perform the activity. An ADL difficulty score was created by counting the number of activities that the beneficiary had difficulty with or was unable to do. The ADL help score was the number of activities for which the beneficiary needed help. Each score ranges from 0 to 6.

3.1.1.4 Background Characteristics

The final section of the questionnaire collected information about demographic characteristics such as race (Hispanic and African American status), educational attainment in years, living arrangements—whether beneficiaries lived alone or with a spouse or a relative—presence and type of health insurance coverage in addition to Medicare, and proxy information.

3.1.2 Analytic Methods

We conducted a series of statistical analyses to explore intervention-comparison differences and CMHCB demonstration intervention effects, including a response propensity analysis and descriptive and scaling analyses. We restrict our discussion in this report to the analyses associated with the outcomes variables.

3.1.2.1 Analysis of Covariance Model for Intervention Effects

We estimated weighted regression models to examine the effects the MGH CMP's interventions on the outcomes appearing in the conceptual model. The research design for this evaluation involved only a single round of the survey conducted during the demonstration period. Baseline levels of the individual study outcomes are not available. To increase the precision of the intervention effect estimates, we constructed multivariable regression models consisting of a broad set of beneficiary characteristics as explanatory covariates. Many of these covariates are drawn from claims data, while other background characteristics are reported in the survey questionnaire.

Two key indicators of initial status are the HCC risk score and PBPM expenditures. Both of these variables are measured for the year prior to the start of the demonstration. The following covariates are used:

- what demographic characteristics (age, gender, Hispanic ethnicity, African American, years of education) were,
- what Medicaid/dual eligible status was,
- whether the beneficiary lived alone,
- whether the beneficiary had health insurance coverage in addition to Medicare or Medicaid,
- whether the beneficiary used a proxy respondent, and
- whether the beneficiary completed a mail survey (versus a telephone survey).

Proxy and mail status are included to capture any systematic differences in responses that can be attributed to response mode. Previous research indicates that, compared with telephone surveys, mail surveys frequently elicit less favorable ratings of health status.

A general Analysis of Covariance (ANCOVA) model for the intervention analyses is

$$Y = a + b_1X_1 + b_kX_k + e,$$

where

Y = outcome measure;

X_1 = intervention status (1 = intervention, 0 = comparison);

X_k = a vector of k covariates;

b_1 and b_k = regression coefficients to be estimated;

a = an intercept term; and

e = an error term.

In this model, coefficient b_1 estimates the overall effect of the intervention in an intent-to-treat (ITT) analysis. The covariate coefficients correspond to direct effects of the mediating variables (e.g., communication with the health care team, self-management, and the helpfulness of health care services). Models in this general format were estimated separately for each CMO to test the impact of the program in each site. A logistic regression model consisting of the same set of covariates was used for dichotomous outcomes. The covariates in the model increase the precision of an intervention effect estimate by accounting for other sources of variation in the outcome measure. As described in ***Chapter 1***, the intervention and comparison beneficiaries were initially matched on either diagnostic status or Medicare expenditure levels. The covariate adjustments therefore control for other factors that may affect beneficiary outcomes and equalize any potential imbalances between the intervention and comparison groups when evaluating the impact of MGH's CMP.

3.1.2.2 Sampling Frame

The first step in the design process was to identify a sample frame for the survey in each of the six demonstration sites. Beneficiaries were eligible for the survey if (1) they were members of the starting intervention or comparison group populations and (2) they met the criteria for inclusion in quarterly monitoring reports at the time the frame was identified. Beneficiaries who met any of the exclusion criteria (death, loss of Part A or B coverage, enrollment in a Medicare Advantage plan, etc.) were ineligible for the survey frame. To maximize the number of eligible respondents in the frame, we performed a Medicare EDB run prior to sampling to identify decedents and other beneficiaries who had recently become ineligible.

3.1.2.3 Data Collection Procedures

We surveyed beneficiaries by mail with a telephone follow-up of nonrespondents. We used a multiple-mode, multiple-contact approach that has proved very successful on surveys conducted with the Medicare population and incorporates suggestions from Jenkins and Dillman's best mail survey practices guidelines (Jenkins and Dillman, 1997). Beneficiaries were surveyed once during the intervention period. The MGH CMP's survey was conducted between January 7, 2008 and May 4, 2008.

3.1.2.4 Sample Size, Statistical Power, Survey Weights, and Survey Response Rate

The target was 300 completed surveys for the intervention and comparison populations. From the sample frame for each group, we randomly selected 300/.7 = 429 beneficiaries. The response rate for MGH's CMP was 71%. The targeted sample size permits us to detect effect sizes (Cohen's d) of 0.23 or more for continuous outcome measures (power = .80, alpha = .05, two-sided tests). For a binary outcome, this is equivalent to the difference between percentages of 61% in the intervention group and 50% in the comparison group. The covariates in the ANCOVA models further increase the precision of coefficient estimates, allowing us to detect even smaller effects for many outcomes. Response weights were computed as the inverse of the probability of response predicted from each site's response propensity model. These weights were then rescaled to reflect the actual number of survey respondents.

3.1.3 Medicare Health Services Survey Results for MGH's CMP

This section presents the results of the Medicare Health Services Survey data analysis for MGH's CMP. We present the ANCOVA results with survey outcomes organized into three domains: beneficiary experience and satisfaction with care, self-management, and physical and mental functioning. Overall, we present results for 19 survey outcomes.

3.1.3.1 Experience and Satisfaction with Care

The primary measure of satisfaction was a rating of experience with health care providers to help the beneficiary cope with his or her condition. The survey also included five other measures of satisfaction with care experience. ***Table 3-1*** displays the satisfaction and experience with care measures for MGH's CMP.

Overall experience: helping beneficiary to cope with chronic condition— The average score for the key satisfaction outcome item that assessed how well the health care team helped beneficiaries cope with their illness was 4.0 for the intervention group, or about midway between "very good" and "good" ratings. The average score for the comparison group was about 3.9. Just over fifty percent of MGH CMP beneficiaries rated their experience as "excellent" or "very good" and about 26% selected "good." It is not uncommon among the elderly to report high satisfaction ratings. For that reason, the mean scale score was used in the analyses so that transitions between all response categories would be captured. For this overall satisfaction measure, we observe no statistically significant intervention effect for MGH's CMP. MGH's intervention was not found to improve beneficiary overall satisfaction in helping them cope with their chronic illness.

Table 3-1
Medicare Health Services Survey: Estimated intervention effects for experience and satisfaction with care, MGH's CMP
(N = 590)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Helping to cope with a chronic condition (1 to 5)	3.97	3.86	0.16	N/S
Number of helpful discussion topics (0 to 5)	2.40	2.45	-0.02	N/S
Discussing treatment choices (1 to 4)	3.29	3.17	0.26	**
Communicating with providers (0 to 100)	80.5	77.8	4.5	*
Getting answers to questions quickly (0 to 100)	70.5	65.4	5.0	N/S
Multimorbidity Hassles score (0 to 24)	2.78	3.22	-0.27	N/S

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; ANCOVA = Analysis of Covariance.

Statistical significance (Stat. sig.): * Indicates significance at the 5% level; ** Indicates significance at the 1% level; otherwise N/S means not statistically significant.

SOURCE: RTI analysis of the Medicare Health Services Survey, 2008. Computer program: CreqD2

Number of helpful discussion topics— For this item, beneficiaries were asked to evaluate five types of services (1) one-on-one educational or counseling sessions, (2) discussions about when and how to take medicine, (3) discussions about dealing with stress or feeling sad, (4) discussions about diet, and (5) discussions about exercise. The mean number of services for which beneficiaries had helpful discussions with their health care team was comparable between the intervention and the comparisons groups (2.4 and 2.5 respectively). For this measure, we observe no statistically significant intervention effect for MGH's CMP.

Discussing treatment choices— For this item, beneficiaries were asked whether health care team talked about pros and cons of each treatment choice with answers ranging from 1 "definitely no" to 4 "definitely yes". The mean score for the intervention group was 3.3, compared to 3.2 for the comparison group, resulting in a statistically significant ANCOVA

adjusted intervention effect. MGH's CMP beneficiaries in the intervention group rated this outcome higher than those in the comparison group.

Communication with health care team—The score for communication with health care team could range from 0 to 100, with 0 indicating never to all items in the composite and 100 indicating always to all items in the composite. Both MGH's CMP intervention and comparison groups reported high average communication scores with 80.5 for the intervention group and about 77.8 for the comparison group. This difference is statistically significant suggesting that the MGH CMP intervention improved perceived beneficiary communication with providers.

Getting Answers to Questions Quickly. The score for getting answers to questions quickly could range from 0 to 100, with 0 indicating never to both items in the composite and 100 indicating always to both items in the composite. MGH's CMP was associated with high average scores for intervention and comparison groups (70.5 and 65.4, respectively). For this measure, the difference between the intervention and comparison groups was not found to be statistically significant.

Multimorbidity Hassles Scale. Multimorbidity Hassles Scale, designed to measure frustrating problems that patients experience in getting comprehensive care for chronic illnesses, is measured on a scale from 0 to 24. High scores indicate more problems. MGH's CMP showed relatively low Multimorbidity Hassles scores for intervention and comparison groups (2.8 and 3.2, respectively). In the original development sample, the mean Hassles score for these six items was 5.86 (Parchman et al., 2005). For this measure, we observe no statistically significant intervention effect for MGH.

In summary, across the 6 measures of experience and satisfaction with care, we observe two statistically significant intervention effects for MGH's CMP. Interventions conducted by MGH were associated with statistically significantly higher ratings on discussing treatment choices and communicating with health providers compared to the comparison group.

A goal of chronic disease management is to improve compliance with self-care activities that may slow the decline in functioning and health status. The survey included three sets of questions related to self-management: receiving help with setting goals and making a care plan, self efficacy ratings, and self-care activities. ***Table 3-2*** displays the self-management measures for MGH's CMP.

Setting goals and making a care plan—The survey included two questions that asked if someone from their health care team helped set goals or a plan to take care of their health problems. For MGH CMP beneficiaries in the intervention group, 57% reported receiving help setting goals and help making a care plan each. In the comparison group, 64% and 58% respectively reported receiving assistance on these self-management activities. The ANCOVA results reveal that these differences are not statistically significant: MGH's CMP did not have a higher proportion of intervention beneficiaries who received help setting goals for self-care management, nor was it associated with a reported increase in providing help in making health care plans.

Table 3-2
Medicare Health Services Survey: Estimated intervention effects, self-management, MGH's CMP (N = 590)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Percent receiving help setting goals	57.1	64.4	-.5.6	N/S
Percent receiving help making a care plan	57.2	57.8	2.3	N/S
Self-efficacy ratings				
Take all medications (1 to 5)	4.62	4.51	0.05	N/S
Plan meals and snacks (1 to 5)	4.17	4.11	0.01	N/S
Exercise 2 or 3 times weekly (1 to 5)	3.69	3.49	0.11	N/S
Self-care activities				
Prescribed medications taken (mean # of days)	6.65	6.69	-0.10	N/S
Followed healthy eating plan (mean # of days)	4.81	4.91	-0.16	N/S
30 minutes of continuous physical activity (mean # of days)	2.85	2.70	0.05	N/S

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; ANCOVA = Analysis of Covariance.

* Indicates significance at the 5 percent level. ** Indicates significance at the 1 percent level.

N/S means not statistically significant.

SOURCE: RTI analysis of the Medicare Health Services Survey, 2008.

Computer program: CreqD2

Self-efficacy ratings—To assess self-efficacy, respondents were asked how sure they were that they could perform each of three specific behaviors: taking medications, planning meals according to dietary guidelines, and engaging in physical exercise. Ratings are made on a five-point scale ranging from 1=very unsure to 5=very sure. Overall, MGH's CMP beneficiaries typically reported relatively high levels of self-efficacy with mean ratings averaging around 4 (somewhat sure of their ability to perform self-care activities) out of a maximum of 5 (very sure). The highest self-efficacy scores were reported for taking medications as prescribed, and the lowest scores were for getting exercise two or three times per week. On average, MGH's CMP beneficiaries in both groups were quite sure they could take their medications as often as prescribed, 4.6 rating for the intervention group compared to 4.5 in the comparison group. Confidence in planning meals and snacks was rated 4.2 and 4.1, respectively, and confidence in

exercising was rated as 3.7 and 3.6, respectively. None of the self-efficacy items yielded statistically significant ANCOVA-adjusted intervention effects.

Self-care activities—A goal of chronic disease management is to promote patient compliance with self-care behaviors and activities that may help to maintain or improve health status. The questionnaire included questions about three self-care behaviors that parallel the items in the self-efficacy ratings. Self-care activities are measured in the number of days in the past week when beneficiaries were compliant and range from 0 to 7. The reported compliance rate for self-care activities ranged from quite high for both groups among some activities (taking medications) to more modest compliance rates for another activity (exercise). For example, the mean number of days that MGH's CMP beneficiaries said they take their medications as prescribed was 6.7 for both groups; the mean number of days that MGH's CMP beneficiaries reported following a healthy eating plan ranged between 4.8 to 4.9, and the mean number of days MGH's CMP beneficiaries reported exercising was 2.7-2.9 days out of 7. There were no significant group differences in the rates for any of these three self-care activities between the intervention and the comparison groups.

Physical and mental function—***Table 3-3*** displays the mental and physical functioning outcomes for MGH. On average, MGH's CMP respondents had the mean PHC score for the intervention group of 32.6, significantly higher when compared to 29.9 for the comparison group, producing a statistically significant ANCOVA intervention effect at the 1 percent level. The mean MHC score for the intervention group was 39.5 and the PHQ-2 score of 1.7, compared to 38.3 and 1.9 for the comparison group. Both mental health function outcome differences were not statistically significant.

Activities of daily living—On average, respondents in MGH's CMP intervention group reported being limited on 1.9 ADLs compared to 2.2 ADLs for the comparison group. MGH's CMP beneficiaries also reported received help with an average of 0.6 to 0.9 ADLs. ANCOVA results indicate that there was no statistically significant difference in functional status (both difficulty and receipt of help measures) between the intervention and the comparison groups.

3.1.4 Conclusions

The CMHCB demonstration employs strategies to improve quality of care for high cost Medicare beneficiaries while reducing costs by empowering Medicare beneficiaries to better manage their care by insuring daily communication between the participant and health care system. Specifically, MGH's CMP provided practice-based care management and patient education and connected patients with resources to address medical and psychosocial needs. MGH's CMP also included components to address mental health issues, evaluate complex pharmaceutical regimens, and support end-of-life decision making. The MGH demonstration staff hypothesized that better communication with providers will mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions.

Table 3-3
Medicare Health Services Survey: Estimated intervention effects, physical and mental health function, MGH's CMP
(N = 590)

Outcome	Intervention mean	Comparison mean	ANCOVA-adjusted intervention effect	Stat. sig.
PHC score (physical health, mean =50, std=10)	32.6	29.9	2.3	**
MHC score (mental health, mean =50, std=10)	39.5	38.3	1.1	N/S
PHQ-2 score (depression, 0 to 6)	1.73	1.87	-0.03	N/S
Number of ADLs difficult to do (0 to 6)	1.91	2.24	-0.28	N/S
Number ADLs receiving help (0 to 6)	0.63	0.90	-0.21	N/S

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; ANCOVA = Analysis of Covariance; PHC = Physical Health Composite; MHC = Mental Health Composite; PHQ-2 = Patient Health Questionnaire 2; ADLs = activities of daily living.

* Indicates significance at the 5 percent level. ** Indicates significance at the 1 percent level.

N/S means not statistically significant.

SOURCE: RTI analysis of the Medicare Health Services Survey, 2008.

Computer program: CreqD2

Among the 19 outcomes covered by the survey, we found three statistically significant ANCOVA-adjusted intervention effects, two in the experience and satisfaction with care domain and one in the physical and mental health function domain. Survey results indicate that members of MGH's CMP intervention group were more satisfied with the discussion of their treatment choices and rated their communication with heath providers higher than MGH's CMP beneficiaries in the comparison group. These achievements, however, were not translated into any improvements in self-efficacy or in self-care activities. In addition, MGH's CMP beneficiaries in the intervention group reported significantly higher PHC scores than those in the comparison group, suggesting that the intervention improved physical functioning of participating Medicare beneficiaries. No other statistically significant outcomes were found in the physical and mental health function domain.

3.2 Provider Satisfaction

RTI conducted two site visits to MGH's CMP and spoke with a small number of primary care physicians during each site visit to gauge their assessment of satisfaction with the demonstration program.

At the time of the first site visit, a small number of physicians expressed concerns about the program. For example, they had questions about whether CMP patients would divert services from other patients in their practices. And, some physicians did not have a full understanding of the role of the case managers. However, as physicians gained experience working with the case managers, the most common concern they voiced was frustration about their inability to include additional patients in the program. One provider noted that for each patient eligible for the program, there are two additional patients in the practice who could benefit from such case management support.

At the time of the second site visit, physicians gathered for the focus group reported great overall satisfaction with the CMP. Below is the summary of the physicians' input:

- "The program 'wraps its arms' around the most difficult and complex patients, those who are poor, with a lot of mental health issues;
- The program signifies a move towards a true medical home model—it is run by a team of providers;
- Case managers take care of things like preauthorization, gathering documentation, medication tracking and other time-consuming issues, allowing PCPS to focus on the relationship with patients and provide real continuity of care;
- The program does what every PCP needs to be doing but cannot do anymore because of the medicine practice and reimbursement realities and primary care provider shortages;
- The program assists with complex needs, including social issues and transportation, that are key for this population;
- Social workers are very proactive and are instrumental in gaining trust and compliance from difficult patients;
- The program only works well when physicians are highly engaged;
- Both patients and physicians love the program as case managers take a lot of burden off both sides;
- Key value of the program is in the help they provide PCPs with medication review and management, the most difficult to resolve issue when PCPs do not have any help;

- The program has done a remarkable job in training and cultivating case managers who are very good at breaking barriers and making it work for the most difficult patients;

- The program is a tremendous asset for the hospital, and a model for the future of primary care delivery;

- Case managers are instrumental in helping participants discuss advanced care planning- accepting and understanding the issues around end of life care, hospice and other difficult issues;

- The program is particularly helpful for homebound persons, sometimes leading to drastic positive changes in the their lives;

- We do not want the program to end- it is very valuable! Once the program is gone, participants will become 'frequent flyers' in emergency department and in the hospital in general;

- Fantastic program, truly saving lives."

Also at the second site visit, MGH's CMP leadership shared with us the results of a physician satisfaction survey they had conducted. In November 2007, the CMP evaluated the level of satisfaction of involved primary care physicians (PCPs). The online survey was emailed to 148 PCPs and leaders within the program, 60% (n=89) of which responded to the survey. The survey addressed a number of topics:

- The overall satisfaction of the PCPs with the Care-Management Program;

- Impacts of the program on the PCPs' practice; and

- The perceived impact the program has on the lives of patients.

The PCPs response to the program was quite positive. Overall, there was satisfaction with the program, with 67% (n=60) of the PCPs agreeing (strongly agree or somewhat agree) that the Care Management Program improved their quality of practice. Additionally, 73% (n=65) of the PCPs favorably agreed (strongly agree or somewhat agree) that the addition of a case manager to their practice positively affected the quality of care for their patients.

CHAPTER 4
PARTICIPATION RATES IN MGH'S CMP AND LEVEL OF INTERVENTION

4.1 Introduction

Our participation analysis is designed to critically evaluate the level of engagement by MGH's CMP in this population-based demonstration and to identify any characteristics that systematically predict participation versus nonparticipation. Furthermore, we seek to evaluate the degree to which beneficiaries who consented to participate were exposed to the MGH CMP's interventions. The analyses are designed to answer a broad policy question about the depth and breadth of the reach into the community: how well did MGH's CMP engage their intended audiences? Specific research questions include the following:

- Were there systematic baseline differences in demographic characteristics and disease burden between the intervention and comparison group beneficiaries at the start of the demonstration?

- How many individuals did MGH's CMP engage, and what were the characteristics of the participants versus nonparticipants (in terms of baseline clinical measures, demographics, and health status)?

- What beneficiary characteristics predict participation in MGH's CMP?

- To what extent were the intended audiences exposed to MGH's CMP interventions? To what extent did participants engage in the various features of the program?

- What beneficiary characteristics predict a high level of MGH's CMP intervention versus a low level of intervention?

The overall design of the CMHCB demonstration follows an intent-to-treat (ITT) model, and all CMOs are held at risk for their monthly management fees based on the performance of the full population of eligible beneficiaries randomized to the intervention group and compared with all eligible beneficiaries in the comparison group. The CMHCB demonstration has been designed to provide strong incentives to gain participation by all eligible beneficiaries in the intervention group. At our March 2007 site visit, MGH's CMP staff reported that 2,151 beneficiaries had consented to participate in their program (about 90% of MGH's intervention population). Approximately 100 beneficiaries had not refused communication with CMP staff but were still considering whether to join the program. The remaining beneficiaries declined to participate, did not have a primary care physician in Partners based on patient self-report, moved out of Massachusetts, or were unreachable (Brody and Bernard, 2007). By September 2008, MGH's CMP had 2,401 participating beneficiaries, with 1,807 from the original cohort and 594 from the refresh cohort (Khatutstky and McCall, 2010). In our first analysis of participation in the CMHCB demonstration, we examined participation during the initial 6-month outreach period of the demonstration (McCall et al., 2008). In this report, we examine the level of participation for the full intervention period and the beneficiary characteristics that predict participation.

We also examine the level of intervention between MGH's CMP and its beneficiaries. The core element of MGH's CMP is the one-on-one relationship between patients and their practice-based case managers, supplemented by support received from the program's mental health, pharmacist, and end-of-life components. During the first 6 months of the demonstration period, case managers focused their efforts almost exclusively on beneficiary outreach and enrollment tasks. Therefore, case management activities began in earnest on February 1, 2007. Thus, we only examine encounter data beginning in month 7 for both the original and refresh populations. For each participating beneficiary, MGH's CMP provided RTI with a count of the number of telephonic contacts to a patient or caregiver, number of office visits by type (e.g., psychiatric, MSW, or end-of-life (EOL)), in-home visits by a physician or nurse practitioner, in-person contact with a patient or caregiver by the case manager, and indicators of participants with a telemonitoring device.

4.2 Methods

4.2.1 Participation Analysis Methods

We determined participation status during the demonstration period using a monthly indicator provided to us by ARC in the *Participant Status* file to align with dates of eligibility for MGH's CMP. We report the percentage of intervention beneficiaries who consented to participate for at least 1 month during the intervention period as well as those who never consented to participate and the reason for nonparticipation (refused or never contacted/unable to be reached). We also report the percentage of beneficiaries who, after initial consent, were continuous participants (while eligible for MGH's CMP) and the percentage of beneficiaries participating for more than 75% of their eligible months.[6] These latter two sets of numbers provide an estimate of the number of beneficiaries with whom MGH's CMP had the greatest opportunity to intervene. Because beneficiaries lose eligibility for various reasons over time (e.g., loss of Part A or Part B benefits, or due to death), we report counts of full-time equivalents (FTEs) or numbers of intervention and comparison beneficiaries weighted by the fraction of the demonstration period each beneficiary was eligible. Only beneficiaries who had at least 1 day of eligibility in both the baseline and demonstration periods are included in these analyses.

We also conduct a multivariate logistic regression analysis to determine the predictors of participation versus nonparticipation among those in the intervention group. The logistic model used in this study to identify differences in the likelihood of a beneficiary being in the participant group versus the nonparticipant group as a function of baseline and intervention period clinical factors, baseline cost, and baseline demographic factors is specified as

$$\text{Log e}\,(p_i / [1 - p_i]) = \beta X_i + \text{error}, \qquad (4\text{-}1)$$

where P_i = the probability that the *i*th individual will consent to participate, βX_i = an index value for the *i*th individual based on the person's specific set of characteristics (represented by

6 A beneficiary becomes ineligible to participate if he/she enrolls in a Medicare Advantage (MA) plan, loses eligibility for Part A or B of Medicare, moves out of the demonstration area, has a new primary payer (i.e., Medicare becomes secondary payer), develops ESRD, elects the hospice benefit, or dies.

the vector), and e = the base of natural logarithms. The probability of a beneficiary being in the participant group is thus explained by the variables.

Logistic regression produces an odds ratio for every predictor variable in the model; that is, an estimate of that variable's effect on the dependent variable, after adjusting for the other variables in the model. The odds ratio is greater than 1.0 when the presence (or higher value) of the variable is associated with an increased likelihood of being in the participant group versus the nonparticipant group; odds ratios less than 1.0 mean that the variable is inversely associated with being in the participant group.

We estimate three participation regression models to allow for evaluation of whether characteristics of participation differed across time (first 6 months versus the full intervention period) and across levels of participation (at least 1 month versus at least 75% of eligible months). The participation model investigates whether group membership is influenced by beneficiary demographic attributes, clinical characteristics, and utilization and cost factors previously defined in ***Chapter 2***. The demographic variables included in the model are defined as follows from the Medicare EDB and determined as of the date of assignment for the original population (July 4, 2006) and the refresh population (July 18, 2007):

- male, a dichotomous variable, set at 1 for males;
- African American/other/unknown, a dichotomous variable, set at 1 for beneficiaries whose race code is African American, other, or unknown;
- aged-in, a dichotomous variable, set at 1 for beneficiaries whose entitlement to Medicare benefits is based on age rather than disability;
- age, three dichotomous variables set at 1 for age less than 65 years, age 75-84, and age greater than or equal to 85 years; age 65-74 is the reference group; and
- Medicaid, a dichotomous variable, set at 1 for beneficiaries enrolled in Medicaid. Medicaid enrollment is based on a beneficiary being enrolled in Medicaid at any point 1 year prior to the go-live date.

Baseline clinical and financial characteristics included in the model are defined as follows:

- baseline HCC score medium and high, two dichotomous variables set at 1 if the prospective HCC score was between 2.0 and 3.1 (medium) and greater than 3.1 (high); HCC score less than 2.0 is the reference group;
- baseline Charlson score medium and high, two dichotomous variables set at 1 if the Charlson index score was equal to 2, 3, or 4 (medium) and 5 or greater than (high); Charlson score of less than 2 is the reference group for the original population. For the refresh population, baseline Charlson scores of 2 or 3 were medium and 4 or greater were in the high group. The reference group was a score of less than 2.

- baseline PBPM costs medium and high, two dichotomous variables set at 1 if the PBPM cost calculated by RTI for a 12-month period prior to the *start* of the MGH's CMP original demonstration program was greater than or equal to $336.70 and less than $1,539 (medium) and $1,539 or greater (high); PBPM costs less than $336.70 is the reference group for the original population. For the refresh population, baseline PBPM costs greater than or equal to $324 and less than $1,408 were assigned to the medium group and $1,408 or greater to the high category; PBPM costs less than $324 is the reference group.

Intervention period beneficiary characteristics included in the model are defined as follows:

- died, a dichotomous variable, set at 1 for beneficiaries who died during the intervention period;

- institutionalized, a dichotomous variable, set at 1 for beneficiaries who were resident in a long-term care setting for any 1 or more months of the initial 6 months of the intervention period; and

- concurrent HCC score medium and high, two dichotomous variables set at 1 if the concurrent HCC score calculated by RTI for the initial 6-month original intervention period was greater than 0.811 but less than 2.013 (medium) and greater than or equal to 2.013 (high); concurrent HCC score less than or equal to 0.811 is the reference group. These scores were re-calculated for the first 6-months of the refresh intervention period with the medium category assigned to values greater than 0.694 but less than 1.693 and values greater than or equal to 1.693 were assigned to the high category; a concurrent HCC score less than or equal to 0.694 is the reference group.

4.2.2 Level of Intervention Analysis Methods

MGH's CMP provided RTI with the number and nature of contacts with participating beneficiaries at the beneficiary level from August 2006 through the end of Phase I of their CMHCB demonstration program. We use these data to develop estimates of the level of intervention provided to MGH's CMP participants. The core element of MGH's CMP is the one-on-one relationship between patients and their practice-based case managers, supplemented by support received from the program's mental health, pharmacist, and end-of-life components. Case managers develop relationships with patients over time through telephone calls and in-person interactions during physician office visits or at the hospital, if they are admitted for an inpatient service. Case managers also conduct visits to patient homes on an as-needed basis (Brody and McCall, 2006). MGH's CMP also had a mental health program to screen the entire intervention population for depression, substance abuse, and dementia and refer beneficiaries to a psychiatrist as needed. Another goal was to facilitate earlier access to end-of-life (EOL) resources and create a seamless transition between hospital and home care for patients approaching the end of life.

Using the encounter data submitted by MGH's CMP, we constructed counts of the number of telephonic contacts with participants, office visits with mental health specialists, social workers, or EOL consultants, in-person contacts, home visits, and the number of active

telemonitoring devices. We report the mean and median number of total contacts and the distribution of beneficiaries across six categories of contacts (0, 1, 2-4, 5-9, 10-19, and 20 or more). We also estimate a multivariate logistic regression model of the likelihood of being in the high total contact category relative to the low total contact category. A dichotomous dependent variable was created and set at 1 for beneficiaries who had a high level of contact with MGH's CMP and 0 for beneficiaries who had a low level of contact. Beneficiaries, who had a medium level of contact with MGH's CMP, were the reference group in the regression analysis. Independent variables in the contact regression model included those that we have described for the participation regression model and two additional demonstration period utilization measures:

- one intervention period hospitalization set at 1 if the beneficiary had one hospitalization in months 7-18 for the original population and months 13-24 for the refresh population; and

- multiple intervention period hospitalizations set at 1 if the beneficiary had more than one hospitalization during the same time periods.

We included these two additional demonstration period intervention variables because MGH's CMP staff attempted to identify beneficiaries at risk of a hospitalization and to intervene to prevent the hospitalization from occurring or to identify beneficiaries at the time of hospitalization or shortly thereafter to intervene to prevent readmission. Thus, we would expect these two variables to be positively associated with being in the high contact group.

4.3 Findings

4.3.1 Participation Rates for MGH's CMP Population

Analyses presented in this section include only beneficiaries who had at least 1 day of eligibility in the year prior to the start of the intervention period and at least 1 day of eligibility in the demonstration. The results are based on the full demonstration period for both the original and refresh populations. The number of months for the full demonstration period for MGH's CMP is 36 months for the original population and 24 months for the refresh population.

Tables 4-1 and 4-2 display the number of beneficiaries included in our participation analyses for the original and refresh populations and illustrates the impact of loss of eligibility by reporting the FTEs. We report

1. Number of beneficiaries. The number of beneficiaries is equal to all beneficiaries who had at least 1 day of eligibility in the 1-year baseline period and had at least 1 day of eligibility in the period tabulated.

2. Full-time equivalents. FTEs defined as the total number of beneficiaries weighted by the number of days eligible in the intervention period divided by the total number of days in the intervention period. For example, a beneficiary in MGH's CMP had a total of 36 months (or 1,096 days) of possible enrollment. If he/she died after 90 days, their FTE value would be 90/1,096 or 0.082 FTEs. If someone were eligible for all 36 months, then his or her value is 1. The sum of this value across all beneficiaries gives the total FTE value reported.

3. Number fully eligible. The number fully eligible is the number of beneficiaries that had no gap in MGH's CMP eligibility during the demonstration period.

The ratio of FTEs to the total number of eligible beneficiaries in the original intervention population is 0.83 for the entire intervention period (months 1-36) compared with higher ratios (0.93 and greater) for each individual year of the demonstration. These differences in ratios illustrate the effect of subsetting to beneficiaries in the different time periods and attrition over time of the original beneficiaries due primarily to death. Beneficiaries also became ineligible for participation in MGH's CMP if they joined a Medicare Advantage (MA) plan, lost Medicare Part A or B eligibility or Medicare became a secondary payer, developed ESRD, or moved out of the service area.

Thirty-four percent of the original intervention and 39% of comparison beneficiaries had a spell of ineligibility. This can be estimated as the difference in the number of eligible beneficiaries and the number of fully eligible beneficiaries. Within the intervention group, eligibility was higher for participants and lower for nonparticipants. The MGH CMP's nonparticipant group was eligible only 62% of all possible days—much lower than the 85% of days for participants. Also, the participant group had a higher rate of beneficiaries being fully eligible for the entire intervention period (68%) compared with 49% for the nonparticipant group.

Table 4-2 displays eligibility data for the refresh population, which is about one-third the size of the original population. The ratio of total number of beneficiaries to FTEs was lower for the full 24 months (0.90) compared to the two 12-month periods (0.95) for the intervention population. This held true for the comparison population as well. The percent of beneficiaries that were fully eligible for the full refresh time period is higher among participants (83%) than nonparticipants (69%) or the comparison group (74%), but the difference narrows by the last 12 months of the demonstration (91%, 90%, and 87%, respectively).

Table 4-1
Number of Medicare FFS beneficiaries eligible for and participating in MGH's CMP: Original population

Characteristics	Months 1-36	Months 1-12	Months 13-24	Months 25-36
Intervention group				
Number eligible[1]	2,584	2,584	2,273	1,994
Full time equivalent[2]	2,137	2,425	2,123	1,863
Number fully eligible	1,705	2,263	1,974	1,726
Participants				
Number eligible	2,280	2,250	2,031	1,765
Full time equivalent	1,949	2,173	1,900	1,653
Number fully eligible	1,556	2,045	1,768	1,530
Participants > 75%				
Number eligible	1,985	1,196	2,010	1,760
Full time equivalent	1,803	1,167	1,880	1,648
Number fully eligible	1,486	1,108	1,749	1,525
Nonparticipants				
Number eligible	304	334	242	229
Full time equivalent	188	252	223	209
Number fully eligible	149	218	206	196
Comparison group				
Number eligible	2,719	2,714	2,307	1,979
Full time equivalent	2,153	2,500	2,121	1,839
Number fully eligible	1,658	2,268	1,936	1,695

NOTES:

FFS = fee-for-service; MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

SOURCES: Medicare claims data, Medicare enrollment database.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-1.sas 27JUN2010.

Table 4-2
Number of Medicare FFS beneficiaries eligible for and participating MGH's CMP: Refresh population

Characteristics	Months 1-24	Months 1-12	Months 13-24
Intervention group			
Number eligible[1]	775	775	691
Full time equivalent[2]	695	734	656
Number fully eligible	623	685	626
Participants			
Number eligible	648	646	586
Full time equivalent	598	629	559
Number fully eligible	536	591	532
Participants > 75%			
Number eligible	574	216	584
Full time equivalent	556	212	557
Number fully eligible	518	198	530
Nonparticipants			
Number eligible	127	129	105
Full time equivalent	97	105	98
Number fully eligible	87	94	94
Comparison group			
Number eligible	766	766	658
Full time equivalent	660	708	612
Number fully eligible	570	655	573

NOTES:

FFS = fee-for-service; MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

SOURCES: Medicare claims data, Medicare enrollment database.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-1.sas 27JUN2010.

Tables 4-3 and 4-4 present participation rates for MGH's CMP original and refresh populations and display the participation status of the beneficiary after verbal consent to participate was given (continuous participation, became a continuous nonparticipant after initial participation period, or intermittent participation). We also display the reasons for nonparticipation and the percent of beneficiaries who participated more than 75% of eligible months. Numbers of participants by selected months are also reported. Continuous versus intermittent participation is important because it affects the ability of MGH's CMP to contact beneficiaries and, ultimately, have any impact on utilization and costs.

Participation rates for MGH's CMP original population. Of all MGH's CMP original intervention group beneficiaries, 88% verbally consented to participate in its program at some point during the intervention period. We previously reported (Brody and Bernard, 2007) that MGH had an ambitious goal to obtain 100% participation among beneficiaries eligible for its demonstration program. As of March 16, 2007, 2,151 beneficiaries consented to participate in the CMHCB program—this represents 83% of MGH's intervention population and we observe a slight increase in MGH's CMP's enrollment over the entire intervention period. Sixty-six percent of beneficiaries were continuous participants (***Table 4-3***), which equates to 75% of participants. Among MGH's CMP beneficiaries, 6% refused to participate. The percent not contacted or unable to be located was also 6%.

Participation rates were heavily influenced by length of eligibility during the intervention period. An alternative measure of participation is the percentage of beneficiaries who participated more than 75% of months they were eligible for the demonstration. Of MGH's CMP original intervention beneficiaries, 77% participated for more than 75% of their eligible months, which is much higher than the continuous participant percentage. ***Table 4-3*** also reports the number of participants over time (for months 6, 12, 24 and 36, the last month of the demonstration). The number of participants declined over time as would be expected given the attrition due to loss of eligibility primarily due to death.

Participation rates for MGH's CMP refresh population. The criteria for selection of the intervention and comparison refresh populations were similar to the criteria used to select the initial populations. The refresh population had a slightly lower participation rate (***Table 4-4***). Overall, 84% of the refresh intervention beneficiaries consented to participate at some point during the 24-month period. Of those, 72% were continuous participants, which equates to 86% of participants. The percent that refused to participate was modestly lower (4%), and the percent not contacted or unable to be contacted was higher at 13%. Roughly the same percentage of refresh intervention beneficiaries participated for more than 75% of their eligible months that we observed for the original intervention beneficiaries.

Table 4-3
Participation in MGH's CMP: Original population

Characteristics	Statistic
Number of intervention months	36
Participation rate (entire demonstration period)	88%
Length of participation	
Continuous participation after engagement	66%
After initial participation, became a continuous nonparticipant	20%
Intermittent participation	2%
Nonparticipation (never agreed)	12%
Refused to participate when contacted	6%
Not contacted/unable to be contacted	6%
Beneficiaries participating more than 75% of eligible months	77%
Number of participants in selected months[1]	
Month 6	2,121
Month 12	2,036
Month 24	1,777
Month 36 (last month)	1,533

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

Data Sources: Medicare claims data, Medicare enrollment database.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-2.sas 27JUN2010.

Table 4-4
Participation in MGH's CMP: Refresh population

Characteristics	Statistic
Number of intervention months	24
Participation rate (entire demonstration period)	84%
Length of participation	
Continuous participation after engagement	72%
After initial participation, became a continuous nonparticipant	11%
Intermittent Participation	0%
Nonparticipation (never agreed)	16%
Refused to participate when contacted	4%
Not contacted/unable to be contacted	13%
Beneficiaries participating more than 75% of eligible months	74%
Number of participants in selected months[1]	
Month 6	618
Month 12	590
Month 24 (last month)	533

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

Data Sources: Medicare claims data, Medicare enrollment database.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-2.sas 27JUN2010.

4.3.2 Characteristics of MGH's CMP Intervention and Comparison Populations

In addition to evaluating the level of initial engagement by MGH's CMP, our participation analysis is designed to confirm that the selection procedures produced similar demographic, disease, and economic burden profiles between the intervention and comparison groups for both the original and refresh populations. Identifying any systematic baseline differences in demographic characteristics, health status, or baseline chronic condition patterns between the intervention and comparison group beneficiaries is important because the contractual and financial benchmarks established as part of the CMHCB demonstration program are based on an ITT framework and an assumption that the intervention and comparison groups are equivalent or essentially equivalent at the start of the demonstration.

We used the go-live date as our reference point and examined claims for 1 year prior to the go-live date. Only beneficiaries that had some eligibility in both the baseline and intervention periods were selected for this analysis. We explore the sufficiency of the assignment procedures for producing similar populations based on the selection strata and other variables. We also examine whether there are any systematic baseline differences in the disease burden between the intervention and comparison group beneficiaries assessed at the start of the demonstration. ***Supplement 4A*** provides tables displaying the percent of beneficiaries by these characteristics for the intervention and comparison populations for both the original and refresh populations.

Characteristics of MGH's CMP original population—In ***Supplement*** 4A, we report the beneficiary characteristics for the intervention and comparison groups for the original and refresh populations. ***Supplement 4A*** also reports the beneficiary characteristics that predict participation in MGH's CMP for both the original and refresh populations. Note that these are univariate tests, so there could be some correlation with other variables, thus we follow up with the multivariate analysis. Among the original population, intervention beneficiaries had lower percentages of the disabled, Medicaid enrollees, beneficiaries under the age of 65, and lower rates of diabetes with and without complications, cardiomyopathy, and acute and chronic renal disease. The refresh population had fewer statistically significant differences in the beneficiary characteristics. However, the intervention population had a lower percentage of beneficiaries in the high baseline HCC score group and a higher percentage of beneficiaries in the medium baseline HCC score group compared with the comparison group. Because such a high percentage of beneficiaries participated, we see few statistical differences between participants and nonparticipants. ***Supplement 4A*** also provides participation rates during the first 6 months of the demonstration by beneficiary demographic characteristics, baseline clinical and financial characteristics, and intervention period health status that we use in the multivariate modeling of participation.

4.3.3 Characteristics of Participants in MGH's CMP Original and Refresh Populations

In order to better understand the characteristics that most strongly predict participation in the demonstration, we estimated three multivariate logistic regression models for both the original and refresh populations:

1. Model 1: Beneficiaries who participated at least 1 month in the first 6 months of the intervention period compared with all other beneficiaries (nonparticipants);
2. Model 2: Beneficiaries who participated at least 1 month during the full intervention period compared with all other beneficiaries (nonparticipants); and
3. Model 3: Beneficiaries who participated at least 75% of eligible months compared with all other beneficiaries (nonparticipants and minimal participants).

Presentation of these regression results allows for a comparison of characteristics of beneficiaries who agreed to participate during the initial 6-month engagement period for at least 1 month versus characteristics of beneficiaries who agreed to participate at any point during the entire intervention period versus those who participated in MGH's CMP more than 75% of their eligible months. Model 1 reflects the initial recruitment emphasis by MGH's CMP, or characteristics of beneficiaries with whom MGH's CMP had the longest potential period of

intervention. Model 3 reflects characteristics of the beneficiaries who demonstrated the greatest willingness or ability to participate in MGH's CMP. For each model, we estimated two equations; an equation with just demographic characteristics and a full model equation that includes baseline and demonstration utilization and health status variables. Because there is correlation between beneficiary characteristics and the other variables, such as health status and baseline characteristics, we were most interested in examining which beneficiary characteristics had the greatest effect on willingness to participate before controlling for these other factors. The results for all three models were very similar in direction and magnitude of effect of beneficiary characteristics on the likelihood of participation so we do not display results of Models 1 and 2 in the body of the text (see ***Supplement 4A***).

Tables 4-5 and 4-6 present the results of the logistic regression analyses that predict participation based on various beneficiary characteristics for the original and refresh populations for Model 3, participants for more than 75% of their eligible months. Model 3a (columns 1 and 2) contains the odds ratio and associated statistical level of significance for the equation with just beneficiary characteristics. Model 3b (columns 3 and 4) contains the odds ratio and associated statistical level of significance for the equation with additional utilization and health status variables. An odds ratio less than 1 means that beneficiaries with a particular characteristic were less likely to participate; an odds ratio greater than 1 means that beneficiaries with the particular characteristic were more likely to participate. In general, the reference group comprises characteristics associated with younger and healthier beneficiaries. Across all three models, the explanatory power of the studied beneficiary characteristics was extremely low. This suggests that there is not a strong set of variables that predict the likelihood of a beneficiary being in the participant group. Pseudo R-squares for all of the models were 0.05 or less for both the original population and refresh populations. ***Supplement 4A*** contains tables that present the odds ratios and levels of significance for Models 1 and 2.

Table 4-5
Logistic regression modeling results comparing beneficiaries that participated at least 75% of eligible months during MGH's CMP intervention period to all other intervention beneficiaries: Original population[1,2]

Characteristics	Model 3A OR	p[3]	Model 3B OR	p[3]
Intercept	6.19	**	7.22	**
Beneficiary characteristics				
Male	0.99	N/S	1.00	N/S
African American/other/unknown	0.93	N/S	0.90	N/S
Age < 65 years	0.84	N/S	0.79	N/S
Age 75-84	0.92	N/S	0.92	N/S
Age 85 + years	0.74	N/S	0.78	N/S
Medicaid	0.68	N/S	0.63	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.35	N/S
Baseline HCC score high	N/I	N/I	2.38	**
Medium baseline PBPM cost	N/I	N/I	0.91	N/S
High baseline PBPM cost	N/I	N/I	0.68	*
Baseline Charlson score medium	N/I	N/I	0.87	N/S
Baseline Charlson score high	N/I	N/I	0.61	*
Demonstration period health status				
Died	N/I	N/I	0.50	**
Concurrent HCC score medium	N/I	N/I	1.01	N/S
Concurrent HCC score high	N/I	N/I	0.91	N/S
Number of cases	2,584	N/A	2,584	N/A
Chi-square (p<)	4.88	**	43.99	**
Pseudo R-square	0.00	N/A	0.02	N/A

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the entire period the Care Management Organization was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

N/I means not included; N/A means not applicable; N/S means not statistically significant.

The baseline HCC score reference group is < 2. The age reference group is 65-74 years. The PBPM reference group is < $336.70. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.811 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: bene02 27JUNE2010, partab3b and partab4b 28JUNE2010.

Table 4-6
Logistic regression modeling results comparing beneficiaries that participated at least 75% of eligible months during MGH's CMP intervention period to all other intervention beneficiaries: Refresh population[1,2]

Characteristics	Model 3A OR	p[3]	Model 3B OR	p[3]
Intercept	5.10	**	5.95	**
Beneficiary characteristics				
Male	0.62	*	0.60	*
African American/other/unknown	0.84	N/S	0.87	N/S
Age < 65 years	1.29	N/S	1.21	N/S
Age 75-84	1.09	N/S	1.08	N/S
Age 85 + years	0.73	N/S	0.84	N/S
Medicaid	1.06	N/S	1.09	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.62	N/S
Baseline HCC score high	N/I	N/I	0.91	N/S
Medium baseline PBPM cost	N/I	N/I	0.78	N/S
High baseline PBPM cost	N/I	N/I	0.96	N/S
Baseline Charlson score medium	N/I	N/I	0.63	N/S
Baseline Charlson score high	N/I	N/I	1.14	N/S
Demonstration period health status				
Died	N/I	N/I	0.24	**
Concurrent HCC score medium	N/I	N/I	1.07	N/S
Concurrent HCC score high	N/I	N/I	1.19	N/S
Number of cases	775	N/A	775	N/A
Chi-square (p<)	8.97	N/S	38.73	**
Pseudo R-square	0.01	N/A	0.05	N/A

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the entire period the Care Management Organization was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

N/I means not included; N/A means not applicable; N/S means not statistically significant.

The baseline HCC score reference group is < 2. The age reference group is 65-74 years. The PBPM reference group is < $324. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.694 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: bene02 27JUNE2010, partab3b and partab4b 28JUNE2010.

Model 3a shows no statistical differences between beneficiaries that participated more than 75% of eligible months and all other beneficiaries (***Table 4-5***). Examining Model 3b for the original population (***Table 4-5***), beneficiaries with high baseline predictive HCC scores were more likely to participate than those with low predictive HCC scores, holding other factors constant. And, somewhat contradictorily, beneficiaries with high baseline PBPM costs, high baseline Charlson scores, and those beneficiaries that died during the demonstration period were less likely to participate for more than 75% of eligible months. One would expect the same direction of influence of baseline HCC and Charlson scores. Although efforts were made to remove beneficiaries institutionalized prior to the demonstration start period, it is likely that those who died during the demonstration may have been more likely a resident of an institutional facility making it more difficult for the MGH CMP to interact with these beneficiaries to gain their participation.

There are fewer statistically significant results for the refresh population (***Table 4-6***). Beneficiaries that were male and that died during the demonstration period were less likely to be participants for long periods of time. Given that 70% of beneficiaries in the intervention group participated more than 75% of eligible months, and the total number of beneficiaries was only 775, it is difficult to capture any statistically significant differences between the participants and nonparticipants.

4.3.4 Level of Intervention

In this section, we report the frequency of interaction between MGH'S CMP and intervention beneficiaries for a subset of original intervention population beneficiaries who were fully eligible and participating from months 7 through 36 for the original population and months 7 through 24 for the refresh population. This essentially removes beneficiaries who died or lost eligibility early in the demonstration period and with whom MGH would have had limited opportunity to intervene. However, we do include contact information for the full demonstration time period as assessment activities were ongoing during the first 6-month period. We also examine whether there is evidence of selective targeting of beneficiaries for intervention contacts based upon level of perceived need as determined by beneficiary demographic, health status, baseline costliness, and acute care utilization during the demonstration period. MGH's CMP target population had a high prevalence of comorbid conditions, such as diabetes and HF. During the second site visit, MGH's CMP staff reported that they terminated their contract with Health Dialog and made several staffing changes - a designated case manager position was added to work specifically on post discharge assessments, a data analytics team was created to develop and strengthen program's reporting capabilities, the role of the community resource specialist was expanded to provide patient support and referral for non-clinical services, and they strengthened their mental health component of the program by adding social work positions. Thus, we expect to see a pattern of higher levels of intervention contacts for beneficiaries in poorer health status or higher users of hospitalization services, especially for the refresh population.

Descriptive statistics were performed using beneficiaries participating in MGH's CMP demonstration program to determine the breadth and depth of contacts related to care management. The data represent beneficiaries who were fully eligible and participating (unless they died) for months 7 through the end of Phase I of the demonstration. ***Table 4-7*** provides a detailed description

of the method of contact and number of contacts during the demonstration for the subset of eligible beneficiaries. Telephonic contact was the dominant form of contact (92% for the original population and 94% for the refresh population). Eight percent of original population contacts were in-person contacts, primarily visits by the case manager. For the refresh population, office visits to a social worker or home visits by a physician or nurse practitioner were the dominant forms of in-person contact. Although the number of contacts across all categories is considerably smaller for the refresh population this is to be expected since the size of the refresh population is only 30% of the original population size and the demonstration period is only 24 months rather than 36 months.

Table 4-7
Frequency distribution of MGH's CMP interactions: Total contacts[1]

Contact	Original Frequency	Percent	Refresh Frequency	Percent
Patient has been referred to and seen by a:				
Psychiatrist	151	1.1	33	1.6
Social Worker	113	0.8	40	2.0
End-of-life program	41	0.3	6	0.3
Home visit by physician or nurse practitioner	140	1.0	47	2.3
In-person visit by case manager	687	4.9	0.0	0.0
Total in-person[2]	1,132	8.0	126	6.3
Telemonitoring	9	0.1	0	0.0
Outbound or inbound telephonic	12,967	91.9	1,876	93.7
Total contacts	14,108	100.0	2,002	100.0

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Beneficiaries had to be fully eligible and full participants months 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[2] Includes psychiatrist, social worker, end-of-life program, home visits, and in-person contacts with the case manager

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010

MGH's CMP defined risk on an ongoing basis using initially the Health Dialog risk stratification and subsequently primary care physician or case manager assessments. ***Table 4-8*** shows the percent of each type of visit that were for MGH-defined high and low risk beneficiaries (medium risk beneficiaries are not shown). For the original population, 60% of total contacts and 76% of in-person visits were for beneficiaries defined as low risk. For the refresh population, only 14% of total contacts were for beneficiaries defined as low risk. Most of the contacts (66%) were with medium-risk beneficiaries (not shown). However, we do note a shift in focus for the home visits for the refresh population, the vast majority of home visits were for high or medium risk patients. For the original intervention population, almost one-third of home visits were for low risk patients.

Table 4-8
Frequency distribution of MGH's CMP interactions by MGH risk status: Total contacts[1]

Contact	Original High Risk Percent	Original Low Risk Percent	Refresh High Risk Percent	Refresh Low Risk Percent
Patient has been referred to and seen by a:				
Psychiatrist	17.2	30.5	15.2	24.2
Social Worker	14.2	29.2	32.5	7.5
End-of-life program	29.3	43.9	16.7	50.0
Home visit by physician or nurse practitioner	27.9	32.9	34.0	2.1
In-person visit by case manager	17.3	75.7	0.0	0.0
Telemonitoring	22.2	33.3	0.0	0.0
Outbound or inbound telephonic	16.5	60.4	19.3	14.5
Total contacts	16.7	60.2	19.9	14.3

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Beneficiaries had to be fully eligible and full participants months 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Table 4-9 displays the distribution of care management-related contacts for the original population. A total of 1,654 unique original population beneficiaries met the selection criteria - fully eligible and participating (unless they died) for months 7 through the end of the demonstration period. Observations were weighted by the fraction of eligible days, accounting for fewer contacts due to attrition because of death, which resulted in 1,551 full-time equivalent beneficiaries. A very high percentage of these beneficiaries had at least one contact during the demonstration period (97%) with the average number of contacts for each beneficiary being 8 and the median number of contacts being 5. On average, beneficiaries were contacted during 4 of the 10 quarters of the demonstration. One-third of beneficiaries had less than 3 contacts and one-third of beneficiaries had 8 or more contacts over the 36-month period.

Table 4-9
Distribution of number of contacts[1] with participants[2] in MGH's CMP: Original intervention population

Statistic	Number	Percent
Number of beneficiaries[3]	1,654	—
FTE beneficiaries[4]	1,551	—
Percent of beneficiaries that had at least one contact	97%	—
Mean number of contacts	8	—
Median number of contacts	5	—
Mean number of quarters of contact	4	—
Median number of quarters of contact	4	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-3 contacts	522	33.7%
4-7 contacts	516	33.3%
8+ contacts	512	33.0%
Total	1,551	100.0%

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program;

FTE = full time equivalent.

[1] Contacts are restricted to in-person and telephonic inbound and outbound. Excludes telemonitoring.

[2] Participants are defined as patients and caregivers in this analysis.

[3] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[4] Beneficiary counts weighted by fraction of eligible days = full-time equivalents.

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH's CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010.

The overall rate of contact reported in ***Table 4-9*** seems low relative to the level of interaction between the case managers and participants reported to us at our site visits and our observation of the level of day-to-day intervention between case managers and participants and case managers and primary care physicians (PCPs). There are likely to be numerous reasons for what we believe to be under-reporting of intervention activities. First, the reported data do not capture interactions between case managers and PCPs in the office through the MGH electronic medical record (EMR) and interactions with specialty physicians via any mode of contact. Given that the case managers are embedded in the clinical practice site we were told by case managers and PCPs about the high level of frequent interactions between the two regarding participants. This was viewed as one of the key advantages of MGH's CMP model that places case managers in practice sites. Second, it is unlikely that we are capturing all of the in-person contacts between case managers and patients that occur in the physician office at the time of visits. Case managers attempt to review the EMR the day patients have visits with their PCPs and/or have face-to-face contacts with patients after their visits to ensure understanding by the patient of changes in the care plan, especially prescription drug changes. And, third, there was a major CMP system redesign part way through the demonstration and in response to the desire to reduce documentation burden and better capture the case managers' interactions with their patients recognizing that there was likely under-reporting. Case managers' top priorities were (1) provision of case management services before documentation in the CMP database and (2) documentation in the EMR before documentation in the CMP database.

Table 4-10 displays distribution of care management-related contacts for the refresh population. A total of 541 unique refresh population beneficiaries met the selection criteria (529 full-time equivalents). A lower percentage of beneficiaries were contacted during the 24-month refresh period (87%). The number of contacts and quarters of contact are decidedly lower for this population likely due to the shorter demonstration time period.

Table 4-10
Distribution of number of contacts[1] with participants[2] in MGH's CMP: Refresh intervention population

Statistic	Number	Percent
Number of beneficiaries[3]	541	—
FTE beneficiaries[4]	529	—
Percent of beneficiaries that had at least one contact	87%	—
Mean number of contacts	4	—
Median number of contacts	3	—
Mean number of quarters of contact	2	—
Median number of quarters of contact	2	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-1 contacts	172	32.6%
2-4 contacts	197	37.3%
5+ contacts	159	30.1%
Total	529	100.0%

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; FTE = full time equivalent.

[1] Contacts are restricted to in-person and telephonic inbound and outbound. Excludes telemonitoring.

[2] Participants are defined as patients and caregivers in this analysis.

[3] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[4] Beneficiary counts weighted by fraction of eligible days = full-time equivalents.

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH's CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010.

Table 4-11 displays the percent of participants with interactions with MGH's CMP staff – in-person visits, telephone contacts, and total contacts (telephonic and in-person) by frequency of contact over the 36 months of the demonstration for the original population. About 1/3 of beneficiaries had no in-person visits. Nineteen percent of beneficiaries had one in-person visit and another 13% of beneficiaries had 2 to 4 in-person visits during the 36-month period.

Table 4-11
Percent distribution of participants[1] with MGH's CMP interactions[2]: Original intervention population

Type and frequency of contact	Number of FTE beneficiaries[3,4]	Percent
In-person		
0	1,030	66.4
1	292	18.8
2-4	195	12.6
5-9	31	2.0
10-19	3	0.2
20+	0	0.0
Telephonic		
0	56	3.6
1	157	10.1
2-4	536	34.6
5-9	469	30.2
10-19	219	14.2
20+	114	7.3
Total telephonic and in-person		
0	45	2.9
1	127	8.2
2-4	511	32.9
5-9	493	31.8
10-19	247	16.0
20+	127	8.2

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; FTE = full time equivalent.

1 Participants are defined as patients and caregivers in this analysis.

2 Contacts are restricted to in-person and telephonic inbound and outbound. Excludes telemonitoring.

3 Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

4 Beneficiary counts weighted by fraction of eligible days = full-time equivalents.

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010.

Participants in the MGH's CMP received more phone calls during the 36 months of the demonstration. All but 4% of beneficiaries received a telephone call from a case manager, while 22% received 10 or more calls. Combining telephone and visit contacts, we observe that 3% of fully eligible and participating beneficiaries had no contact for the 36-month period and another 75% had fewer than 10 contacts. Yet at the same time, we observe 8% of beneficiaries had 20 or more contacts with the majority being telephone contacts. We find higher percentages of beneficiaries receiving no contact in ***Table 4-12***, which provides the same distributions for the refresh population

Table 4-12
Percent distribution of participants[1] with MGH'S CMP interactions[2]: Refresh intervention population

Type and frequency of contact	Number of FTE beneficiaries[3,4]	Percent
In-person		
0	442	83.5
1	63	12.0
2-4	24	4.5
5-9	0	0.0
10-19	0	0.0
20+	0	0.0
Telephonic		
0	80	15.1
1	107	20.2
2-4	197	37.3
5-9	110	20.9
10-19	35	6.6
20+	0	0.0
Total telephonic and in-person		
0	68	12.9
1	104	19.7
2-4	197	37.3
5-9	123	23.3
10-19	35	6.6
20+	1	0.2

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; FTE = full time equivalent.

[1] Participants are defined as patients and caregivers in this analysis.

[2] Contacts are restricted to in-person and telephonic inbound and outbound. Excludes telemonitoring.

[3] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[4] Beneficiary counts weighted by fraction of eligible days = full-time equivalents.

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010.

Table 4-13 displays the frequency of total contacts by baseline HCC score and type of contact for the original population. Contact by mode was not mutually exclusive in that a beneficiary could have a combination of telephone and visit contacts any time during the 36 months of the demonstration period. Beneficiaries were stratified into three HCC categories ranging from an HCC score greater than 3.1 to less than 2.0 using baseline prospective HCC scores calculated for the 12 months prior to the beginning of the demonstration period. One-third of beneficiaries are distributed across the three categories with a slightly higher percentage of beneficiaries in the low HCC risk group (36%) compared to the high HCC risk group (29%).

Table 4-13
Frequency of MGH'S CMP contacts by HCC score:
Original intervention population

Contact mode	HCC Score High (>3.1) N = 454 Frequency	%	HCC Score Medium (2 to ≤3.1) N = 544 Frequency	%	HCC Score Low (<2) N = 552 Frequency	%
In-person						
0	294	64.8	368	67.5	368	66.7
1	80	17.6	99	18.2	113	20.4
2-4	65	14.3	65	12.0	65	11.7
5-9	13	2.9	12	2.2	5	0.9
10-19	2	0.4	0	0.0	1	0.2
20+	0	0.0	0	0.0	0	0.0
Telephonic						
0	15	3.3	17	3.1	24	4.4
1	42	9.3	56	10.3	59	10.7
2-4	129	28.4	197	36.3	209	37.9
5-9	148	32.6	153	28.2	167	30.3
10-19	74	16.3	78	14.3	68	12.3
20+	46	10.2	43	7.8	24	4.4
Total telephonic and in-person						
0	13	2.8	15	2.8	17	3.2
1	32	7.0	47	8.6	49	8.8
2-4	129	28.4	184	33.8	198	35.9
5-9	148	32.6	168	30.8	177	32.0
10-19	82	17.9	85	15.7	81	14.6
20+	51	11.2	46	8.4	30	5.5

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program;
HCC =Hierarchical Condition Category; N = number of beneficiaries.

[1] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[2] Beneficiary counts weighted by fraction of eligible days = full-time equivalents

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010

In-person visits—The percentage distribution of in-person visits is similar across all three risk categories. High risk beneficiaries have a slightly lower percentage of beneficiaries receiving no in-person contact. There is no clearly defined targeting activity based on the distribution of in-person visits.

Telephone contacts—We also observe fairly similar distributional patterns across risk categories in number of telephone contact. Low baseline HCC risk score beneficiaries were a bit more likely to have had no telephonic contact in comparison with medium and high HCC risk score beneficiaries. When examining the two highest categories of outbound calls, the high risk group has a higher percentage of participants that received 10 or more calls during the 36-month period compared to the low risk group (27% compared to 17%) with 10% of beneficiaries in the high HCC risk group receiving 20 or more phone calls compared with 4% of beneficiaries in the low HHC risk group. These findings suggest that MGH's CMP made a focused effort to contact their higher acuity beneficiaries on a more regular basis. Because total contacts are driven by telephonic contact, the distributional patterns of total contacts are similar to telephone contacts.

Table 4-14 displays the frequency of total contacts by baseline HCC score and type of contact for the refresh population. The distribution of refresh beneficiaries across the three HCC risk categories is not as even as we observed for the original population. Almost 40% of beneficiaries are in each of the medium and low HHC risk groups (39% and 36%, respectively) compared to 25% of beneficiaries in the high HCC risk group. We observe greater percentages of refresh beneficiaries in all three HCC risk groups with no in-person or telephonic contacts than we observed for the original population. And, we observe more similar percentages of refresh beneficiaries receiving 10 or more telephone calls in the high and low HCC risk groups as compared to the original beneficiaries; 9.1% versus 6.8% for the refresh population, respectively, and 27% and 17% for the original population, respectively. No refresh beneficiary received 20 or more telephone calls during the 24-month period regardless of HCC risk group. In contrast, 10% of original high HCC risk score beneficiaries received 20 or more telephone contacts.

Table 4-14
Frequency of MGH'S CMP contacts by HCC score:
Refresh intervention population

Contact mode	HCC Score High (>3.1) N = 132 Frequency	%	HCC Score Medium (2 to ≤3.1) N = 207 Frequency	%	HCC Score Low (<2) N = 189 Frequency	%
In-person						
0	110	83.4	173	83.6	158	83.5
1	16	12.1	24	11.6	23	12.3
2-4	6	4.5	10	4.8	8	4.2
5-9	0	0.0	0	0.0	0	0.0
10-19	0	0.0	0	0.0	0	0.0
20+	0	0.0	0	0.0	0	0.0
Telephonic						
0	19	14.1	28	13.7	33	17.3
1	25	18.9	42	20.1	40	21.2
2-4	44	33.0	83	40.0	71	37.4
5-9	33	25.0	44	21.4	33	17.4
10-19	12	9.1	10	4.8	13	6.8
20+	0	0.0	0	0.0	0	0.0
Total telephonic and in-person						
0	18	13.4	24	11.8	26	13.7
1	17	12.9	44	21.0	44	23.1
2-4	49	36.8	80	38.6	69	36.3
5-9	37	28.0	49	23.9	37	19.5
10-19	11	8.3	10	4.8	14	7.3
20+	1	0.8	0	0.0	0	0.0

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; HCC =Hierarchical Condition Category; N = number of beneficiaries.

[1] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[2] Beneficiary counts weighted by fraction of eligible days = full-time equivalents

Data Sources: RTI analysis of 2007-2009 Medicare enrollment, eligibility, and MGH'S CMP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/enctab2 19AUG2010.

To more directly examine the targeting strategy of MGH's CMP, a multivariate logistic regression model was estimated with the number of total contacts as the dependent variable. The model estimates the likelihood of a participant receiving a high number of contacts. The medium contact group was omitted, thus comparing the high contact group to the low contact group. ***Tables 4-15*** (original population) ***and 4-16*** (refresh population) display the odds ratios for discrete categories of demographic characteristics, baseline health status, baseline Medicare payments, and demonstration health status. Beneficiaries were weighted by their period of eligibility during the last 30 months of the demonstration for the original population and last 18 months for the refresh population, and their number of contacts categorized either as low or high. Low contacts are defined as 0-3 contacts for the original population and less than 2 for the refresh population. Original population participants with eight or more contacts are captured in the high contact category, while 5 or more contacts are defined as high contact for the refresh population. Odds ratios are partial in the sense that all other variables are held constant. For example, the odds of a beneficiary younger than 65 years of age experiencing a high contact rate are 1.9 times greater than those for a beneficiary age 65 and older, adjusting for any baseline difference in other beneficiary characteristics and demonstration period health status.

For the original population, disabled beneficiaries (as defined by beneficiaries younger than age 65) and beneficiaries with high baseline PBPM costs were found to be statistically significant indicators of the likelihood of being in the high contact category (***Table 4-15***). A high concurrent HCC score, or health status measured during the first 6 months of the demonstration period, was found to be a positive predictor of being in the high contact group. Demonstration period acute care utilization was also a strong predictor of a high level of contact and reflects MGH's CMP staff's successful effort to target participants that had been to an emergency room or hospitalized. Beneficiaries who had one hospitalization were 2 times more likely to be in the high contact group while beneficiaries with multiple hospitalizations were 5 times more likely to be in the high contact group than those who had no hospitalizations. Beneficiaries who died during the demonstration were less likely to be in the high contact category.

For the refresh population, there are fewer statistically significant results but we do observe a relationship between demonstration period acute care utilization and likelihood of being in the high contact group (***Table 4-16***). Beneficiaries who had one hospitalization were almost 2 times more likely to be in the high contact group while beneficiaries with multiple hospitalizations were nearly 6 times more likely to be in the high contact group (odds ratio = 5.82), than those who had no hospitalizations. There were no other statistically significant indicators of high contact. However, the number of beneficiaries included in this analysis is very low (337). These findings suggest that MGH's CMP continued to be successful contacting the refresh beneficiaries who were at high risk of hospitalization or re-hospitalization.

Table 4-15
Logistic regression modeling results comparing the likelihood of being in MGH's CMP high contact category relative to the low contact category: Original intervention population

Characteristics	Odds ratio[1,2]	*p*[3]
Intercept	0.49	**
Beneficiary characteristics		
Male	0.57	**
African American/other/unknown	1.00	N/S
Age <65	1.88	**
Age 75-84	1.11	N/S
Age 85+ years	0.86	N/S
Baseline characteristics		
Baseline HCC score medium	1.11	N/S
Baseline HCC score high	1.17	N/S
Medium base PBPM cost	1.25	N/S
High base PBPM cost	1.70	*
Baseline Charlson score medium	1.07	N/S
Baseline Charlson score high	0.81	N/S
Demonstration period health status		
Died	0.46	*
Concurrent HCC score medium	1.17	N/S
Concurrent HCC score high	1.65	**
One hospitalization	2.14	**
Multiple hospitalizations	5.05	**
Number of cases	1,116	N/A
Chi-square (p<)	141.31	**
Pseudo R2	0.12	N/A

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

1 Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

2 Beneficiary counts weighted by fraction of eligible days = full-time equivalents

3 * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

N/A means not applicable; N/S means not statistically significant.

The baseline HCC score reference group is < 2. The age reference group is 65-74 years. The PBPM reference group is < $336.70. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.811 or less.

Data Sources: RTI analysis of 2004-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/ enctab3a 25AUG2010.

Table 4-16
Logistic regression modeling results comparing the likelihood of being in MGH's CMP high contact category relative to the low contact category: Refresh intervention population

Characteristics	Odds ratio[1,2]	p[3]
Intercept	0.41	*
Beneficiary characteristics		
Male	0.92	N/S
African American/other/unknown	0.89	N/S
Age <65	1.63	N/S
Age 75-84	1.46	N/S
Age 85+ years	2.04	N/S
Medicaid	0.92	N/S
Baseline characteristics		
Baseline HCC score medium	1.36	N/S
Baseline HCC score high	1.67	N/S
Medium base PBPM cost	0.93	N/S
High base PBPM cost	1.09	N/S
Baseline Charlson score medium	0.77	N/S
Baseline Charlson score high	0.92	N/S
Demonstration period health status		
Died	0.18	N/S
Concurrent HCC score medium	0.89	N/S
Concurrent HCC score high	1.50	N/S
One hospitalization	1.84	*
Multiple hospitalizations	5.82	**
Number of cases	337	N/A
Chi-square (p<)	40.31	**
Pseudo R2	0.11	N/A

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Beneficiaries had to be fully eligible and full participants in the last 7 through 36 for the original population and months 7-24 for the refresh population of the MGH'S CMP.

[2] Beneficiary counts weighted by fraction of eligible days = full-time equivalents

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

N/A means not applicable; N/S means not statistically significant.

The baseline HCC score reference group is < 2. The age reference group is 65-74 years. The PBPM reference group is < $324. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.694 or less.

Data Sources: RTI analysis of 2004-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/MGH/ enctab3a 25AUG2010.

4.4 Summary

The MGH CMP had an ambitious goal of gaining participation from 100% of its original population beneficiaries. It was successful in recruiting 88% of its original population beneficiaries and 84% of its refresh population beneficiaries. Because of high levels of participation we found few statistically significant differences between participants and nonparticipants in either the original or refresh intervention populations. In multivariate modeling of factors that predict likelihood of participation, we had low explanatory power suggesting that MGH's CMP was able to recruit broadly across its intervention population as no particular set of factors strongly predicted participation. The substantially smaller sample size for the refresh population also limited our ability to detect participation factors.

A cornerstone of MGH's CMP was one-on-one relationship between patients and their practice-based case managers, supplemented by support received from the program's mental health, pharmacist, and end-of-life components. Telephone contact was the most dominant form of contact. Nearly every participating original population beneficiary received at least one call or in-person visit from a case manager during the 36 months of the demonstration. However, average number of contacts was relatively low for this chronically ill complex population, 8 contacts on average. When examining rate of contact by baseline health status measured by the HCC risk score, we found evidence that MGH's CMP made a focused effort to contact their higher acuity original population beneficiaries on a more regular basis. The high HCC risk score group had a larger percentage of participants who received 10 or more calls during the 36-month period compared to the low HCC risk score group (27% compared to 17%) with 10% of beneficiaries in the high HCC risk score group receiving 20 or more phone calls compared with 4% of beneficiaries in the low HHC risk score group. We observe greater percentages of refresh beneficiaries with no in-person or telephonic contacts than we observed for the original population. No refresh beneficiary received 20 or more telephone calls during the 24-month period regardless of HCC risk score. And, we observe less targeting by HCC risk score.

In our multivariate regression modeling of likelihood of being in a high contact versus low contact group for both the original and refresh populations, we found hospitalizations during the demonstration period to be very strong predictors of contact. A major focus of MGH's CMP was to prevent hospitalizations or re-hospitalizations. These findings suggest that MGH's CMP was successful in their effort to contact beneficiaries who had been hospitalized and at high risk of re-hospitalization in an effort to prevent readmissions.

CHAPTER 5
CLINICAL QUALITY PERFORMANCE

5.1 Introduction

RTI's analysis of quality of care focuses on measuring effectiveness of MGH's CMP as part of the CMHCB demonstration by answering the following evaluation question:

- *Clinical Quality of Care:* Did MGH's CMP improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?

In this chapter, we present analyses related to clinical quality performance during MGH's CMP by examining changes in the rate of receipt of three evidence-based, process-of-care measures during the demonstration, relative to a 12-month baseline period in both the intervention and comparison populations. We selected these annual measures appropriate for different populations of elderly beneficiaries, in general, and with chronic conditions such as diabetes and ischemic vascular disease (IVD): influenza vaccine for all beneficiaries; low-density lipoprotein cholesterol (LDL-C) testing for beneficiaries with diabetes or IVD; and rate of annual HbA1c testing for beneficiaries with diabetes.

However, we do note that in our site visits to MGH, the CMP leadership felt that there was a very good reason standard quality measures are not part of this demonstration's outcomes. Program leaders reported that such measures are not good quality of care indicators for the program's population. For their group of patients, something like testing for HbA_{1c} levels is not a relevant measure of how well the program is managing the care of their very sick and complex patients. The CMP leadership and MGH leadership believe that ER use and acute hospitalizations are in essence the measures that need to be used. In addition to these outcomes, other types of measures related to care coordination that they believe are highly relevant to this population include how fast case managers follow up on patient-initiated calls, can appointments be consolidated so frail beneficiaries do not have to drive to the hospital 3 times a week, and so forth. To this end, MGH's CMP developed a weekly and monthly clinical dashboard to monitor patients' utilization of medical services. Both dashboards were built using MGH internal data. The monthly dashboard included: population characteristics, top 5 diagnoses for inpatient discharges, inpatient and psychiatric admissions, readmissions, primary and urgent care visits, average length of stay, and preventable hospitalizations. The data do not include ambulatory visits or laboratory or radiology service utilization. However, MGH's electronic medical record does include all laboratory results so case managers could see laboratory test results flagged as critical. Thus, it is likely that we will not observe significant changes in rate of receipt of the studied clinical measures during the demonstration period for the intervention beneficiaries.

Under an ITT model and our difference-in-differences evaluation approach, we require information for the pre- and demonstration periods and for both the intervention and comparison populations for our measures. Therefore, in our evaluation, we selected measures that we believed could be reliably calculated using Medicare administrative data to assess improvements in quality of care and health outcomes. Further, these data are available for both the intervention and comparison populations and do not require medical record abstraction or beneficiary self-report. Medical record data are not available to us for either the intervention or comparison

populations, and beneficiary self-report data would only available for the intervention beneficiaries who participated during the demonstration. Further, beneficiary self-report is subject to recall error and to the willingness of beneficiaries to provide the information.

5.2 Methods

We created the process-of-care measures for the 12-month period immediately prior to the go-live date for MGH's CMP for its original and refresh populations and for two intervention periods (months 7-18 and months 25-36) for its original population and for one intervention period (months 13-24, or the last 12 months of the demonstration) for its refresh population. Only beneficiaries who had at least 1 day of eligibility in both baseline and in each of the intervention periods were included in the analysis of each measure. ***Table 5-1*** provides the number of beneficiaries who were included in the analyses of the quality of care measures, in total, and by two disease cohorts: diabetes and ischemic vascular disease (IVD).

Medicare claims for the full baseline and intervention period were included regardless of beneficiary eligibility for MGH's CMP (e.g., claims were included even if beneficiaries did not pay the Part B premium for 1 or 2 months). This allowed us to provide credit to MGH's CMP for services received after exposure to their intervention and possibly as a result of the intervention. To the extent that the service was included in the Medicare claims files during a period of ineligibility for MGH's CMP—or as a denied claim due to disenrollment from Part B, for example—it reflects actual receipt of the service and was therefore included in our analyses.

Rates per 100 beneficiaries are reported for the intervention and comparison groups for the 12-month baseline period and for the intervention periods, weighted by beneficiary eligibility in each time period. For each measure, the difference-in-differences rate is reported and reflects the growth (or decline) in the intervention group's mean rate of receipt of care relative to the growth (or decline) in the comparison group's mean rate. A positive intervention effect for the guideline-concordant care measures occurred if the intervention group's mean rate increased more than the comparison group's mean rate, or declined less, during the demonstration period. A negative intervention effect occurred if the intervention group's mean rate increased less than the comparison group's mean rate, or declined more, during the demonstration period.

Statistical testing of the change in the rate of receipt of the quality of care measures was performed at the individual beneficiary level. The standard method for modeling a binary outcome, such as receiving an HbA1c test or not, is logistic regression. The experimental design for the CMHCB demonstration also requires that the variance of the estimates be properly adjusted for the repeated (pre- and post-) measures observed for each sample member within a nested experimental design. MGH's CMP was based on two nested cohort samples of Medicare beneficiaries who were assigned to intervention and comparison groups. In addition, an eligibility fraction ranging from 0 to 1 was included as the weight to reflect the period of time during which the beneficiary met MGH's CMP eligibility criteria in the baseline and demonstration periods. STATA SVY was used to fit the model with robust variance estimation.

Table 5-1
Number of beneficiaries included in analyses of guideline concordant care and acute care utilization for MGH's CMP

Statistics	All	Diabetes	Ischemic vascular disease
Original beneficiaries			
Months 7-18			
Intervention			
Total number of beneficiaries	2,427	817	895
Full time equivalents[1]	2,425	817	895
Comparison			
Total number of beneficiaries	2,490	1,036	924
Full time equivalents[1]	2,488	1,035	923
Months 25-36			
Intervention			
Total number of beneficiaries	1,994	670	714
Full time equivalents[1]	1,992	670	714
Comparison			
Total number of beneficiaries	1,979	824	715
Full time equivalents[1]	1,977	823	715
Refresh beneficiaries			
Months 13-24			
Intervention			
Total number of beneficiaries	691	205	252
Full time equivalents[1]	690	204	251
Comparison			
Total number of beneficiaries	658	260	242
Full time equivalents[1]	657	259	241

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

[1] Full Time Equivalent for the intervention group during the baseline period is the total number of beneficiaries weighted by their period of eligibility for the demonstration.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: gcc01, gcc02 27JUNE2010, gcctab, gcc_rob, gcctab1 28JUNE2010.

Logistic regression produces an odds ratio for every predictor variable in the model; that is, an estimate of that variable's effect on the dependent variable, after adjusting for the other variables (randomization factors) in the model. The odds ratio is greater than 1.0 when the presence of the variable is associated with an increased likelihood of receiving the service; an odds ratio less than 1.0 means that the variable is inversely associated with receiving the test. The statistical test determines whether the odds ratio is 1.0. We report the odds ratio associated

with the difference-in-differences interaction term, or the test of the difference-in-differences of the rate, and the odds ratio's associated *p* value and 95% confidence level.

To better understand the movement underlying the reported difference-in-differences rates, we stratified MGH's CMP original and refresh beneficiaries into four categories based upon whether or not they received each of the quality of care measures during the pre-demonstration baseline period and the last 12 months of the demonstration: compliant in both the baseline and demonstration periods; compliant in the baseline period but not in the demonstration period; not compliant in the baseline period but compliant in the demonstration period; and not compliant in both periods. We report on the natural trends observed in the comparison and intervention populations over the 3-year period.[7] Only beneficiaries who had at least 1 day of eligibility in both baseline and the last 12 months of the demonstration were included and the percentages were weighted by eligibility in each of the periods.

5.3 Findings

Process-of-care rates per 100 beneficiaries are reported for MGH's CMP original population in ***Table 5-2***. We report the baseline and intervention period rates for the intervention and comparison groups as well as the difference-in-differences rates (baseline period intervention versus comparison rate difference minus intervention period intervention versus comparison rate difference). Positive difference-in-differences rates per 100 beneficiaries indicate that the intervention group's mean rate improved more than the comparison group's mean rate or the intervention group's mean rate declined at a lower rate than the comparison group's mean rate. Negative difference-in-differences rates per 100 beneficiaries indicate that comparison group exhibited higher rates of growth or less of a decline than the intervention group.

Rates of three of the measures calculated for the pre-demonstration period in the original intervention and comparison groups are relatively high and similar, ranging from 77% for LDL-C testing for beneficiaries with IVD to 95% for LDL-C testing for beneficiaries with diabetes. However, the baseline rates for influenza vaccine were lower, 50% for the comparison group and 60% for the intervention group.

Over the course of the two demonstration periods for the original comparison population, we generally observe stable or decreasing rates of receipt for all measures except influenza vaccination; we observe a 6 percentage point increase during the first demonstration period and an 8 percentage point increase during the second demonstration period. Yet, the rates of influenza vaccination receipt within the comparison group remain lower than the rates observed within the intervention group. Of the eight measures evaluated for the original population, we observe only one statistically significant negative difference-in-differences rate; whereby the intervention group's rate of influenza vaccination increased less than the rate for the comparison group during months 25-36 of the demonstration. However, the original intervention group's rate

[7] We do not conduct statistical testing of the differences in distributions. Our formal test of quality improvement is conducted on the difference-in-differences rates using a model based test of statistical significance to allow for robust variance estimation. These data are provided for illustrative purpose only to better understand the natural movement in rate of receipt of quality of care measures in a cohort of elderly, ill FFS beneficiaries.

of receipt of influenza vaccination at baseline was 10 percentage points higher than the comparison group's rate.

Table 5-2
Comparison of rates of process-of-care measures for the first and last 12 months of the MGH CMP demonstration period with rates for a 1-year period prior to the start of the MGH CMP demonstration: Original and refresh populations

Process-of-care measures	Rate per 100 Baseline I[1]	Rate per 100 Baseline C[1]	Rate per 100 Demo period I[1]	Rate per 100 Demo period C[1]	D-in-D Rate per 100	D-in-D OR	D-in-D p	D-in-D CI Low	D-in-D CI High
ORIGINAL POPULATION									
Months 7-18									
All beneficiaries									
Influenza vaccine	60	50	64	56	-1.87	0.94	0.43	0.79	1.10
Beneficiaries with diabetes									
HbA1c test	95	95	91	91	0.40	1.05	0.84	0.62	1.78
LDL-C test	80	83	80	82	0.29	1.02	0.91	0.73	1.43
Beneficiaries with IVD[2]									
LDL-C test	77	80	76	76	2.93	1.19	0.28	0.87	1.63
Months 25-36									
All beneficiaries									
Influenza vaccine	60	50	63	58	-5.94	0.79	0.01	0.66	0.95
Beneficiaries with diabetes									
HbA1c test	95	95	90	90	0.01	0.99	0.98	0.56	1.76
LDL-C test	82	83	80	83	-2.57	0.85	0.39	0.58	1.24
Beneficiaries with IVD[2]									
LDL-C test	80	82	76	80	-1.68	0.92	0.66	0.63	1.33
REFRESH POPULATION									
Months 13-24									
All beneficiaries									
Influenza vaccine	61	46	64	60	-11.37	0.64	0.00	0.46	0.87
Beneficiaries with diabetes									
HbA1c test	94	92	89	89	-2.13	0.70	0.47	0.27	1.84
LDL-C test	81	86	82	79	7.80	1.72	0.12	0.86	3.42
Beneficiaries with IVD[2]									
LDL-C test	79	77	80	72	6.19	1.40	0.28	0.76	2.58

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; I = intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odds ratio;

LDL-C = low-density lipoprotein cholesterol; IVD = ischemic vascular disease.

[1] All rates are per 100 beneficiaries and are adjusted for periods of demonstration eligibility during the one-year period prior to the start of the demonstration and each set of months the care management organization was active in the program. Only beneficiaries who had at least one day of eligibility in both the baseline and demonstration periods are included in this analysis.

[2] Ischemic Vascular Disease is defined using the National Qualify Forum definition.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: gcc01, gcc02 27JUNE2010, gcctab, gcc_rob, gcctab1 28JUNE2010.

For the refresh population, we generally observe similar levels of compliance with the evidence-based care guidelines at baseline and similar patterns of change during the demonstration as noted for the original population. Of the four measures evaluated for the refresh population, we also observe one statistically significant negative difference-in-differences rate. The intervention group's rate of influenza vaccination increased less than the rate for the comparison group during months last 12 months of the demonstration. The refresh intervention group's rate of receipt of influenza vaccination at baseline was 15 percentage points higher than the refresh comparison group's rate.

Table 5-3 displays the percentages of the MGH CMP's original and refresh populations who did or did not receive one of the process-of-care measures during the baseline period and the last 12 months of each population's respective demonstration period. We display the distribution of intervention and comparison beneficiaries across four categories of compliance:

- always compliant, meaning compliant in both baseline and intervention periods;
- became noncompliant, meaning compliant in the baseline period but noncompliant in the intervention period;
- never compliant, meaning noncompliant in both the baseline and intervention periods; and
- became compliant, meaning noncompliant in the baseline period but compliant in the intervention period.

The first column for each quality of care measure contains the percentage distributions for the comparison populations and the second column displays the percentage distributions for the intervention populations. The top half displays rates of compliance for the original population and the bottom half for the refresh population.

For the original population and within both the intervention and comparison groups, there are high rates of always being compliant in receipt of three of the process of care measures with the noted exception of influenza vaccination. Only 1-in-3 or 1-in-4 beneficiaries were always compliant for influenza vaccination. We observe similar percentages of intervention and comparison beneficiaries becoming compliant (between 3 and 23%) or becoming noncompliant (between 7 and 17%) during the demonstration period across the four measures. During the last year of the demonstration, only 10% of intervention beneficiaries with diabetes were not compliant in receipt of annual HbA1c testing, 20% and 24% of intervention beneficiaries with diabetes or IVD, respectively, were not compliant in receipt of annual LDL-C testing, and 36% of intervention beneficiaries were not compliant in receipt of an annual influenza vaccine. A similar pattern is observed among the refresh beneficiaries.

Table 5-3
Percentage of comparison and intervention beneficiaries meeting process-of-care standards in the baseline year and last 12 months of MGH's CMP: Original and refresh populations

	HbA1c testing[1,2]	HbA1c testing[1,2]	LDL-C diabetes	LDL-C diabetes	LDL-C IVD	LDL-C IVD	Influenza vaccine	Influenza vaccine
Original population	C	I	C	I	C	I	C	I
Always compliant	87%	88%	71%	69%	67%	65%	34%	42%
Became noncompliant	7	7	12	13	14	14	16	17
Became compliant	3	3	11	11	11	11	23	21
Refresh population	C	I	C	I	C	I	C	I
Always compliant	84	87	71	73	62	70	36	44
Became noncompliant	8	8	16	9	16	10	10	17
Became compliant	5	3	8	9	10	10	24	20

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; LDL-C = low-density lipoprotein cholesterol; IVD = ischemic vascular disease; C = comparison population; I= intervention population.

[1] All percentages are adjusted for periods of beneficiary CMHCB demonstration eligibility during the one-year period prior to the start of the demonstration and the last 12 months the Care Management Organization was active.

[2] Only beneficiaries who had at least one day of eligibility in both the baseline and demonstration periods are included in this analysis.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: gcc01, gcc02 27JUNE2010, gcctab, gcc_rob, gcctab3 28JUNE2010.

5.4 Summary of Findings and Conclusion

In this chapter, we report on RTI's assessment of the effect of MGH's CMP on quality of care. Specifically, we report findings for the key research question: did MGH's CMP improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care? Rates of three of the measures calculated for the pre-demonstration period in the original and refresh intervention and comparison groups are relatively high and similar. Baseline rates for influenza vaccine were considerably lower for both the intervention and comparison populations.

Over the course of the two demonstration periods for the original and refresh comparison populations, we generally observe stable or decreasing trends in the rates of receipt for all measures except influenza vaccination. For both the original and refresh populations, we observe statistically significant negative difference-in-differences rates whereby the intervention group's rate of influenza vaccination increased less than the rate for the comparison group during the last 12 months of the demonstration. However, the original and refresh intervention groups' rates of influenza vaccination at baseline were far higher than the comparison groups' rate.

For the original and refresh populations and within both the intervention and comparison groups, there are high rates of always being compliant in receipt of three of the process-of-care measures with the noted exception of influenza vaccination. Thus, there was limited room for improvement in either population. During the last year of the demonstration, only 10% of intervention beneficiaries with diabetes were not compliant in receipt of annual HbA1c testing and 20% and 24% of intervention beneficiaries with diabetes or IVD, respectively, were not compliant in receipt of annual LDL-C testing.

Although 36% of intervention beneficiaries were not compliant in receipt of an annual influenza vaccine, the observed rate of compliance might be lower than the actual rate of compliance if the vaccine is given by providers for whom we cannot observe this service in Medicare Part B data. Vaccines provided during an inpatient stay or in an emergency room are not captured in the Part B data. Nor would we observe influenza vaccines provided to beneficiaries at senior centers or in drug stores or grocery stores.

Lastly, we did not expect to observe significant changes in rate of receipt of the studied clinical measures during the demonstration period for the intervention beneficiaries. MGH's CMP leadership felt that a measure like testing for HbA_{1c} levels is not a relevant measure of how well the program is managing the care of their very sick and complex patients. The CMP leadership and MGH leadership believe that ER use and acute hospitalizations are in essence the measures that need to be used. We examine acute care utilization in ***Chapter 6***.

CHAPTER 6
HEALTH OUTCOMES

6.1 Introduction

RTI's analysis of health outcomes focuses on measuring effectiveness of MGH's CMP as part of the CMHCB demonstration by answering the following evaluation questions:

- Did the MGH's CMP improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and emergency room (ER) utilization?
- Did MGH's CMP improve health outcomes by decreasing mortality?

In this chapter, we present analyses related to intermediate clinical health outcomes by examining relative to a 12-month baseline period changes in the rate of hospitalizations, ER visits, and readmissions during months 7-18 and the last 12 months of MGH's CMP for the original population and the last 12 months of the demonstration for the refresh population. We also examine differences in the rate of mortality between the intervention and comparison original and refresh beneficiaries during the entire demonstration period.

6.2 Methods

6.2.1 Rates of Hospitalizations and Emergency Room Visits

Rates of hospitalization and ER visits were constructed for the 12-month period immediately prior to the launch of MGH's CMP date, for months 7-18 for the original population, and the last 12 months of the intervention period for both the original and refresh populations. We constructed rates of all-cause hospitalization and ER visits and a combined utilization measure for 10 ambulatory care sensitive condition (ACSC) reasons for admission—heart failure, diabetes, asthma, cellulitis, chronic obstructive pulmonary disease (COPD) and chronic bronchitis, dehydration, bacterial pneumonia, septicemia, ischemic stroke, and urinary tract infection—using the primary diagnosis on the claim. Only claims that occurred during periods of eligibility were included in the utilization measures and only beneficiaries who had at least 1 day of eligibility in both baseline and the demonstration periods are included in these analyses. ***Table 5-1*** in ***Chapter 5*** displays the number of beneficiaries who were included in these utilization analyses.

All-cause and 10 ACSC rates of hospitalization and ER visits per 1,000 beneficiaries are reported for the intervention and comparison groups for the 12-month baseline period and for intervention periods, weighted by beneficiary eligibility in each time period. For each measure, the difference-in-differences rate is reported and reflects the decline (or growth) in the intervention group's mean rate of utilization relative to the decline (or growth) in the comparison group's mean rate.

We performed statistical testing of the change in the utilization rates at the individual beneficiary level. The distributional properties of the data led us to select a negative binomial generalized linear model to account for the presence of beneficiaries with no hospitalizations or ER visits in one time period or the other, as well as heterogeneity in rates of acute care service

use. As with the process-of-care measures, STATA SVY was used to fit the model with robust variance estimation to adjust for the repeated (pre- and post-) measures and multiple hospitalizations or ER visits observed for sample members within a nested experimental design. An eligibility fraction ranging from 0 to 1 was assigned to the pre- and post-time periods for each beneficiary and was included as the weight to reflect the period of time the beneficiary met MGH's CMP eligibility criteria in the baseline and demonstration periods.

Negative binomial regression models produce an incidence rate ratio (IRR) that is an estimate of that variable's effect on the dependent variable, after adjusting for the other variables in the model. An IRR greater than 1.0 is associated with an increased likelihood of acute care utilization; an IRR less than 1.0 means that the variable is inversely associated with utilization. We report the IRR associated with the test of the Difference-in-differences of the rate of hospitalizations and ER visits, and the incidence rate ratio's associated *p* value and 95% confidence interval.

6.2.2 Rates of 90-Day Readmissions

We estimated the percent of beneficiaries with at least one readmission within 90 days of discharge and the readmission rate per 1,000 beneficiaries. Readmissions are estimated for index admissions that occurred during 12-month spans in the baseline and demonstration periods. For the baseline, we included index admissions in the 12-month period immediately prior to the go-live date of MGH's CMP. Therefore, readmissions for baseline period admissions were counted through the first 3 months of the demonstration period. The intervention periods for the original population examined admissions during the periods of months 7 through 18 and months 22 through 33 and included readmissions through months 21 and 36, respectively. The intervention period for the refresh population examined admissions during months 10 through 21 and readmissions through month 24.

For all admissions, we calculated readmissions for any diagnosis (all-cause readmissions). For the subset of admissions for the 10 ACSC conditions, we calculated readmissions with a primary diagnosis in the same ACSC category (same cause readmissions). Because readmissions can only occur if there is an initial admission, admission rates can influence readmission rates. To provide context for readmission rate estimates, we estimated the percent of beneficiaries with an admission for any diagnosis and the percent with an admission for one of the 10 ACSC categories.

Readmission estimates were weighted by the fraction of days eligible until a readmission occurred or up to 90 days following an index hospitalization discharge, if there were no readmission within 90 days. For beneficiaries with more than one index hospitalization, the fraction was calculated by summing eligible days following each admission. To equalize the impact of differences in days of eligibility on readmission rates per 1,000 beneficiaries, counts of admissions were inflated by the fraction of days eligible following index hospitalizations.

The percent of beneficiaries with an admission, the percent with a readmission, and the readmission rate per 1,000 beneficiaries are presented for the baseline and demonstration periods for the intervention and comparison groups. For each measure, we compare the change from the baseline to the demonstration period for the intervention group relative to the comparison group

and test for the significance of this difference-in-differences rate between the groups. If MGH's CMP reduced admissions and readmissions, we expect to observe negative difference-in-differences rates, reflecting greater reductions or smaller increases in the intervention group relative to the comparison group.

6.2.3 Mortality

Another outcome metric in this evaluation is mortality. We constructed mortality rates per 100 beneficiaries and compare differences in mortality rates between the original and refresh intervention and comparison groups between the go-live date and the end of the demonstration period. Date of death was obtained from the Medicare EDB. Statistical comparison of the mortality rates was made using a *t*-test of differences in mean rates between the intervention and comparison groups.

We also conducted multivariate regression analysis to determine the predictors of mortality controlling for baseline differences in beneficiary demographic and health status characteristics between the intervention and comparison groups with inclusion of an indicator for being in the intervention group. Both a logistic model of the likelihood of death and a Cox proportional hazard model of survival were estimated testing the relationship of a large number of independent variables with likelihood of death or time to death. The independent variables used in the final multivariate Cox proportional hazard model of survival reported are defined as follows:

- male, a dichotomous variable, set at 1 for males;
- African American/other/unknown, a dichotomous variable, set at 1 for beneficiaries whose race code is African American, other, or unknown;
- age, three dichotomous variables set at 1 for age less than 65 years, age 75-84, and age greater than or equal to 85 years; age 65-74 is the reference group; and
- Medicaid, a dichotomous variable, set at 1 for beneficiaries enrolled in Medicaid. Medicaid enrollment is based on a beneficiary being enrolled in Medicaid at any point 1 year prior to the go-live date.

Baseline clinical and financial characteristics included in the model are defined as follows:

- baseline HCC score medium and high, two dichotomous variables set at 1 if the prospective HCC score was between 2.0 and 3.1 (medium) and greater than 3.1 (high); HCC score less than 2.0 is the reference group;
- baseline Charlson score medium and high, two dichotomous variables set at 1 if the Charlson index score was equal to 2, 3, or 4 (medium) and 5 or greater than (high); Charlson score of less than 2 is the reference group for the original population. For the refresh population, baseline Charlson scores of 2 or 3 were medium and 4 or greater were in the high group. The reference group was a score of less than 2.

- baseline PBPM costs medium and high, two dichotomous variables set at 1 if the PBPM cost calculated by RTI for a 12-month period prior to the *start* of the MGH's CMP original demonstration program was greater than or equal to $336.70 and less than $1,539 (medium) and $1,539 or greater (high); PBPM cost less than $336.70 is the reference group for the original population. For the refresh population, baseline PBPM costs greater than or equal to $324 and less than $1,408 were assigned to the medium group and $1,408 or greater to the high category; PBPM cost less than $324 is the reference group.

6.3 Findings

6.3.1 Rates of Hospitalizations and Emergency Room Visits

Rates of hospitalization and ER visits per 1,000 original population beneficiaries for the year prior to go-live and MGH's CMP demonstration periods are presented in ***Table 6-1***. Rates of hospitalization and ER visits are presented for all causes and for the 10 ACSCs. Next to the columns of the utilization rates are the difference-in-differences rates of change observed between the baseline and demonstration intervention periods. Negative difference-in-differences rates indicate that the intervention group's mean rate of hospitalization or ER visits declined more than the comparison group's mean rate or the intervention group's mean rate of hospitalization or ER visits grew at a lower rate than the comparison group's mean rate. Positive difference-in-differences rates, as statistically determined through the IRR, indicate that the comparison group exhibited either lower rates of growth or greater decline of hospitalization or ER visits than the intervention group. We report the odds ratio (OR) from the logistic regressions and the incidence rate ratio (IRR) from the negative binomial regressions of the difference-in-differences test along with the associated *p* value and 95% confidence interval. The last four columns contain the IRR and its statistical level of significance (*p*) value as well as the 95% confidence interval for the IRR.

Table 6-1
Comparison of rates of utilization for months 7-18 and the last 12 months of MGH's CMP with rates of utilization for a 1-year period prior to the start of the MGH's CMP: Original population

Utilization	Baseline rate per 1,000 I[1,2,3]	Baseline rate per 1,000 C[1,2,3]	Demo period rate per 1,000 I[1,2,3]	Demo period rate per 1,000 C[1,2,3]	D-in-D	IRR[4]	*p*-value	Low CI	High CI
Months 7-18									
Hospitalizations									
All-cause	767	726	863	1,072	-249	0.76	0.00	0.67	0.87
10 ACSCs[5]	202	203	284	391	-105	0.73	0.01	0.58	0.92
ER/Obs visits									
All-cause	1,207	1,256	1,397	1,695	-250	0.86	0.02	0.75	0.98
10 ACSCs	253	276	332	460	-105	0.79	0.03	0.64	0.97
Months 25-36									
Hospitalizations									
All-cause	677	670	878	1,072	-202	0.81	0.01	0.70	0.94
10 ACSCs	153	170	282	360	-61	0.87	0.31	0.66	1.14
ER/Obs visits									
All-cause	1,106	1,200	1,457	1,684	-133	0.94	0.41	0.81	1.09
10 ACSCs	203	255	317	443	-73	0.90	0.43	0.70	1.16

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ER/Obs = emergency room visits, including observation bed stays.

[1] The baseline period is the one-year period prior to the go-live date of the demonstration.

[2] Rates are per 1,000 beneficiaries adjusted for periods of program eligibility for the 1-year period prior to the start of the demonstration and for the demonstration program eligibility during two intervention periods.

[3] Only beneficiaries who at least 1 day of eligibility in the baseline and demonstration period are included in this analysis.

[4] Statistical testing of the difference-in-differences is conducted in STATA using negative binomial regression for rates/1,000 beneficiaries with robust variance estimation. The IRR is reported for negative binomial regressions. The p-value and confidence interval is reported for the IRR.

[5] The 10 ambulatory care sensitive conditions are as follows: Heart failure, Diabetes, Asthma, Cellulitis, COPD and Chronic Bronchitis, Dehydration, Bacterial Pneumonia, Septicemia, Ischemic Stroke, and urinary tract infection.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 27JUNE2010; acsctab acsc acsctab1 28JUNE2010.

Not unexpectedly, the baseline rates of hospitalization and ER visits were high in MGH's CMP original intervention and comparison populations. The baseline rate of all-cause hospitalization was 767 per 1,000 original intervention group beneficiaries (***Table 6-1***). And, the baseline rate of all-cause ER visits was 1,207 per 1,000 original intervention beneficiaries. Original population beneficiaries eligible for the later months of the demonstration had modestly lower baseline utilization rates reflecting the attrition through death of higher users of services. The 10 ACSC reasons for hospitalization combined accounted for roughly one-fifth to one-quarter of all-cause hospitalizations and all-cause ER visits. Thus, Medicare FFS beneficiaries in MGH's CMP were being treated in acute care settings quite frequently for prevalent chronic medical conditions, such as heart failure, diabetes, and COPD, as well as prevalent acute medical conditions such as pneumonia.

The rate of all-cause and ACSC hospitalization increased in the original intervention and the comparison groups between the baseline and both demonstration periods. However, the difference-in-differences rates are negative for all-cause and ACSC hospitalizations and ER visits during both demonstration periods indicating that the intervention rates increased less than the comparison group's rates. During months 7-18 of the demonstration, we observe statistically significant and clinically substantial lower difference-in-differences rates of growth within the intervention group for all four acute care measures. The rate of all-cause hospitalization grew 48% in the original comparison population during months 7-18 of the demonstration. In contrast, the rate of all-cause hospitalization grew only 13% in the intervention population. The rate of all-cause ER visits grew 35% in the original comparison population during months 7-18 of the demonstration while the rate of all-cause ER visits grew only 16% in the intervention population. We observe almost a doubling of the rate of hospitalization for ACSCs between baseline and months 7-18 of the demonstration within the comparison group in contrast to a 40% growth rate within the intervention group. A similar pattern is observed for ACSC ER visits in both groups.

During the last 12 months of the demonstration, there is less separation in rates of growth in acute care utilization between the intervention and comparison beneficiaries. During the first half of the demonstration, the rate of growth in the intervention group was one-third to one-half lower than the rate of growth in the comparison group. Although the rate of growth in the comparison group was even higher in the last 12 months of the demonstration than during months 7-18, the intervention group's rate of growth accelerated during months 7-18. For example, the rate of all-cause hospitalization grew 13% between baseline and months 7-18 and 30% between baseline and months 25-36 within the intervention group. In contrast, the rate of all-cause hospitalization grew 48% between baseline and months 7-18 and 60% between baseline and months 25-36 within the comparison group. There is a similar pattern across the other three acute care utilization measures. Although all four difference-in-differences rates are clinically significant, only one is statistically significant. The rate of all-cause hospitalization growth was lower within the intervention group than within the comparison group, 30% versus 60%, yielding a statistically significant difference-in-differences rate of -202 per 1,000 beneficiaries.

Rates of hospitalization and ER visits per 1,000 refresh population beneficiaries for the year prior to go-live and months 13-24 of MGH's CMP are presented in ***Table 6-2***. We observe roughly similar baseline rates of all-cause and ACSC hospitalizations and ER visits within the refresh intervention and comparison groups as we do for the original intervention and comparison groups. And, we observe similar patterns of increases in the hospitalization and ER

visit rates for both the intervention and comparison refresh groups during the demonstration period, and in a manner similar to the original population.

Table 6-2
Comparison of rates of utilization for the last 12 months of MGH's CMP with rates of utilization for a 1-year period prior to the start of MGH's CMP: Refresh population

Utilization	Baseline rate per 1,000 I[1,2,3]	Baseline rate per 1,000 C[1,2,3]	Demo period rate per 1,000 I[1,2,3]	Demo period rate per 1,000 C[1,2,3]	D-in-D	IRR[4]	*p*-value	Low CI	High CI
Months 13-24									
Hospitalizations									
All-cause	660	728	687	1,002	-247	0.76	0.04	0.58	0.99
10 ACSCs[5]	149	187	222	360	-99	0.78	0.29	0.49	1.24
ER/Obs visits									
All-cause	1,193	1,223	1,248	1,696	-419	0.75	0.04	0.57	0.99
10 ACSCs	193	256	256	425	-106	0.80	0.32	0.51	1.25

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ER/Obs = emergency room visits, including observation bed stays.

[1] The baseline period is the one-year period prior to the go-live date of the demonstration.

[2] Rates are per 1,000 beneficiaries adjusted for periods of program eligibility for the one-year period prior to the start of the demonstration and for demonstration program eligibility during the last 12 months MGH's CMP was active in the program.

[3] Only beneficiaries who at least one day of eligibility in the baseline and demonstration period are included in this analysis.

[4] Statistical testing of the difference-in-differences is conducted in STATA using negative binomial regression for rates/1,000 beneficiaries with robust variance estimation. The incidence rate ratio (IRR) is reported for negative binomial regressions. The p-value and confidence interval is reported for the IRR.

[5] The 10 ambulatory care sensitive conditions are as follows: Heart failure, Diabetes, Asthma, Cellulitis, COPD and Chronic Bronchitis, Dehydration, Bacterial Pneumonia, Septicemia, Ischemic Stroke, and urinary tract infection.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 27JUNE2010; acsctab acsc acsctab1 28JUNE2010.

We observe statistically significant differential rates of all-cause hospitalizations and ER visits during the demonstration period relative to the baseline period. The rate of all-cause hospitalization growth was lower within the intervention group than within the comparison group, 4% versus 38%, yielding a difference-in-differences rate of -247 per 1,000 beneficiaries. The rate of all-cause ER visit growth was lower within the intervention group than within the comparison group, 5% versus 39%, yielding a difference-in-differences rate of -419 per 1,000 beneficiaries.

6.3.2 Rates of 90-Day Readmissions

Table 6-3 displays the number of beneficiaries included in the readmission analyses. ***Table 6-4*** displays the percent of original population beneficiaries with an admission and 90-day readmission and rate of 90-day readmission per 1,000 beneficiaries. Data are displayed for all-cause and ACSC admissions and readmissions. In general, we observe a pattern of similar increases in the percent of both intervention and comparison beneficiaries being hospitalized or having a readmission over the course of the full demonstration with two noted exceptions. Relative to the baseline period, the percent of intervention beneficiaries with an all-cause or ACSC admission decreases during both demonstration periods while the percent of comparison beneficiaries with an all-cause or ACSC admission increases. The -4 to -11 percentage point differences are all statistically significant. We observe no statistically significant differences in the percentage of original intervention or comparison beneficiaries with an all-cause or ACSC same-cause readmission or the rate of readmission per 1,000 beneficiaries during the early stage of the demonstration (months 7-18) or during the last 12 months of the demonstration.

Table 6-3
Number of beneficiaries included in analyses of readmissions for MGH's CMP

Counts of beneficiaries	Intervention	Comparison
Original beneficiaries		
Months 7-18		
Total number of beneficiaries	2,427	2,490
Full time equivalents[1]	2,425	2,488
Months 22-33		
Total number of beneficiaries	2,054	2,045
Full time equivalents[1]	2,052	2,044
Refresh beneficiaries		
Months 10-21		
Total number of beneficiaries	716	680
Full time equivalents[1]	715	679

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program.

[1] Full Time Equivalent for the intervention group during the baseline period is the total number of beneficiaries weighted by their period of eligibility for the demonstration.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: readm01 readm02 27JUNE2010; readmtab1 28JUNE2010.

Table 6-4
Change in 90-day readmission[1] rates between the year prior to MGH's CMP and months 7-18 and months 22-33 of the demonstration: Original population

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo period rate per 1,000[1,2,3] I	Demo period rate per 1,000[1,2,3] C	D-in-D	OR/IRR[4]	*p*	Low CI	High CI
Months 7-18									
Hospitalizations									
Percent with an admission	42	37	39	45	-11	0.63	0.00	0.53	0.74
Percent with ACSC[5] admission	14	13	16	21	-6	0.68	0.00	0.55	0.85
All-cause 90-day readmission									
Percent with readmission	34	38	42	45	1	1.04	0.79	0.80	1.34
Readmission rate / 1,000	643	794	887	1055	-16	1.04	0.74	0.83	1.31
ACSC same-cause 90-day readmission									
Percent with readmission	13	14	13	18	-4	0.72	0.27	0.41	1.29
Readmission rate / 1,000	192	229	177	297	-83	0.71	0.31	0.37	1.37
Months 22-33									
Hospitalizations									
Percent with an admission	40	35	39	45	-10	0.65	0.00	0.55	0.78
Percent with ACSC admission	12	12	15	20	-4	0.73	0.02	0.57	0.95
All-cause 90-day readmission									
Percent with readmission	31	36	44	45	4	1.18	0.25	0.89	1.58
Readmission rate / 1,000	590	742	896	1057	-10	1.06	0.64	0.82	1.38
ACSC same-cause 90-day readmission									
Percent with readmission	11	11	16	16	-0	0.98	0.96	0.48	2.00
Readmission rate / 1,000	169	197	243	239	32	1.19	0.70	0.49	2.85

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odd ratio; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition.

1 Readmissions are defined as admissions that occur within 90 days after the discharge date of an index admission.

2 Rates are per 1,000 beneficiaries adjusted for periods of program eligibility for the one-year period prior to the start of the demonstration and for program eligibility during the demonstration period.

3 Only beneficiaries who at least one day of eligibility in the baseline and demonstration period are included in this analysis.

4 Statistical testing of the difference-in-differences is conducted in STATA using logistic regression for percentages and negative binomial regression for rates/1,000 beneficiaries. Robust variance estimation is used for both logistic and negative binomial regressions. The OR is reported for logistic regressions; the IRR is reported for negative binomial regressions. The *p*-value and confidence interval is reported for odds ratios and IRRs.

5 The 10 ambulatory care sensitive conditions are as follows: Heart failure, Diabetes, Asthma, Cellulitis, COPD and Chronic Bronchitis, Dehydration, Bacterial Pneumonia, Septicemia, Ischemic Stroke, and urinary tract infection.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: readm01 readm02 27JUNE2010; readmtab1 28JUNE2010.

Table 6-5 displays the percent of refresh population beneficiaries with an admission and readmission and rate of readmission per 1,000 beneficiaries. As with the original population, there is a general trend of increasing utilization over time. Relative to the baseline period, the percent of intervention beneficiaries with an all-cause admission decreases during the last 12 months of the demonstration period while the percent of comparison beneficiaries with an all-cause admission increases. The -10 percentage point difference is statistically significant (p=0.01). We observe no other statistically significant differences in the percentage of original intervention or comparison beneficiaries with an ACSC hospitalization, all-cause or ACSC same-cause readmission, or the rate of readmission per 1,000 beneficiaries during the last 12 months of the demonstration.

6.3.3 Mortality

Mortality rates during MGH's CMP for the original and refresh intervention and comparison populations are displayed in ***Table 6-6***. Over the 36-month demonstration period for the original population, 28% of beneficiaries in the intervention group died and 30% of beneficiaries in the comparison group died. The difference is not statistically significant. During the 24-month demonstration period for the refresh population, 16% of beneficiaries in the intervention group died while 20% of beneficiaries in the comparison group died. The 4 percentage point lower rate of mortality among intervention group beneficiaries is a statistically significant difference ($p<0.05$).

We further explored mortality in both the original and comparison populations by estimating a multivariate Cox proportional hazard model of survival. ***Figures 6-1*** and ***6-2*** display survival curves for the original and refresh populations, respectively. Each survival curve has two lines displaying the intervention and comparison groups' unadjusted survival differences. Not unexpectedly, the LifeTest procedure reveals that there is no statistically significant difference in survival between the original intervention and comparison groups; however, we do observe a statistically significant difference between the refresh intervention and comparison groups.

Table 6-7 displays two Cox Proportional Hazard multivariate models of survival for the original population. ***Table 6-8*** displays similar models for the refresh population. The censoring variable is death. Model 1 has a single dichotomous variable whereby intervention group status=1; comparison group status=0. Model 2 steps in a number of baseline covariates to control for any differences between the two groups at baseline. The hazard ratios and associated p values are displayed for both sets of models' independent variables. The hazard ratio can be interpreted as the odds that an individual in the group with the higher hazard reaches the endpoint first, and vice versa. In our case, the endpoint is death.

Table 6-5
Change in 90-day readmission[1] rates between the year prior to MGH's CMP and months 10-21 of the demonstration: Refresh population

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo period rate per 1,000[1,2,3] I	Demo period rate per 1,000[1,2,3] C	D-in-D	OR/IRR[4]	p	Low CI	High CI
Months 10-21									
Hospitalizations									
Percent with an admission	40	36	37	43	-10	0.66	0.01	0.48	0.90
Percent with ACSC[5] admission	11	13	14	21	-4	0.79	0.28	0.51	1.21
All-cause 90-day readmission									
Percent with readmission	27	43	32	45	4	1.22	0.44	0.73	2.02
Readmission rate / 1,000	552	843	684	1015	-41	1.03	0.91	0.65	1.63
ACSC same-cause 90-day readmission									
Percent with readmission	10	10	10	13	-3	0.72	0.62	0.19	2.65
Readmission rate / 1,000	169	110	188	205	-76	0.60	0.47	0.15	2.40

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odd ratio; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition.

[1] Readmissions are defined as admissions that occur within 90 days after the discharge date of an index admission.

[2] Rates are per 1,000 beneficiaries adjusted for periods of CMHCB program eligibility for the one-year period prior to the start of the demonstration and for CMHCB program eligibility during the demonstration period.

[3] Only beneficiaries who at least one day of eligibility in the baseline and demonstration period are included in this analysis.

[4] Statistical testing of the difference-in-differences is conducted in STATA using logistic regression for percentages and negative binomial regression for rates/1,000 beneficiaries. Robust variance estimation is used for both logistic and negative binomial regressions. The OR is reported for logistic regressions; the IRR is reported for negative binomial regressions. The *p*-value and confidence interval is reported for odds ratios and IRR.

[5] The 10 ambulatory care sensitive conditions are as follows: Heart failure, Diabetes, Asthma, Cellulitis, COPD and Chronic Bronchitis, Dehydration, Bacterial Pneumonia, Septicemia, Ischemic Stroke, and UTI.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: readm01 readm02 27JUNE2010; readmtab1 28JUNE2010.

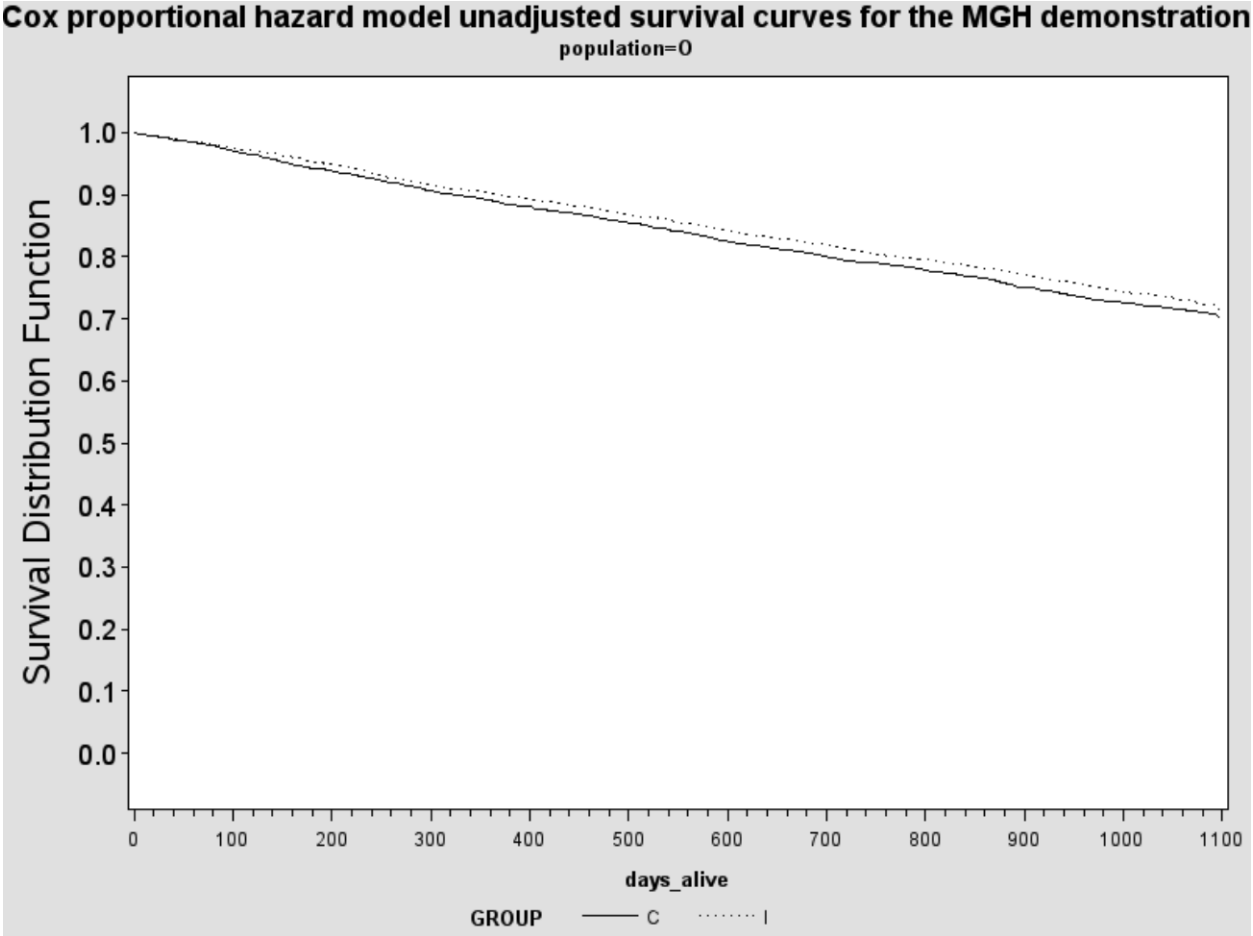
Table 6-6
Mortality rates during MGH's CMP: Original and refresh populations

Description	Intervention number of deaths	Percent	Comparison number of deaths	Percent	Difference	P value
Original population (36 months)	721	27.9%	803	29.5%	-1.63	0.19
Refresh population (24 months)	121	15.6%	150	19.6%	-3.97	0.04

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program.

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: mortality.sas 27JUNE2010.

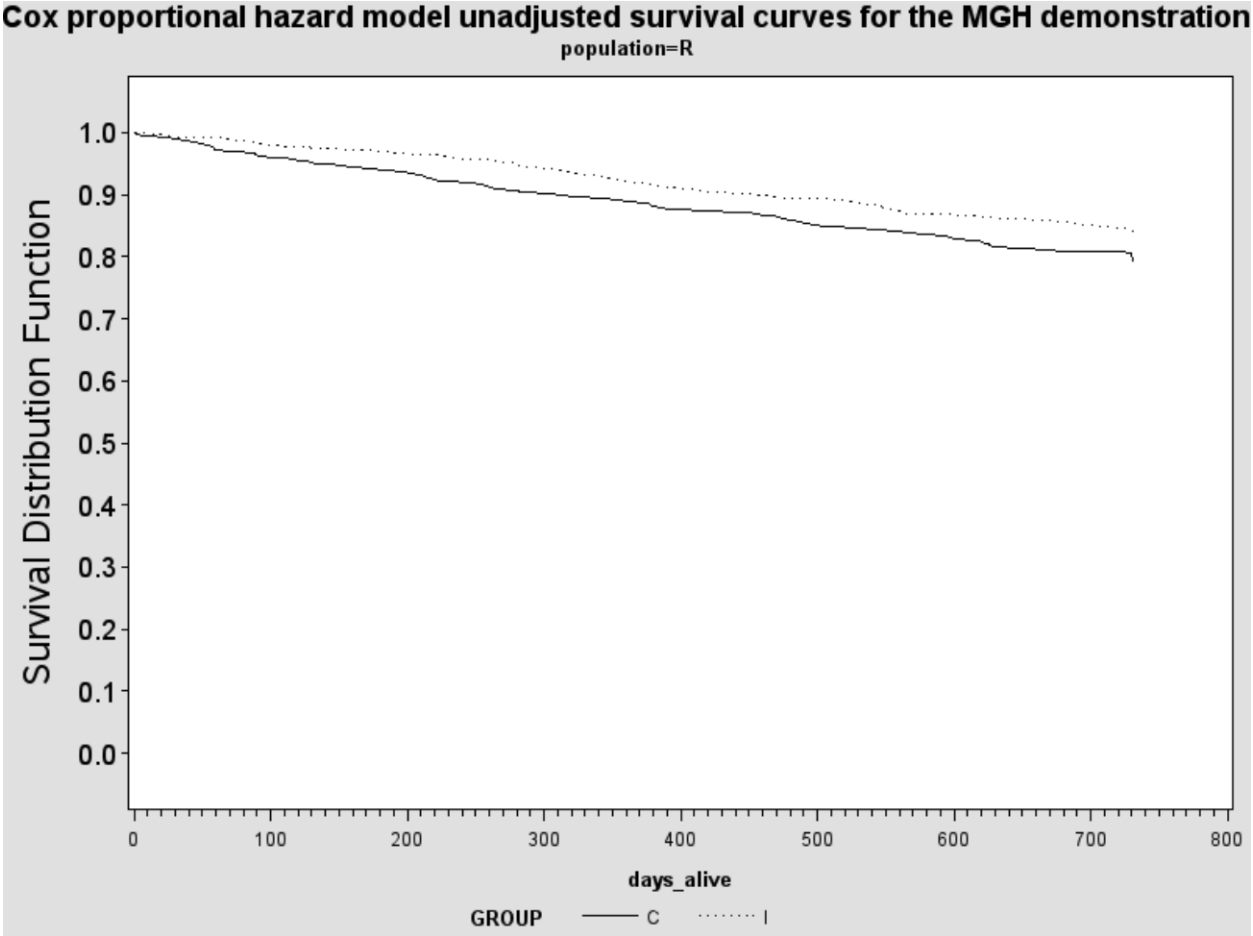
Figure 6-1
Cox proportional hazard model unadjusted survival curves for MGH's CMP demonstration original population



NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program.

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: dietab3.sas 30JULY2010.

Figure 6-2
Cox proportional hazard model unadjusted survival curves for MGH's CMP demonstration refresh population



NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program.

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: dietab3.sas 30JULY2010.

Table 6-7
Cox Proportional Hazard Survival Models for MGH's CMP: Original Population

Original	Model 1 Hazard Ratio	Model 1 p value	Model 2 Hazard Ratio	Model 2 p value
Intervention	0.939	0.2098	0.894	0.0286
Male	N/I	N/I	1.140	0.0118
African American/other/unknown	N/I	N/I	0.932	0.3939
Age < 65 years	N/I	N/I	0.774	0.0134
Age 75-84	N/I	N/I	1.615	<.0001
Age 85 + years	N/I	N/I	2.745	<.0001
Medicaid	N/I	N/I	0.984	0.9054
Baseline HCC score medium	N/I	N/I	0.961	0.6328
Baseline HCC score high	N/I	N/I	1.280	0.0079
Medium baseline PBPM cost	N/I	N/I	0.944	0.4613
High baseline PBPM cost	N/I	N/I	1.519	<.0001
Baseline Charlson score medium	N/I	N/I	1.021	0.8022
Baseline Charlson score high	N/I	N/I	1.517	<.0001

NOTES: MGH'S CMP = Massachusetts General Hospital's Care Management Program; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: dietab4.sas 31August2010.

Table 6-8
Cox Proportional Hazard Survival Models for MGH's CMP: Refresh Population

Original	Model 1 Hazard Ratio	Model 1 p value	Model 2 Hazard Ratio	Model 2 p value
Intervention	0.743	0.0127	0.791	0.0531
Male	N/I	N/I	1.026	0.8349
African American/other/unknown	N/I	N/I	0.992	0.9640
Age < 65 years	N/I	N/I	0.368	0.0004
Age 75-84	N/I	N/I	1.219	0.2142
Age 85 + years	N/I	N/I	2.293	<.0001
Medicaid	N/I	N/I	1.364	0.2561
Baseline HCC score medium	N/I	N/I	0.963	0.8612
Baseline HCC score high	N/I	N/I	1.860	0.0064
Medium baseline PBPM cost	N/I	N/I	0.760	0.1752
High baseline PBPM cost	N/I	N/I	1.560	0.0256
Baseline Charlson score medium	N/I	N/I	0.724	0.1520
Baseline Charlson score high	N/I	N/I	1.224	0.3692

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: dietab4.sas 31August2010.

In ***Table 6-7***, Model 1, we observe that the intervention variable has a hazard ratio of 0.939 that is not statistically different from 1. In Model 2, we now observe a survival advantage among the original intervention beneficiaries when baseline covariates are added adjusting for any imbalances between intervention and comparison groups at baseline. The hazard ratio declines modestly to 0.89 and is statistically significant. ***Table 6-8*** displays regression results for the refresh population. In Model 1, we observe that the intervention variable has a hazard ratio of 0.743 that is significantly lower than 1 and consistent with the unadjusted mortality rates reported in ***Table 6-6***. However, in Model 2, when baseline covariates are added the hazard ratio rises to 0.791 and is statistically significant at the 0.053 level. However, the modeling of mortality was quite sensitive to model specification.

In provider-based analyses without randomization within practice, any positive intervention effect may be due to the quality of the care provided by the participating practice(s) relative to the practices selected for the comparison groups. Being loyal to the intervention practices could be the intervention effect or we could be observing a CMP case manager intervention effect. However, we cannot determine from the structure of the demonstration whether we are observing a practice or case manager effect. However, the positive mortality finding, albeit sensitive to the estimation method, suggests that further investigation of the appearance of a survival benefit is desirable through analysis of the MGH Phase II demonstration mortality experience.

6.3.4 Hospice

A focus of MGH's CMP was encouraging appropriate end-of-life-care planning, including use of the hospice benefit. We examine rates of hospice use between the intervention and comparison groups for both the original and refresh populations. ***Table 6-9*** provides the hospice rates and the mean and median days in hospice. We observe low use rates of the Medicare hospice benefit among the original and refresh intervention and comparison populations, ranging from 6% to 12% (statistically insignificant). Length of hospice is also quite similar between the intervention and comparison groups (no differences are statistically significant) for both the original and refresh populations.

Table 6-9
Comparison of Hospice Use among Beneficiaries that Died During MGH's CMP Compared to Those That Remained Alive

	Intervention N	Comparison N	Hospice Rate I	Hospice Rate C	I vs. C	p value	Mean Days I	Mean Days C	I vs. C	p value	Median Days I	Median Days C	I vs. C	p value
Original population														
All	2,584	2,719	12%	11%	1.4	0.12	65	53	12	0.18	19	18	1	0.78
Alive	1,863	1,916	2%	1%	0.5	0.18	242	198	44	0.53	187	94	93	0.31
Deceased	721	803	40%	34%	5.5	0.03	46	42	5	0.46	17	16	1	0.90
Refresh population														
All	775	766	6%	9%	-3.5	0.01	60	50	10	0.48	24	20	4	0.68
Alive	654	616	1%	1%	-0.2	0.70	182	140	42	0.59	133	148	-15	1.00
Deceased	121	150	32%	43%	-10.4	0.08	41	40	1	0.92	14	19	-5	0.93

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: hsp01 24JUNE2010, hospicetab 27JUNE2010, hsptest 09JULY2010.

6.4 Conclusions

RTI's analysis of health outcomes focuses on measuring effectiveness of MGH's CMP within the CMHCB demonstration by answering the following evaluation questions:

- Did MGH's CMP improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?

- Did MGH's CMP improve health outcomes by decreasing mortality?

During the course of MGH's CMP, we generally observe increasing rates of all-cause and ACSC hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations. However, MGH's CMP was successful at substantially reducing the rate of increase in all-cause and ACSC hospitalizations and ER visits among the original intervention beneficiaries during the first half of the demonstration. We also observe statistically significant declines in the percent of original intervention beneficiaries with an all-cause or ACSC admission during the first half of the demonstration. However, we observe no statistically significant differences in the percentage of original intervention or comparison beneficiaries with an all-cause or ACSC same-cause readmission or the rate of readmission per 1,000 beneficiaries during months 7-18 of the demonstration.

During the last 12 months of the demonstration, there is less separation in rates of growth in acute care utilization between the original intervention and comparison beneficiaries. Yet, we still observe statistically significant reductions in the rate of all-cause hospitalization growth and the percent of beneficiaries with an all-cause or ACSC admission. We continue to observe no statistically significant differences in the percentage of original intervention or comparison beneficiaries with an all-cause or ACSC same-cause readmission or the rate of readmission per 1,000 beneficiaries.

Also during the last 12 months of the demonstration, we observe statistically significant lower rates of all-cause hospitalization and ER visit growth and percent of beneficiaries with an all-cause admission within the refresh intervention group. There are also clinically meaningful reductions in the rate of growth of ACSC hospitalizations and ER visits but they do not achieve statistical significance likely due to the small size of the refresh population. We do not observe statistically significant differences in the percentage of refresh intervention or comparison beneficiaries with an all-cause or ACSC same-cause readmission or the rate of readmission per 1,000 beneficiaries.

The success of reducing hospitalizations but not readmissions is somewhat surprising given the emphasis of MGH's CMP on managing care transitions across settings. In the CMP, case managers follow a protocol for every transition of care between settings. All case managers in the program are trained and expected to adhere to these protocols. The protocols include step by step daily workflow instructions for the following transitions of care: emergency department admissions, inpatient MGH admissions, post hospital discharges to home from MGH, post hospital discharges to other facilities, and post discharge from post acute facility to home. In addition a post-episode assessment is completed within 24-72 hours of the patient's discharge

from the MGH emergency department or inpatient unit, and other acute or post-acute care facilities, if known. The post-episode assessment determines the patient's overall health status; knowledge of their discharge instructions, medication changes, home care services, and follow-up appointments with a primary care provider and/or specialists. The post-episode case manager conducts a follow-up with patients 24-72 hours post discharge. The patient's primary care-based case manager interfaces with the MGH inpatient case manager during the admission and prior to discharge. Pertinent information is shared with the post-episode case manager to assist her with the assessment. If the patient's follow up appointments have not already been scheduled, the post-episode case manager informs the primary case manager who takes responsibility for scheduling. This level of effort to prevent readmissions and the disappointing results suggests that broadly reducing readmissions among chronically ill Medicare FFS beneficiaries may be a far bigger challenge than has been envisioned by MGH's CMP leadership. In Phase II, MGH's CMP has been granted a demonstration waiver to allow for direct admission to a skilled nursing facility (SNF) beneficiaries who meet specific clinical criteria. It will be important to examine if the SNF waiver is a tool that can be used to reduce readmissions of patients who become clinically unstable after discharge.

One component of MGH's CMP was end-of-life planning including advance directives and use of hospice. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rate of the Medicare hospice benefit or in mean or median number of days of hospice.

Another key outcome metric is mortality. Over the 36-month demonstration period for the original population, 28% of beneficiaries in the intervention group died and 30% of beneficiaries in the comparison group died. This is not a statistically significant differential rate of mortality until adjustment for differences in baseline characteristics is introduced through the estimation of a Cox proportional hazard model. For the refresh population, we do observe a lower unadjusted rate of mortality within the refresh intervention group than the comparison group. During the 24-month demonstration period for the refresh population, 16% of beneficiaries in the intervention group died while 20% of beneficiaries in the comparison group died. Introduction of baseline characteristics in the multivariate modeling mitigates somewhat the statistical significance of this difference. Thus, there is evidence of a mortality benefit, albeit sensitive to the estimation method.

CHAPTER 7
FINANCIAL OUTCOMES

7.1 Introduction

In this section, we present *final* evaluation findings on levels and trends in Medicare costs for the year prior to the go-live date and over the full 36 months that MGH's CMP was in operation (or 24 months for the refresh population). The evaluation questions we address are:

- What were the Medicare costs per beneficiary per month (PBPM) in the base year versus the first 36 or 24 months of the demonstration for the intervention and the comparison groups?

- What were the levels and trends in PBPM costs for intervention group participants and nonparticipants? Did nonparticipation materially reduce the intervention's overall cost savings?

- How variable are PBPM costs in this high cost, high risk, population? What was the minimal detectable savings rate given the variability in beneficiary PBPM costs?

- How did Medicare savings for the 36- or 24-month period compare with the fees that were paid out? How close was MGH's CMP in meeting budget neutrality?

- How balanced were the intervention and comparison group samples prior to the demonstration's start date? How important were any measured imbalances to the estimate of savings?

- Did the intervention have a differential effect on high-cost and high-risk beneficiaries?

- What evidence exists for regression-to-the-mean (RtoM) in Medicare costs for beneficiaries in the intervention and comparison groups?

The cost analyses presented in this section differ from those conducted for financial reconciliation by ARC under contract to CMS. ARC determined savings based on the demonstration's terms and conditions negotiated between CMS and MGH. RTI's estimation of savings, detailed subsequently, differs in that

- differences in savings rates between intervention and comparison groups are first determined at the beneficiary level and are then tested using statistical confidence intervals,

- beneficiary PBPM costs are not trimmed using a 1% outlier dollar threshold, and

- both base year and demonstration period PBPM costs are weighted by each beneficiary's fraction of eligible days during the demonstration period.

A more detailed explanation and justification for these differences is provided in ***Section 7.3***.

The rest of this chapter has five sections. The next two sections describe our data sources, variable construction, and analytic methods. ***Section 7.4*** presents our primary findings on trends in PBPM costs between base and demonstration periods. ***Section 7.5*** shows PBPM cost savings in relation to average monthly fees and whether MGH's CMP achieved budget neutrality using RTI's costing methods. ***Section 7.6*** displays stratified PBPM costs and savings by high cost and high risk categories to test for possible imbalances in the intervention and comparison groups. ***Section 7.7*** examines regression-to-the-mean (RtoM) effects. ***Section 7.8*** uses multivariate regression to control for any imbalances between intervention and comparison groups that might affect t-tests of mean differences in PBPM growth rates. The chapter concludes in ***Section 7.9*** with a summary of key findings.

7.2 Data and Key Variables

7.2.1 Population Frame and Data

The data used in RTI's analysis of PBPM costs are Medicare Parts A and B claims extracted for all eligible beneficiaries in the original and refresh intervention and comparison groups as described in ***Chapter 1***. Because of more than a year's gap between selection for and the start of the original demonstration, a new base year of claims data were extracted for the intervention and comparison populations. Consequently, it is likely that some beneficiaries who originally qualified during the randomization process would no longer qualify for the demonstration during the base period just 1 year before the MGH CMP's start date. They still remain in the intervention and comparison groups, however, for our analysis.

We restrict all analyses to beneficiaries who were alive at the start date of the demonstration. Claims costs are accumulated until a beneficiary dies or otherwise becomes ineligible (e.g., joins a Medicare Advantage plan). Claims represent utilization anywhere in the United States, not just the target area of the MGH CMP. Medicare costs are based on eligible claims submitted during the full demonstration period plus 12 months prior to the start date. A 9-month "run-out" period after the demonstration ended assures a complete set of costs.

7.2.2 Constructing PBPM costs

All financial analyses were conducted on a PBPM cost basis, or the ratio of eligible Medicare costs to eligible months. The baseline period is defined as 365 days (or 1 year) prior to the MGH CMP's start date. The 36-month demonstration period for the original population includes 1,095 days (36 months × 30.42 days/month) after the start date. The refresh population covers 24 months, or 730 days.

Medicare program costs in the numerator of PBPM costs include

- only Medicare program Part A and B payments; patient obligations and Part C (managed care) and D (drugs) are excluded;

- only claims for utilization of beneficiaries when they are eligible for the demonstration[8]; and

- only claims for eligible services; end-stage renal disease [ESRD] services are excluded.

To statistically test hypotheses regarding *trends* in beneficiary costs, average PBPM costs first must be calculated at the beneficiary level. Constructing individual PBPM costs required dividing a beneficiary's total cost during eligible periods by his or her own fraction of eligible months during the base year and the demonstration period. Most beneficiaries had 12 months of base year eligibility and 36 or 24 months of demonstration period eligibility. However, some beneficiaries had fewer than the maximum number of eligible months (or days), usually due to death. At the extreme, a beneficiary could have a 10-day hospital admission at the beginning of the intervention period with a combined Part A and B payment of $30,000 before dying. This $30,000 outlay is divided by approximately 1/3 (10 days / 30.42 days), resulting in an adjusted PBPM cost outlay of $90,000. Consequently, (unweighted) PBPM costs exhibit substantial variation that, in turn, reduces the likelihood of finding statistical differences.

Table 7-1 shows unweighted mean *intervention* group PBPM costs in MGH's original population (2,584 with eligible days in both the base and intervention period) stratified by beneficiaries' number of eligible days in the demonstration period (1,095 maximum). The 6 beneficiaries with <10 eligible days averaged $15,686 compared with beneficiaries eligible for a year or more who averaged PBPM costs of $2,194. Beneficiaries with truncated eligibility averaged monthly costs 7.1 times greater than those with more than a year's eligibility. Roughly 3% of the sample was eligible less than 3 months. (See ***Section 7.3.2*** for statistics on PBPM cost variation.) Maximum demonstration period PBPM costs were $66,132.

Table 7-2 shows the unweighted cost effects of short-term eligible beneficiaries in the much smaller (775 beneficiaries) *refresh* population. Again, short-eligibility refresh beneficiaries were several times as costly per month as those with more than 1 year of eligibility. Maximum PBPM costs for the refresh population were $42,603.

Variation can be reduced by trimming high PBPM cost outliers at the 99th percentile, as done by CMS for financial reconciliation. While the 1% trim reduces the MGH CMP's financial risk, we wanted to avoid biasing comparisons against interventions that constrained spending among the most expensive beneficiaries.

---

8 For example, if a beneficiary joined a managed care plan for a few months then returned to fee for service (FFS) Medicare, any claims for plan services were excluded.

Table 7-1
MGH's CMP PBPM mean costs by eligible days, intervention group, demonstration period: Original population

Eligible days[1]	N (%)	PBPM Cost	Range
< 10	6 (0.2%)	$15,686	$0–66,132
11–30	17 (0.7)	6,766	0–32,199
31–60	12 (0.5)	17,439	184–41,798
61–90	16 (0.6)	10,533	410–49,476
91–365	198 (7.7)	6,594	5–36,277
366+	2,335 (90.3)	2,194	3–26,908
Mean/Total	2,584	2,716	0–66,132

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N (%) = number of beneficiaries (percent of all eligibles).

[1] Number of days beneficiary eligible for intervention.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/27/10).

Table 7-2
MGH's CMP PBPM mean costs by eligible days, intervention group, demonstration period: Refresh population

Eligible days[1]	N (%)	PBPM Cost	Range
< 10	0 (0.0)	$–	$–
11–30	3 (0.4)	17,218	2,839–42,603
31–60	5 (0.7)	8,019	31–15,645
61–90	1 (0.1)	93	93–93
91–365	40 (5.2)	6,682	33–32,969
366+	726 (93.7)	1,826	0–20,813
Mean	775	2,174	0–42,603

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N (%) = number of beneficiaries (percent of all eligibles).

[1] Number of days beneficiary eligible for intervention.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/27/10).

Instead of trimming or deleting outliers, RTI weighted PBPM mean costs and standard errors by each beneficiary's eligible fraction of days, or exposure to the intervention. In the previous example, the beneficiary's adjusted $90,000 PBPM cost is weighted by 10/1,095 = 0.009 in the original population, or roughly 110-times less than beneficiaries with full eligibility through the entire demonstration period. This weighting method is equivalent to simply adding the beneficiary's $30,000 and 10 eligible days to total costs and days of fully eligible beneficiaries and then calculating the combined PBPM cost.

7.2.3 Monthly Fees

Demonstration Care Management Organizations (CMOs) proposed monthly fees when submitting their applications for the demonstration program to the CMS Office of Demonstrations. CMS then negotiated final fees as part of each CMO's agreed-upon contract terms and conditions. RTI benchmarked savings against each CMO's initially negotiated fee. For MGH's CMP, its negotiated management fee was a constant $120 for all 3 demonstration years for the original intervention group. Fees for the refresh intervention group also were $120 for the last 2 years of the demonstration. To be consistent with the calculation of gross savings, these two fees were weighted by the share of fee-bearing to all eligible months in the intervention group.

7.3 Analytic Methods

RTI's analytic approach is based on a *comparison of growth rates in PBPM costs at the individual beneficiary level*. This approach has two principal strengths:

- First, it controls in a more precise, beneficiary-specific manner for any differences in PBPM costs between the base year and the demonstration period that are not accounted for through the selection process.

- Second, by calculating changes in PBPM costs at the beneficiary level (i.e., "paired" base-demonstration period PBPM costs), we can conduct statistical *t*-tests of the differences in spending growth rates between intervention and comparison groups.

In addition to answering the question of whether any or all of the CMHCB demonstration programs achieved budget neutrality (or even any savings), we also are interested in *generalizing* results to future care management activities by answering the question, "What savings are likely to be realized if the demonstration is expanded?" This question necessarily requires testing the hypothesis that any savings in a sample of beneficiaries during a particular time period could have been caused by chance with no long-run implications. RTI conducted a range of analyses to answer the key financial questions.

7.3.1 Tests of Gross Savings

Gross savings to Medicare is defined as the difference between the claims costs of the intervention and comparison groups. There are two ways to calculate these differences. Assuming that the selection process balanced the intervention and comparison populations, PBPM cost differences between the two groups can be based solely on the demonstration period. That is, MGH's CMP was neither advantaged nor disadvantaged by the costliness of their intervention group relative to their comparison group. However, more than 1 year passed between the time the beneficiaries were assigned to the intervention and comparison groups and when MGH's CMP began recruiting beneficiaries to the intervention. Also, because we wanted to conduct statistical tests of intervention effects, it was necessary to construct PBPM cost estimates at the beneficiary level and then use variation in the observations to produce confidence intervals around the estimates.

Recognizing that base year costs may be different between intervention and comparison populations, we used a mixed paired sample approach. First, we used each beneficiary's own mean PBPM costs in the base year just prior to the MGH CMP's start date and the intervention period to construct a change in costs. This was done for all beneficiaries in both the intervention and comparison groups, thereby producing a paired comparison within group. Next, we determined the mean difference in the differences in PBPM cost growth rates for each group, treating the mean differences as independent samples.[9] The strength of first calculating the change in PBPM costs at the beneficiary level is that it completely controls for any unique clinical and socioeconomic characteristics that might differ between the intervention and comparison groups. Any imbalances in beneficiary characteristics that might produce inter-

[9] For a more detailed description of this approach, see Rosner (2006, chapter 8).

temporal differences in medical utilization or costs are factored out using first-differencing. Our gross savings rate, in equation form, is

$$\text{Gross Savings} = \text{Diff}[I] - \text{Diff}[C] = [I_t^* - I_b^*] - [C_t^* - C_b^*] = \Delta I^* - \Delta C^* \qquad (7.1a)$$

$$\text{Gross Savings} = [I_t^* - C_t^*] - [I_b^* - C_b^*], \qquad (7.1b)$$

where * = the mean difference in PBPM costs within all intervention (I) or comparison (C) beneficiaries, t and b = demonstration and base periods, and Δ = the change in PBPM costs between the base and demonstration periods. Savings, as the difference-in-(paired) differences, is equivalent to adjusting the difference in intervention and comparison means during the demonstration by the mean difference that existed in the base year (eq. 7.1b).

In calculating mean changes in PBPM costs across beneficiaries, each beneficiary's *change* needs to be weighted to produce an unbiased estimate of the overall mean change. We used the beneficiary's fraction of eligible days during the demonstration period as weights. This effectively weights each beneficiary's base period PBPM costs by their proportion of days during the demonstration period. Consequently, early demonstration dropouts (usually due to death) will have their base period PBPM costs underweighted relative to their actual contribution when displaying base period mean costs for intervention or comparison groups. As early demonstration dropouts tend to be more costly in the base period, our mean base year costs will appear lower than actuarial means based on their proportion of days during the base period. It did not seem reasonable to give beneficiaries with only a few days involvement in the actual demonstration full credit in calculating mean base year costs even if they had 12 months of base year Medicare eligibility.

7.3.2 Detectable Savings

In all of the analyses in this chapter, we test the hypothesis of whether gross savings is statistically different from zero, or no savings. Gross savings must be sufficiently greater than zero to assure CMS that the measured savings rate was not due to chance.[10] A critical evaluation question is the power we had to detect relatively small savings rates. By "detectable" we mean the rate of savings that would force us to reject the null hypothesis of no savings at all. Having completed the demonstration, we now have the information on both the level and variation in savings rates that allows us to calculate the detectable savings threshold for MGH's CMP.

The fundamental test statistic is the Z-ratio of gross savings (see eq. 7.1a) to its standard error (SE)

$$Z = [\Delta I - \Delta C]/SE_{[\Delta I - \Delta C]} \qquad (7.2)$$

[10] Chance savings can occur primarily because of random fluctuations in the utilization of health services required in the intervention and comparison groups. It is possible that random declines in health in the intervention group unrelated to the intervention could explain lower savings rates.

$$SE_{[\Delta I - \Delta C]} = [SE_{\Delta I}{}^2 + SE_{\Delta C}{}^2]^{0.5}. \qquad (7.3)$$

A two-sided test[11] of intervention savings uses the following confidence interval:

$$-1.96\ SE_{[\Delta I - \Delta C]} <= \text{Savings} <= 1.96\ SE_{[\Delta I - \Delta C]}, \qquad (7.4)$$

and the detectable threshold is

$$\text{Detectable Threshold (DT)} = -1.96\ SE_{[\Delta I - \Delta C]}. \qquad (7.5)$$

Intervention savings must equal or exceed -1.96 times the standard error of the difference in the growth in intervention and comparison PBPM costs. (Savings are expressed in negative terms if intervention PBPM cost growth is less than the comparison group cost growth.) The detectable threshold (DT) is approximately double the standard error of the difference in mean growth rates, which in turn varies with the square root of the intervention and comparison group sample sizes. It is also convenient for some analyses to express the DT as a percent of the comparison group's demonstration mean PBPM cost, or $DT/PBPM_C$.

Table 7-3 and ***7-4*** show the variation that exists in the (unweighted) PBPM costs in the base year prior to the start date and the demonstration period for the MGH CMP's intervention and comparison, original and refresh populations. Mean PBPM costs in the base period ranged from a low of $0 to a high of $27,438 in the original comparison group. The coefficient of variation (CV), or the standard deviation of beneficiary-level PBPM costs divided by the mean, is fairly large (about 1.47) in the base year (standard deviations roughly 47% greater than mean costs). CVs in the original and refresh comparison groups were little changed during the demonstration period. Some of the variation is reduced after weighting observations when determining intervention savings later in this chapter.

[11] A reasonable argument can be made that the detectable threshold should be based on a one-sided *t*-test if one assumes that any chronic care management intervention would not be expected to *increase* Medicare outlays. If an intervention is likely only to reduce costs, a one-sided test effectively puts all 5% of the possible error on the negative side, resulting in a detectable threshold only -1.68 times the standard error.

Table 7-3
MGH's CMP PBPM cost distribution thresholds, comparison and intervention group, base, and demonstration period: Original population

Quantiles[1]	Base year comparison	Base year intervention	Demonstration period comparison	Demonstration period intervention
(N)	(2,719)	(2,584)	(2,719)	(2,584)
Minimum	$0	$0	$0	$0
<10%	0	0	302	244
<25%	241	235	682	514
Median	716	745	1,707	1,459
>75%	2,179	2,202	3,866	3,422
>90%	4,823	4,413	7,443	6,313
Maximum	27,438	28,268	116,201	66,132
Mean	1,732	1,660	3,211	2,716
CV	1.47	1.42	1.60	1.46

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N = number of beneficiaries; CV = coefficient of variation.

[1] <10%, <25%, >75%, >90%: PBPMs below or above percentage.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/27/10).

Table 7-4
MGH's CMP PBPM cost distribution thresholds, comparison and intervention group, base and demonstration period: Refresh population

Quantiles[1]	Base year comparison	Base year intervention	Demonstration period comparison	Demonstration period intervention
(N)	(766)	(775)	(766)	(775)
Minimum	$0	$0	$0	$0
<10%	0	0	237	173
<25%	223	220	487	360
Median	718	654	1,483	977
>75%	2,167	1,942	3,783	2,655
>90%	5,105	4,292	7,393	5,576
Maximum	53,801	30,647	56,618	42,603
Mean	1,827	1,569	3,057	2,174
CV	1.75	1.61	1.56	1.63

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N = number of beneficiaries; CV = coefficient of variation.

[1] <10%, <25%, >75%, >90%: PBPMs below or above percentage.

SOURCE: Medicare 2004-2008 Part A & B claims; .COSTRUN2 (6/27/10).

The difference between median and mean PBPM costs indicates how skewed costs actually are. Mean costs are usually more than double median costs in the original population's base year with little change during the intervention period, indicating a strong right tail of very high costs. Costs were similarly skewed in the refresh group (***Table 7-4***). Note that 25% of refresh comparison beneficiaries had base year PBPM costs less than $223 (roughly $3,000 on an annual basis). Maximum values show how high PBPM costs can be before weighting. As shown earlier in ***Table 7-1***, these costs are often incurred by beneficiaries with very short eligibility who died very early in the demonstration period. Weighting these short-eligible, very high cost beneficiaries reduces overall variance and produces lower detectable thresholds.

Because of the relatively large variances in the base year PBPM costs (CV[original comparison] = 1.47), coupled with adjustments for the repeated nature of the experimental design, the power afforded by the original population sizes was very low, i.e., about 10%.[12]

7.3.3 Budget Neutrality

Each CMO is obligated to produce net savings for the Medicare program. The net savings requirements for those CMOs that complete the full demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. Thus, to avoid paying back any fees in the original population,

$$PBPM_I <= 0.95PBPM_c - MF \quad (7.6a)$$

or as a fraction of the comparison PBPM cost,

$$PBPM_I/PBPM_c <= 0.95 - (MF/PBPM_c), \quad (7.6b)$$

where $PBPM_I$, $PBPM_C$ = average monthly costs in the intervention and comparison groups, MF = the average monthly fee.

For example, if a CMO's monthly fee were 5% of the comparison PBPM cost, then intervention PBPM costs would have to be 90% or less of monthly comparison costs to avoid paying back fees. Debt obligation per intervention beneficiary month is the positive difference:

$$PBPM_I - [0.95PBPM_c + MF]. \quad (7.6c)$$

RTI's conclusion regarding budget neutrality will differ from those of CMS during financial reconciliation, given the way we adjust for unequal base period costs, how fees are calculated, the lack of an outlier trim, and a few other minor differences. Because we use statistical confidence intervals to judge the extent of gross savings, we test whether a CMO achieved any savings at all: the Z-test against zero savings.

In addition to Z-tests of mean cost differences between the entire intervention group and the comparison group, we also tested for differences in PBPM cost growth rates between intervention beneficiary participants and nonparticipants relative to the comparison group. If the intervention had more success with those beneficiaries it actively engaged, then savings should be greater for participants than nonparticipants.

---

[12] Power for a comparison of two mean changes in PBPMs is given by $\Phi[-1.96 + (\sqrt{n}\Delta/(\sigma_d\sqrt{2})]$ (Rosner, 2006, p. 336). $\sigma_d = [\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2]^{0.5}$, where subscript 1 and 2 pertain to variances in intervention and comparison PBPMs costs, and ρ = correlation between observations between the base and intervention periods. The intervention and comparison standard deviations in the base period were $2,354 and $2,403, respectively. Assuming a .33 intra-patient correlation, σ_d = $3,075. If there were no increase in the comparison group's PBPM over time, then Δ = .05($1,563) = $78 (see Table 7-5). The treatment n = 2,584. Thus, power = $\Phi[-1.96 + (\$78 \bullet 50.8/3{,}075 \bullet 1.41 = 0.914) = 1 - \Phi[1.05] = 0.147$. With the MGH intervention sample, we had 15% likelihood of finding a significant difference if the true mean change in the intervention PBPM cost was $78 less than the change in the comparison PBPM cost.

7.3.4 Adjusting for Unbalanced Intervention and Comparison Groups

Two approaches were used to test the effects of imbalances between the intervention and comparison groups in base year characteristics. First, we produced frequency distributions of key beneficiary characteristics between the two groups. Second, we used multivariate regressions to quantify the effects of any imbalances on trends in PBPM costs. We pooled base and demonstration period observations and regressed each beneficiary's own demonstration period PBPM cost on group status (I = intervention; C = comparison); each beneficiary's own base period $PBPM_{pb}$ cost; the beneficiary's high cost or high risk group eligibility status in the base year, $Risk_{pr}$; and a vector of base period beneficiary characteristics ($\varphi Char$):

$$PBPM_{pt} = \alpha + \beta Status_p + \gamma PBPM_{pb} + \Sigma_r \rho_r Risk_{pr} + \Sigma_k \delta_k \varphi Char_{pk} + \varepsilon_{pt}. \qquad (7.7)$$

The intercept, α, is the original comparison group's average PBPM cost in the base year, while γ = each beneficiary's average dollar increase in PBPM costs over 24 months (i.e., the sixth month of the base year to the eighteenth mid-period month of the demonstration). γ provides a test of RtoM effects. The smaller is γ, the greater is RtoM. The *t*-value for β tests the differences in intervention and comparison demonstration cost growth, while ρ_r tests for the difference in the growth rates for the "*r*" cost-risk groups. By including each beneficiary's age, gender, race, urban/rural residence, disabled status, Medicaid eligibility, and institutional status at the start of the demonstration, we purge the Status and other coefficients of any systematic differences between the intervention and comparison groups that remained at the start of the demonstration. Inclusion of these variables also narrows the confidence intervals around the other coefficients, thereby reducing detectable thresholds that give more precise estimates of mean intervention effects (Greene, 2000, chapter 6).

7.4 PBPM Cost Levels and Trends

7.4.1 Original Population

Table 7-5 displays PBPM cost levels and rates of growth in average PBPM costs between the 12-month base year and the 36-month demonstration period for the original population. Results are shown for the entire intervention group and for participating and nonparticipating beneficiaries, separately. PBPM costs in both periods have been weighted by the fraction of days beneficiaries were eligible in the demonstration period so as not to overweight beneficiaries who were exposed to the intervention for shorter periods. Only beneficiaries with at least 1 day of demonstration eligibility in both periods were included.

Table 7-5
MGH's CMP PBPM cost growth rates between base year and demonstration period, intervention and comparison groups: Original population

Study group	Beneficiaries	Base year PBPM Mean[1]	Base year PBPM SE	Demo PBPM Mean[1]	Demo PBPM SE	Differences in means	SE
Intervention	2,584	$1,500	41.8	$2,022	47.9	$522**	55.2
Participants	2,280	1,507	44.8	2,024	50.3	517**	58.3
Nonparticipants	304	1,434	112.7	2,005	157.4	571**	173.7
Comparison	2,719	1,563	46.1	2,373	54.2	810**	60.5
Differences							
I – C	—	−63	62.3	−350**	72.5	−288**	82.1
Participants - C	—	−56	64.6	−349**	74.6	−293**	84.6
Nonparticipants – C	—	−129	158.8	−368**	188.9	−239	210.7
Participants - Nonparticipants	—	73	147.6	20	169.1	−53	195.2

NOTE: MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; I = intervention; C = comparison.

[1] Means weighted by beneficiary fraction of eligible days in demonstration period.

*p < .05; **p < .01.

SOURCE: Medicare Part A&B claims; run costrun1(6/27/10).

Overall. The weighted base year average PBPM cost was -$63 (4%) less ($p$=insig) in the intervention group versus the comparison group ($1,500 versus $1,563). The intervention-comparison difference in PBPM Medicare costs increased to -$350 ($p$<.01) in the demonstration period ($2,022 versus $2,373). Intervention beneficiaries, who were 4% less costly on a weighted basis at baseline, became nearly 15% less costly, on average, than the comparison group after 36 months. Between the base year and the end of the 36-month demonstration period, the average comparison group PBPM cost increased significantly by $810 ($p$<.01), while the intervention group's PBPM average Medicare costs rose more slowly by $522 ($p$<.01). Consequently, the intervention group's PBPM cost rose -$288 more slowly ($p$<.01) than the comparison group's PBPM cost.

Participation Status. The participation rate, based on beneficiaries used in this cost analysis, was 88% (2,280/2,584). Participant costs in MGH's CMP intervention group were about 4% lower (-$56; p=insig) than in the comparison group in the base period. Nonparticipants were -$129 less costly (p = insig). Participant costs rose $517 over the demonstration period compared with $810 in the comparison group, resulting in a growth difference of -$293 (p<.01). Nonparticipants became -$239 less costly (p=insig) during the demonstration period. Thus, the -$288 slower growth in overall intervention PBPM costs appears to be due in large part to slower growth in the participant group (-$293; p<.01).

7.4.2 Refresh Population

Overall. ***Table 7-6*** displays PBPM cost levels and rates of growth in average PBPM costs between the 12-month base year and the end of the 24-month demonstration period for the refresh population. The weighted base year average PBPM cost was $-182 less (*p*=insig) in the intervention versus comparison group ($1,399 versus $1,581). The intervention-comparison gap in PBPM Medicare costs widen to -$537 in the demonstration period ($1,716 versus $2,253). The average comparison group PBPM costs increased $672 (p<.01) while the intervention group's PBPM average Medicare costs increased $317 (p<.01). As a result, the intervention group's PBPM costs increased -$355 slower (p<.05) compared with the comparison group's PBPM costs. Intervention beneficiaries, who were 11.5% less costly at baseline, were 23.8% less costly than the comparison group, on average, after 18 months between the mid-points of the baseline and demonstration periods.

Table 7-6
MGH's CMP PBPM cost growth rates between base year and demonstration period, intervention and comparison groups: Refresh population

Study group	Beneficiaries	Base year PBPM Mean[1]	Base year SE	Demo PBPM Mean[1]	Demo PBPM SE	Differences in means	SE
Intervention	775	$1,399	75.3	$1,716	84.8	$317**	100.7
Participants	648	1,396	80.0	1,745	90.8	349**	105.8
Nonparticipants	127	1,423	216.9	1,541	234.9	117	302.7
Comparison	776	1,581	88.3	2,253	104.2	672**	121.8
Differences							
I – C	—	−182	115.8	−537**	134.0	-355*	157.6
Participants - C	—	−186	120.3	−508**	139.9	-323*	163.4
Nonparticipants - C	—	−158	244.9	−713**	286.3	−555	338.5
Participants - Nonparticipants	—	−28	217.8	204	245.1	232	291.2

NOTE: MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; I = intervention; C = comparison.

[1] Means weighted by beneficiary fraction of eligible days in demonstration period.

*$p < .05$; **$p < .01$.

SOURCE: Medicare Part A&B claims; run costrun1 (6/27/10).

Participation Status. The participation rate, based on beneficiaries used in the refresh population cost analysis, was 84% (648/775). Participants in the base period in MGH's CMP intervention group were -$186 less costly (p=insig) than comparison group beneficiaries and nonparticipants were -$158 less costly (p=insig). Participants became -$508 less costly (p<.01) during the demonstration period. Nonparticipants became $713 less costly (p<.01) during the demonstration period. Consequently, the participant group's PBPM cost rose -$323 more slowly (p<.05) than the comparison group's cost while the non-participant group's PBPM cost rose -$555 (p=insig) slower than the comparison group's PBPM cost.

7.5 Savings and Budget Neutrality

7.5.1 Original Population

Table 7-7 presents summary statistics on savings from the MGH CMP's original intervention population. It also includes the minimum level of savings necessary to achieve statistical significance, expressed in negative terms, and as a percentage of the comparison group's PBPM cost. The MGH CMP's monthly fee is reported also as a percentage of the comparison group's PBPM cost. Over the course of the 36-month intervention, average monthly costs increased $522 in the intervention group and $810 in the comparison group. The result was a -$288 relative decrease in PBPM cost growth in the intervention group. This negative difference implies *gross savings* at a rate of 12.1% of the comparison group's demonstration period PBPM cost. These savings were statistically significant.

With 2,854 beneficiaries in the intervention group and only 2,719 in the comparison group, the minimal detectable savings threshold was -$161 at the 95% confidence level. This threshold rate was 6.1% of the comparison group's PBPM cost, implying that the intervention would have had to achieve this level of savings to be considered statistically reliable in repeated patient samples.[13]

The MGH CMP's average monthly fee was $109 when averaged over all intervention beneficiaries, which amounted to 4.6% of the comparison group's PBPM cost during the demonstration period. It was less than $120 because it was paid only for participating intervention beneficiaries each month. Thus, MGH's CMP would have had to achieve 9.6% (4.6% + 5%) savings in order to retain all of its fees—at least according to RTI's calculations, which are not official under financial reconciliation. MGH actually saved 12.1%, exceeding the minimum required level of savings using RTI's method.

If one accepted the MGH CMP's intervention savings of -$288, then the net fee to Medicare would be -$179 instead of +$109. Medicare's rate of return on investment would be 2.65.

[13] If minimal savings were based just on differences in PBPM costs during the demonstration period, the intervention would have to achieve a 6.0% savings rate (72.5(1.96)/$2,373) based on RTI's weighting method.

Table 7-7
MGH's CMP average PBPM gross savings, fees, and budget neutrality status: Original population

Description	PBPM cost change
Intervention group	$522
Comparison group	810
Difference	−$288
Gross (dis)saving %[1]	−12.1%
Minimal Detectable Savings[2]	
Absolute	−$161
% of comparison PBPM[3]	−6.8%
Monthly Fee	
Absolute[4]	$109
% of comparison PBPM	4.6%
Net Fee	
Absolute[5]	-$179
% of comparison PBPM[3]	-7.5%
Return on Investment (RoI)[6]	2.65

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison demonstration PBPM (= $2,373). Negative values imply savings.

[2] Minimum Detectable Savings = 1.96*standard error of difference in mean PBPM changes.

[3] % Comparison PBPM = Absolute variable as % of comparison PBPM ($2,373) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of uniform $120 fee paid in outreach period and thereafter through month 36. Weights = fee-eligible members.

[5] Absolute Net Fee = Absolute Monthly Fee + Difference in PBPM outlay change.

[6] RoI = Gross savings difference/Absolute Monthly Fee.

SOURCE: Medicare 2004-2008 Part A&B claims; PBPM cost changes and detectable savings: Table 7-5; monthly fees: Fees and members taken from ARC Final Reconciliation for MGH Phase I, July 21, 2010, Tables 3, 5 and 6.

7.5.2 Refresh Population

Table 7-8 presents summary statistics on savings from the MGH intervention with the refresh population. Over the course of the 24-month intervention, average monthly costs increased $317 in the intervention group and $672 in the comparison group. The result was a -$355 lower relative increase in PBPM costs in the intervention group. This negative difference implies *gross savings* at a rate of 15.8% of the comparison group's PBPM cost.

With only 775 beneficiaries in the intervention group, the minimal detectable refresh savings threshold was -$309 at the 95% confidence level. This rate is -13.7% of the comparison group's PBPM cost, implying that the intervention would have had to achieve this level of savings to be considered statistically reliable in repeated patient samples. MGH exceeded this threshold by -$46 (-$355/-$309). The net fee to Medicare was reduced from $106 per beneficiary per month to -$249, resulting in a net Medicare cost of -11.1% of the comparison group's average monthly outlay on claims. Medicare's return on investment was 3.35, implying net savings of $2.35 on every dollar of Medicare fees paid out.

Table 7-8
MGH's CMP average PBPM gross savings, fees, and budget neutrality status: Refresh population

Description	PBPM cost change
Intervention group	$317
Comparison group	$672
Difference	−$355
Gross (dis)saving %[1]	−15.8%
Minimal Detectable Savings[2]	
Absolute	−$309
% of comparison PBPM[3]	−13.7%
Monthly Fee	
Absolute[4]	$106
% of comparison PBPM[3]	4.7%
Net Fee	
Absolute[5]	-$249
% of comparison PBPM[3]	-11.1%
Return on Investment (RoI)[6]	3.35

NOTES: MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison PBPM (= $2,253). Negative values imply true savings.

[2] Minimum Detectable Savings = 1.96*standard error of difference in mean PBPM changes.

[3] % Comparison PBPM = Absolute variable as % of comparison PBPM ($2,253) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of uniform $120 fee paid in outreach period and there after through months 13-36. Weights = fee-eligible members.

[5] Absolute Net Fee = Absolute Monthly Fee + Difference in PBPM outlay change.

[6] RoI = Gross savings difference/Absolute Monthly Fee.

SOURCE: Medicare 2004-2008 Part A&B claims; PBPM cost changes and detectable savings: Table 7-6; monthly fees: Fees and members taken from ARC Final Reconciliation for MGH Phase I, July 21, 2010, Tables 3, 5 and 6.

7.6 Imbalances between Intervention and Comparison Populations

Because the MGH CMP's comparison group was not based on random sampling, it is possible that material imbalances remained between intervention and comparison groups simply by chance. If the distribution of high cost and high risk beneficiaries differs between the MGH CMP's intervention group and its comparison group, then demonstration period PBPM cost comparisons could be biased against the intervention, if it had a disproportionate number of high risk, more cost-increasing, beneficiaries. We created four, mutually exclusive, high-low cost-risk groups. The high-cost threshold was set at $26,000/month, or the top 25% of cases in either population based on their costs the year prior to selection. The high-risk threshold was set at an HCC score of 2.75.

For differences in other beneficiary characteristics to have any effect on intervention savings, two things must happen. First, one or more characteristics must have a statistically important effect on PBPM cost growth rates. Second, unless the same important characteristics also significantly differ, numerically, between the intervention and comparison groups, they will not affect the intervention savings rates. Because most characteristics are simple binary (0, 1) indicators, there must be substantial numbers of "costly" beneficiaries involved and not just a large differences in relative frequencies.

7.6.1 Frequencies of Beneficiary Characteristics

Table 7-9 and ***7-10*** show some imbalances in the intervention and comparison groups. Because of the roughly one year lag between randomization of the original population into intervention and comparison groups and the official base year, intervention beneficiaries, compared with comparison beneficiaries, were more likely to be high cost and less likely to be just high risk. They were less likely to be disabled, minority, and eligible for Medicaid prior to the demonstration period. These differences remained in the two refresh groups.

Table 7-9
MGH's CMP frequency distribution of beneficiary characteristics, intervention and comparison groups, base year: Original population

Characteristics	Intervention (%)	Comparison (%)
COST-RISK Group		
High-cost > = $26,000	11.2%	9.5%
Both	11.8	12.5
High-risk: HCC > 2.75	9.3	11.6
Neither	67.8	66.5
Age Group		
<65	12.6	17.5
65-69	11.3	11.6
70-74	18.4	16.0
75-79	22.4	20.4
80-84	19.2	19.2
85+	16.3	15.3
Gender		
Female	52.5	52.9
Male	47.8	47.1
Race		
Minority	8.4	18.4
White	91.6	81.6
MEDICAID Eligible		
No	96.4	93.2
Yes	3.7	6.8
DISABLED		
No	87.6	82.5
Yes	12.5	17.5
Urban residence		
No	0.0	0.0
Yes	100.0	100.0
Long-term care		
No	95.6	96.8
Yes	4.4	3.2
SNF		
No	91.5	92.3
Yes	8.5	7.7

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. MGH's CMP = Massachusetts General's Hospital Care Management Program; HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/27/10).

Table 7-10
MGH's CMP frequency distribution of beneficiary characteristics, intervention and comparison groups, base year: Refresh population

Characteristics	Intervention (%)	Comparison (%)
COST-RISK Group		
High-cost > = $26,000	10.2%	8.6%
Both	9.6	13.2
High-risk: HCC > 2.75	5.3	11.1
Neither	74.9	67.2
Age Group		
<65	13.4	17.2
65-69	12.2	9.9
70-74	15.5	13.1
75-79	20.3	19.7
80-84	20.1	20.9
85+	18.5	19.3
Gender		
Female	54.9	53.2
Male	45.1	46.8
Race		
Minority	7.9	17.6
White	92.1	82.4
MEDICAID Eligible		
No	95.5	93.8
Yes	4.5	6.2
DISABLED		
No	86.4	82.3
Yes	13.6	17.7
Urban residence		
No	0.0	0.0
Yes	100.0	100.0
Long-term care		
No	97.3	97.2
Yes	2.7	2.8
SNF		
No	92.3	92.9
Yes	7.1	7.1

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period.
MGH's CMP = Massachusetts General Hospital's Care Management Program;
HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/27/10).

7.6.2 PBPM Cost Levels and Trends by Cost and Risk Group

7.6.2.1 Original Population

Table 7-11 displays PBPM costs stratified by cost and risk group. Extreme cost differences are found between the high-cost and high-risk groups in the base year. High-risk only intervention beneficiaries averaged PBPM costs of just $1,015 in the base year compared with $3,724 for high-cost only beneficiaries (3.7 times greater) and both high-cost and high-risk beneficiaries ($5,269; 5.2 times greater). Both high-cost intervention groups experienced large declines in their PBPM costs while the high-risk–only group's PBPM cost more than doubled. The comparison group showed similar patterns of cost levels and trends. Costs in the base period's neither high-cost nor high-risk group rose the fastest, which is suggestive of RtoM effects. Focusing on the difference in trends at the bottom of ***Table 7-11***, we observe, despite large negative values, only the neither high-cost, high-risk group shows statistically significant differences between the original intervention and comparison group growth rates. This is likely due to having a much larger population than the other three groups.

7.6.2.2 Refresh Population

Table 7-12 presents similar results on PBPM cost trends by the four cost-risk groups for the refresh population. Both the high-cost, high-risk and high-risk-only refresh groups showed costs rising slower in the intervention group (bottom row). Yet, due to small population sizes, statistically significant cost savings came just from the neither group (-$487; $p<.01$). The large standard errors for the refresh population are noteworthy. We had little power, except in the neither group, to detect savings rates even as large as several hundred dollars per month given the small sample sizes and high cost variance from year to year.

Table 7-11
MGH's CMP PBPM costs by cost and risk group, intervention and comparison groups, base and demonstration periods: Original population

Description	High-cost and high-risk PBPM	High-cost and high-risk SE	High-cost only PBPM	High-cost only SE	High-risk only PBPM	High-risk only SE	Neither PBPM	Neither SE
Intervention (N, %)	(368; 14%)	—	(291: 11%)	—	(256: 10%)	—	(1,669; 65%)	—
Base Year	$5,269	155.2	$3,724	105.2	$1,015	37.1	544	13.3
Demonstration	3,575	189.6	2,250	132.2	2,294	153.4	1,667	50.3
Difference	−1,694**	223.9	-1,474**	164.9	1,279**	150.2	1,133**	51.2
% Change	−32%	—	−39%	—	126%	—	208%	—
Comparison (N, %)	(414; 15%)	—	(267; 10%)	—	(327; 12%)	—	(1,711; 63%)	—
Base Year	5,660	183.6	4,112	116.8	1,045	34.1	523	13.1
Demonstration	4,327	212.2	2,696	158.4	2,543	163.0	1,931	54.9
Difference	−1,333**	252.4	-1,416**	186.6	1,498**	159.8	1,408**	55.6
% Change	−23%	—	−34%	—	143%	—	269%	—
Difference-in-Differences	−361	339.6	−58	248.3	−219	223.9	−275**	75.6

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; SE = standard error; N = number of beneficiaries; HCC = Hierarchical Condition Category.

High-Cost: Beneficiaries with annual healthcare spending greater than $26,000 in base period (top 25%).

High-Risk: HCC > 2.75 in base period.

% Change: Difference/Base Year.

*p<.05; **p<.01

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/27/10).

Table 7-12
MGH's CMP PBPM costs by cost and risk group, intervention and comparison groups, base and demonstration periods: Refresh population

Description	High-cost and high-risk PBPM	High-cost and high-risk SE	High-cost only PBPM	High-cost only SE	High-risk only PBPM	High-risk only SE	Neither PBPM	Neither SE
Intervention (N, %)	(92; 12%)	—	(83; 11%)	—	(41; 5%)	—	(559; 72%)	—
Base Year	$5,522	353.7	$3,725	199.5	$1,129	90.2	575	24.5
Demonstration	3,048	397.7	1,989	235.8	2,030	338.9	1,487	88.2
Difference	−2,474**	488.0	−1,736**	274.0	901**	337.5	912**	90.7
% Change	−45%	—	−47%	—	80%	—	159%	—
Comparison (N, %)	(124; 16%)	—	(67; 9%)	—	(86; 11%)	—	(489; 63%)	—
Base Year	5,719	321.5	4,265	262.8	1,168	64.1	498	23.6
Demonstration	3,897	320.9	2,410	361.5	2,345	324.7	1,897	116.8
Difference	−1,822**	438.3	−1,854**	406.6	1,177**	320.5	1,399**	118.3
% Change	−32%	—	−43%	—	101%	—	281%	—
Difference-in-Differences	−652	658.1	118	475.7	−276	513.8	−487**	147.1

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; SE = standard error; N = number of beneficiaries; HCC = Hierarchical Condition Category.

*p<.05; **p<.01

High-Cost: Beneficiaries with annual healthcare spending greater than $26,000 in base period (top 25%).

High-Risk: HCC > 2.75 in base period.

% Change: Difference/Base Year.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/27/10).

7.7 Regression-to-the-Mean

Tables 7-13 and ***7-14*** demonstrate that extensive RtoM is occurring in this high cost population. Changes in comparison group PBPM costs are stratified by base period cost group from low to high in $250 increments. Using comparison group data avoids any effects the intervention might have on the underlying RtoM phenomenon. Unweighted mean costs were $1,732 in the comparison group's base period in the original population (***Table 7-13***), with an overall increase of $1,479. Cost increases are inversely correlated with a beneficiary's base period PBPM costs—especially at the extremes. At the extremes, beneficiaries with less than $250 in base period PBPM costs saw their average costs increase by $2,410 while those with initial costs greater than $4,000 experienced average decreases of -$991. Mean costs in both periods are roughly double median costs and indicate a strong skewness in PBPM costs. The large number of beneficiaries with positive increases in costs during the demonstration period, including those with base period PBPM costs between $3,000 and $3,750, indicates a strong upward trend in Boston area hospital costs. This increase appears to dominate any negative expected change due to RtoM.

Regression-to-the-mean effects also are reflected in the refresh comparison population (***Table 7-14***) but, again, at the extremes. Unweighted mean costs increased $1,230 due mostly by much larger cost increases for beneficiaries with base year costs under $2,500 per month. Cost reductions of nearly -$1,400 per month were evident in beneficiaries with base costs exceeding $4,000. This suggests that for the intervention to be successful, it would need to identify initially lower cost beneficiaries most likely to experience major cost increases. Targeting initially very high cost beneficiaries would require exceptional cost reductions.

Table 7-13
MGH's CMP Regression to the Mean in comparison group PBPM costs: Original population

Base year PBPM cost level	N	Base year PBPM cost	Demonstration period PBPM cost	Change
< $250	701	$70	$2,479	$2,410
251-500	428	370	2,180	1,810
501-750	263	616	2,450	1,834
751-1,000	160	872	2,393	1,520
1,001–1,250	131	1,120	2,844	1,724
1,251-1,500	112	1,372	3,255	1,882
1,501-1,750	93	1,640	3,229	1,599
1,751-2,000	100	1,873	3,218	1,345
2,001-2,250	65	2,118	4,489	2,371
2,251-2,500	61	2,369	3,224	855
2,501-2,750	54	2,614	3,621	1,007
2,751-3,000	53	2,875	4,016	1,141
3,001-3,250	41	3,123	4,272	1,148
3,251-3,500	31	3,381	3,928	547
3,501-3,750	36	3,613	5,929	2,316
3,751-4,000	44	3,873	3,575	−298
> 4,000	346	7,087	6,096	−991
Mean	2,719	1,732	3,211	1,479
Median	—	716	1,707	991

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/27/10).

Table 7-14
MGH's CMP Regression to the Mean in comparison group PBPM costs: Refresh population

Base year PBPM cost level	N	Base year PBPM cost	Demonstration period PBPM cost	Change
< $250	204	$70	$2,831	$2,760
251-500	120	359	1,832	1,474
501-750	72	623	2,077	1,453
751-1,000	50	860	2,082	1,222
1,001–1,250	40	1,117	2,311	1,195
1,251-1,500	29	1,385	2,957	1,572
1,501-1,750	22	1,631	2,675	1,044
1,751-2,000	18	1,876	3,416	1,540
2,001-2,250	24	2,110	2,919	810
2,251-2,500	15	2,369	3,724	1,356
2,501-2,750	13	2,655	2,783	128
2,751-3,000	10	2,871	3,200	329
3,001-3,250	10	3,122	2,452	−671
3,251-3,500	5	3,418	2,886	−532
3,501-3,750	15	3,640	3,768	128
3,751-4,000	10	3,828	4,583	754
> 4,000	109	7,420	6,022	−1,398
Mean/total	766	1,827	3,057	1,230
Median/total	—	718	1,483	765

NOTES: Observations unweighted. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/27/10).

7.8 Multivariate Regression Tests of Intervention Savings

7.8.1 Original Population

Two sets of regression coefficients in ***Table 7-15*** test the intervention effect by using the beneficiary's base year PBPM cost (PBPM_base) to explain each beneficiary's demonstration period PBPM cost. The first column shows the gross effect of the intervention controlling only for each beneficiary's own base PBPM cost. Coefficients in both columns can be interpreted as differences between each beneficiary's average demonstration and base year PBPM costs.

In the first column of results controlling only for each beneficiary's base period PBPM cost, the Intervention coefficient of -331 is statistically significant at the 1% confidence level implying reliable success in slowing beneficiary cost increases. The base period PBPM cost coefficient (0.309; $p < .01$), when combined with the intercept coefficient, implies substantial RtoM effects on costs (0.309 - 1 = -0.691, the RtoM effect). Imagine two comparison group beneficiaries, one with a relative low ($500) and another with a relatively high ($4,000) PBPM cost in the base period. The predicted PBPM cost of the initially "low cost" comparison beneficiary would increase over 3-fold during the intervention period, while the "high cost" beneficiary's PBPM cost would decline by 30%.[14] Whereas example cost differences were 8:1 in the base period, they would now be compressed to 1.6:1.

RtoM effects are quite substantial but not in one direction. Including only high cost beneficiaries in the intervention group would clearly have produced even greater declines in comparison group PBPM costs during the demonstration. Major cost increases did occur among initially lower cost beneficiaries, as evidenced in ***Table 7-13***.

The second regression model controls for which cost-risk group the beneficiary was in during the base period as well as several other beneficiary characteristics. The Intervention coefficient is reduced to -$289 (13% less than -$331) when applying the controls yet remains statistically significant. This reduction is due to imbalances that remained between the intervention and comparison groups (e.g., higher extreme base costs, less disabled, minority, or Medicaid enrollee. The PBPM base coefficient declines to 0.223, implying more RtoM within each of the cost-risk groups.

[14] The calculation is as follows based on Table 7-15, column 1:

PBPM[base]	PBPM[demo]	PBPM Change	%Change
$500	$1,713	$1,213	+243%
$4,000	$2,795	-$1,205	-30%

Table 7-15
MGH's CMP regression results: Intervention gross savings controlling for base period PBPM cost and beneficiary characteristics: Original population

Independent Variable	PBPM_ Demo Coefficient	PBPM_ Demo *t*-stat	PBPM_ Demo Coefficient	*t*-stat
Intercept	1,559**	28.5	1,574	1.3
Intervention	−331**	4.7	−289**	4.1
PBPM_Base	0.309**	20.1	0.223**	8.9
High-cost–high risk			1,106**	6.7
High-cost			−34	0.2
High-risk			511**	4.4
Male			−56	0.8
Minority			142	1.3
Age 65-69			−1	0.0
70-74			−213	−0.2
75-79			−118	−0.1
80-84			37	0.3
85+			118	0.1
Medicaid			282	1.8
Disabled			−47	0.0
SNFB			−122	0.9
R^2	.075		.091	
N	5,302		5,302	

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

*p<.05; **p<.01.

Observations weighted by beneficiary's fraction of eligible days during demonstration.

PBPM_Demo: Dependent variable: Beneficiary's average PBPM during demonstration.

PBPM_Base: Beneficiary's average PBPM in base period just prior to start date.

High-Cost-High Risk: PBPM > $26,000 and HCC > 2.75 in base year.

High-Cost: PBPM > $26,000 and HCC< 2.75.

High-Risk: PBPM < $26,000 and HCC > 2.75.

SNFB = 1 if beneficiary had SNF payments in base year.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/27/10); final/cost5 (5/18/10).

7.8.2 Refresh Population

In the first column of refresh results in ***Table 7-16,*** controlling only for each beneficiary's base period PBPM cost, the Intervention coefficient of -493 is significant at the 1% confidence level, implying highly reliable statistical cost trend savings in the intervention group. The base period PBPM cost coefficient (0.243, $p < .01$), when combined with the intercept coefficient, again implies substantial RtoM of costs in the refresh population (= 0.243 - 1 = -0.757, the RtoM effect).

The second regression model controls for cost-risk group and other patient characteristics determined during the base period. Again, the Intervention coefficient declines somewhat (-439) after controlling for population group imbalances but still remains highly significant.

With the one exception of the high-cost, high-risk group that shows higher cost growth, most of the remaining control variables were statistically insignificant. This is due mainly to the overall balance between the refresh intervention and comparison groups. The high-cost, high-risk group can exhibit higher absolute cost increases while still being negatively affected by RtoM effects because the regression intercept ($2,020) is well below mean costs of the group.

Table 7-16
MGH's CMP regression results: Intervention gross savings controlling for base period PBPM cost and beneficiary characteristics: Refresh population

Independent variable	PBPM_ Demo Coefficient	PBPM_ Demo *t*-stat	PBPM_ Demo Coefficient	PBPM_ Demo *t*-stat
Intercept	1,377**	13.8	2,020	1.7
Intervention	−493**	3.8	−439**	3.3
PBPM_Base	0.243**	8.4	0.147**	3.0
High-Cost-High Risk			1,062**	3.4
High-Cost			−35	0.1
High-Risk			477	1.9
Male			−214	1.6
Minority			194	1.0
Age 65-69			−603	0.5
70-74			−494	0.4
75-79			−641	0.5
80-84			−700	0.6
85+			−172	0.1
Medicaid			236	0.8
Disabled			−888	0.7
SNFB			162	0.6
R^2	.054		.074	
N	1,540		1,540	

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. MGH's CMP = Massachusetts General Hospital's Care Management Program; PBPM = per beneficiary per month; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

*p<.05; **p<.01.

Observations weighted by beneficiary's fraction of eligible days during demonstration.

PBPM_Demo: Dependent variable: Beneficiary's average PBPM during demonstration.

PBPM_Base: Beneficiary's average PBPM in base period just prior to start date.

High-Cost-High Risk: PBPM > $26,000 and HCC > 2.75 in base year.

High-Cost: PBPM > $26,000 and HCC< 2.75.

High-Risk: PBPM < $26,000 and HCC > 2.75.

SNFB = 1 if beneficiary had SNF payments in base year.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1(6/27/10).

7.9 Conclusion

PBPM costs showed considerable variability because of the nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. Cost trends in MGH's market area also showed strong positive growth. Nevertheless, substantial, statistically significant, savings were found for the intervention in the original population. Costs rose -$288 slower in the original intervention group (12.1% of comparison costs) while gross savings needed to exceed just 6.8% to be considered statistically significant. Medicare's return on investment was 2.65. For every dollar invested in MGH's CMP management fees, Medicare received $2.65 in savings on beneficiary health care services.

If anything, MGH's CMP performed even better with its refresh population. Gross savings averaged -$355 (15.8% of comparison monthly costs). Based on an average monthly management fee of $120 paid on 84% of participating intervention eligible beneficiaries, Medicare's return on investment was 3.35. For every management dollar spent, Medicare received $3.35 in return in the form of lower cost increases.

A few material imbalances were found between intervention and comparison groups in the base period. However, controlling for imbalances had little effect on our overall final conclusion of statistically significant savings.

The CMHCB demonstration program at MGH exhibited strong regression to the mean effects while average costs also were increasing rapidly in the market area. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. This churning occurred around a strong local upward trend in costs. Costs continue to rise because the reduction in costs in the very high cost group is more than offset by smaller increases among the greater majority of beneficiaries. Increases in demonstration period costs in initially less costly beneficiaries make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater is the potential for regression-to-the-mean, the greater is the challenge to identify lower cost, lower utilizing beneficiaries initially to avoid expensive hospitalizations in the near future. Nevertheless, it appears that MGH's CMP staff was able to work successfully across a broad cost range of their patients, resulting in a financially successful outcome.

CHAPTER 8
KEY FINDINGS FROM THE MASSACHUSETTS GENERAL HOSPITAL AND THE MASSACHUSETTS GENERAL PHYSICIANS ORGANIZATION CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION EVALUATION

The purpose of this report is to present the findings from RTI International's evaluation of the Massachusetts General Hospital and the Massachusetts General Physicians Organization (MGH) Care Management Program (CMP). Our evaluation focuses upon three broad domains of inquiry:

Implementation. To what extent was MGH able to implement its program?

Reach. How well did MGH's CMP engage its intended audience?

Effectiveness. To what degree was MGH's CMP able to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings?

Organizing the evaluation into these areas focuses our work on the policy needs of the Centers for Medicare & Medicaid Services (CMS) as it considers the future of population-based care management programs or other interventions in Medicare structured as pay-for-performance initiatives. We use both qualitative and quantitative research methods to address a comprehensive set of research questions within these three broad domains of inquiry.

8.1 Key Findings

In this section, we present key findings based upon the 36 months of MGH's CMP operations with its original population and 24 months with its refresh population. Our findings are based on the experience of approximately 6,800 ill Medicare beneficiaries split across 4 groups for analysis purposes (original and refresh intervention and comparison groups) limiting statistical power somewhat within the substantially smaller refresh population (only 30% the size of the original population) to detect differences. Eight findings on participation, intensity of engagement in MGH's program, beneficiary satisfaction and experience with care, provider satisfaction, clinical quality, acute care utilization, health outcomes, and financial outcomes have important policy implications for CMS and future disease management or care coordination efforts among Medicare fee-for-service (FFS) beneficiaries. The CMHCB demonstration program holds MGH financially responsible for financial savings but does not hold MGH financially responsible for quality of care improvements.

Key Finding #1: MGH's CMP achieved a high participation level that reached broadly across its intervention population in terms of beneficiary demographic characteristics, prior health status and health care costs, and health status measured during the early months of its demonstration.

The MGH CMP had an ambitious goal of gaining participation from 100% of its original population beneficiaries. It was successful in recruiting 88% of its original population beneficiaries and 84% of its refresh population beneficiaries. We found few statistically

significant differences between participants and nonparticipants in either the original or refresh intervention populations. In multivariate modeling of factors that predict likelihood of participation, we had low explanatory power suggesting that MGH's CMP was able to recruit broadly across its intervention population as no particular set of factors that we tested strongly predicted participation. The substantially smaller sample size for the refresh population also limited our ability to detect participation factors.

Key Finding #2: MGH's CMP successfully targeted beneficiaries with high rates of acute care utilization.

A cornerstone of MGH's CMP was one-on-one relationship between participants and their practice-based case managers. Telephone contact was the most dominant form of contact. In our multivariate regression modeling of likelihood of being in a high contact versus low contact group for both the original and refresh populations, we found hospitalizations during the demonstration period to be very strong predictors of contact. A major focus of MGH's CMP was to prevent hospitalizations or re-hospitalizations. These findings reveals that MGH's CMP was successful in their effort to contact beneficiaries who had been hospitalized and at high risk of re-hospitalization albeit MGH was unsuccessful at reducing rates of readmission.

Key Finding #3: MGH's CMP improved beneficiary reported satisfaction with helpfulness of discussions and communication with their health care team. MGH's CMP also improved physical functioning. MGH's CMP did not improve beneficiary reported ability to cope with their chronic condition nor improve self-efficacy or self-care activities or mental health functioning.

The beneficiary survey was designed to obtain assessments directly from beneficiaries about key outcomes of beneficiary experience of care, self-management, and physical and mental function. We asked beneficiaries about the extent to which their health care providers helped them to cope with their chronic condition. We supplemented this item with questions related to two key components of MGH's CMP intervention: helpfulness of discussions and quality of communication with their health care team. In addition, the survey instrument collected information about beneficiary self-care frequency and self-efficacy related to medications, diet, and exercise and Clinician and Group Adult Primary Care Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

Among the 19 outcomes covered by the survey, we found three statistically significant ANCOVA-adjusted intervention effects, two in the experience and satisfaction with care domain and one in the physical and mental health function domain. Survey results indicate that members of MGH's CMP intervention group were more satisfied with the discussion of their treatment choices and rated their communication with heath providers higher than beneficiaries in the comparison group. These achievements, however, were not translated into any improvements in self-efficacy or in self-care activities. In addition, MGH's CMP beneficiaries in the intervention group reported significantly higher Veterans RAND-12 Physical Health Composite (PHC) scores than those in the comparison group, suggesting that intervention improved physical functioning of participating Medicare beneficiaries. No other statistically significant outcomes were found in the physical and mental health function domain.

Key Finding #4: MGH's CMP improved primary care provider (PCP) assessment of the quality of medical practice and quality of care for their patients.

In addition to improving the quality of care and outcomes for Medicare beneficiaries, MGH's CMP aimed to improve the quality of work life of primary care physicians and ultimately attract more physicians to the field of primary care. It is one of several initiatives in development at MGH to improve the challenging work life of primary care physicians. Ultimately, these initiatives are part of a larger vision for Partners HealthCare to restructure the practice model for primary care practice characterized by high patient and physician satisfaction, work flow and process improvement, and the delivery of evidence-based care.

RTI conducted two site visits to MGH's CMP and spoke with a small number of primary care physicians during each site visit to gauge their assessment of satisfaction with the demonstration program. At the time of the first site visit, a small number of physicians expressed concerns about the program. For example, they had questions about whether CMP patients would divert services from other patients in their practices. And, some physicians did not have a full understanding of the role of the case managers. However, as physicians gained experience working with the case managers, the most common concern they voiced was frustration about their inability to include additional patients in the program. One provider noted that for each patient eligible for the program, there are two additional patients in the practice who could benefit from such case management support.

At the time of the second site visit, physicians gathered for the focus group reported great overall satisfaction with the CMP. The following first three quotes highlight the essence of their satisfaction with MGH's CMP with the fourth quote expressing a widely held view among the interviewed physicians:

- "The program 'wraps its arms' around the most difficult and complex patients."
- "The program signifies a move towards a true medical home model-it is a team of providers. The program does what every PCP needs to be doing but cannot do anymore because of the medicine practice and reimbursement realities and primary care provider shortages."
- "The program has done a remarkable job in training and cultivating case managers who are very good at breaking barriers and making it work for the most difficult patients."
- "We do not want the program to end—it is very valuable! Once the program is gone, participants will become 'frequent flyers' in the emergency department and hospital."

Key Finding #5: MGH's CMP did not improve rate of compliance in four quality-of-care process measures.

We have defined quality improvement for this evaluation as an increase in the rate of receipt of claims-derived, evidence-based quality-of-care measures. We selected three measures appropriate for different populations of elderly beneficiaries: influenza vaccine for all beneficiaries; low-density lipoprotein cholesterol (LDL-C) testing for beneficiaries with diabetes

or ischemic vascular disease (IVD); and rate of annual HbA1c testing for beneficiaries with diabetes. During the demonstration, we find no evidence of systematic improvement in quality of care among the intervention beneficiaries.

For the original and refresh populations and within both the intervention and comparison groups, there were high rates of "always being compliant" in receipt of three of the process-of-care measures with the noted exception of influenza vaccination. Thus, there was limited room for improvement in either population. During the last year of the demonstration, only 10% of intervention beneficiaries with diabetes were not compliant in receipt of annual HbA1c testing, and 20% and 24% of intervention beneficiaries with diabetes or IVD, respectively, were not compliant in receipt of annual LDL-C testing. For influenza vaccination, the original and refresh intervention groups' rate increased during the demonstration but increased less than the rates for the comparison groups. However, baseline rates for the original and refresh intervention groups were far higher than the comparison groups' rate.

This finding is not unexpected. MGH's CMP leadership felt that there was a very good reason standard quality measures are not part of this demonstration's outcomes. Program leaders reported that such measures are not good quality of care indicators for the program's population. For their group of patients, something like testing for HbA_{1c} levels is not a relevant measure of how well the program is managing the care of their very sick and complex patients. The CMP leadership and MGH leadership believe that ER use and acute hospitalizations are in essence the measures that need to be used. In addition to these outcomes, other types of measures related to care coordination that they believe are highly relevant to this population include how fast case managers follow up on patient-initiated calls, can appointments be consolidated so frail beneficiaries do not have to drive to the hospital 3 times a week, and so forth.

Key Finding #6: MGH's CMP was successful reducing the rate of increase in acute care hospitalizations and ER visits but not 90-day readmissions. MGH's CMP did not impact use of the Medicare hospice benefit.

During the course of MGH's CMP, we generally observe increasing rates of all-cause and ambulatory care sensitive condition (ACSC) hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations. However, MGH's CMP was successful at substantially reducing the rate of increase in all-cause and ACSC hospitalizations and ER visits among the original and refresh intervention beneficiaries. We observe no statistically significant difference in the rate of readmission between the intervention and comparison original and refresh populations. During the last 12 months of the demonstration, rates of growth in acute care utilization narrowed between the original intervention and comparison beneficiaries; yet the lower rates of growth among original intervention beneficiaries remain statistically significant.

One component of MGH's CMP was end-of-life planning including advance directives and use of hospice. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their use of the Medicare hospice benefit or in mean or median number of days of hospice.

Key Finding #7: MGH's CMP was successful at reducing the mortality rate within the intervention group of beneficiaries.

Another key outcome metric is mortality. Over the 36-month demonstration period for the original population, 28% of beneficiaries in the intervention group died and 30% of beneficiaries in the comparison group died (p=0.19). MGH CMP's mortality rate was statistically lower than in the comparison group after adjusting for differences in baseline characteristics using Cox proportional hazard modeling (p<0.05). For the refresh population, we do observe a lower rate of mortality in the intervention group. During the 24-month refresh period, 16% of beneficiaries in the intervention group died while 20% of beneficiaries in the comparison group died (p=0.04). Controlling for baseline characteristics in the multivariate modeling narrows the statistical significance of this difference, yet the mortality difference remains statistically significant at the p=0.05 level for the refresh intervention group.

Key Finding #8: MGH's CMP achieved substantial, statistically significant savings. The Medicare program's return on investment (ROI) was 2.65 for MGH's original intervention group and 3.35 for MGH's refresh intervention group.

PBPM costs showed considerable variability because of the nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. Cost trends in MGH's market area also showed strong positive growth in both groups. Nevertheless, substantial, statistically significant, savings were found for the intervention in the original population. Relative costs (or gross savings) rose -$288 slower in the original intervention group (12.1% of monthly comparison costs); yet needed to exceed just 6.8% to be considered statistically significant at the 95% confidence level. For every dollar invested in MGH's CMP management fees, Medicare received $2.65 in savings on beneficiary health care services.

If anything, MGH's CMP performed even better with its refresh population. Gross savings averaged -$355 (15.8% of comparison monthly costs). Based on an average monthly management fee of $120 paid on 84% of participating intervention eligible beneficiaries-, Medicare's return on investment was 3.35. For every management dollar spent, Medicare received $3.35 in return in the form of lower cost increases.

A few material imbalances were found between the intervention and comparison groups in the base period. However, controlling for imbalances had little effect on our overall final conclusion of statistically significant savings.

The CMHCB demonstration program at MGH exhibited strong regression-to-the-mean effects while average beneficiary costs also were increasing rapidly in the greater Boston area. Intervention group costs continued to rise because minor reductions in costs in the very high cost group were more than offset by larger increases among the greater majority of beneficiaries. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. Regression-to-the-mean effects make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater is the potential for regression-to-the-mean, the greater is the challenge to identify lower cost, lower utilizing beneficiaries initially to avoid expensive hospitalizations in the near future.

Nevertheless, it appears that MGH's CMP staff was able to work successfully across the full spectrum of low-to-high cost beneficiaries, resulting in a financially successful outcome.

8.2 Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that MGH's CMP had success at improving primary care providers' satisfaction with their quality of work life and improving some measures of beneficiary experience with care and functional status. We also find that MGH's CMP had substantial success reducing acute care hospitalizations and ER visits and mortality, and achieving substantial cost savings. We find these latter successes within both the original and refresh intervention groups. The financial savings is particularly noteworthy given the relatively small sample sizes and regression to the mean effects. PBPM costs showed considerable variability because of the nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. With only roughly 2,600 original and 800 refresh intervention beneficiaries and 2,700 original and 800 refresh comparison beneficiaries, we had limited our power to detect significant savings in the refresh population in particular. Gross savings had to be at least 6.8% in the original intervention population and 13.7% or more in the refresh intervention population to be considered significant at the 95% confidence level.

What might explain the observed success in MGH's demonstration program? Two explanations may be the depth of institutional support to (1) develop an MGH-specific program, and (2) to fully integrate the CMP into MGH's health care system. Based upon interviews with senior MGH and CMP leadership, it was noted that from the beginning the CMP had the complete backing from the Board of Trustees and hospital and physician leadership. As one example, MGH physicians received communications about the program directly from the hospital's and the physician organization's leadership. And, MGH invested considerable time and resources in their CMHCB program development before launching the program in Eastern Massachusetts. Pre-launch activities included conducting a pilot study of practice-based care management (PBCM), conducting focus groups to inform the design of the Care Management Program, and hiring and training staff to implement the program.

Approximately 2.5 years prior to the initiation of the launch of the CMP, MGH began the planning for this type of demonstration. As one example, MGH conducted a pilot study of the impact of PBCM at the MGH Revere Healthcare Center. An experienced case manager was placed within the health center to provide care management support services to patients. Physicians at the practice were asked to identify patients who were most likely to be admitted to the hospital within the next 6 to 8 weeks, and those individuals were invited to participate in the PBCM program, regardless of their insurance coverage. The case manager conducted assessments of participants to identify gaps in care and served as a physician extender helping patients deal with issues such as transportation to the physician's office and access to prescription medications. The PBCM pilot required physicians to spend time initially to discuss the organization and content of the case management assessments and services. An evaluation of the program showed that physicians were very satisfied, referring to the case manager as a "fairy godmother."

MGH also convened a series of focus groups, referred to as capstone groups, to obtain input from physicians and other MGH clinical staff about their priorities for PBCM so MGH could include useful interventions in its CMP. Initially, MGH conducted multidisciplinary capstone group sessions with representatives from social work, mental health, and the MGH case management department in addition to leadership from primary care practices. A second round of focus groups was conducted with physician groups to specifically discuss how the CMP could add value to their practices.

In addition to providing input about the design of the CMP, the capstone groups provided an opportunity to obtain physician buy-in to the PBCM program. Despite the fact that some physician practices already had case managers, CMP management observed that most physician practices were apprehensive about changes such as the introduction of new staff into their practice. CMP leadership used a tailored approach to discuss the project with each practice, offering positive anecdotes from the PBCM pilot project as appropriate. In addition, CMP leadership identified a physician champion for the CMP within each physician practice that had at least 25 or more CMP patients at the start if the project to further ease the transitions involved in the introduction of a case manager into the practice. During program implementation physician champions provided insight about the best way to incorporate case managers into the practice and encourage colleagues to take advantage of services available from the case managers.

At the time of the program launch, strong integration support from MGH leadership afforded the case managers physical entry into the primary care practice settings whereby the case managers were co-located with the primary care physicians ultimately becoming a part of the beneficiaries' primary health care teams. At the time of the first site visit, a small number of physicians expressed concerns about the program. However, as physicians gained experience working with the case managers, the most common concern they voiced was frustration about their inability to include additional patients in the program. At the time of the second site visit, physicians gathered for the focus group reported great overall satisfaction with the CMP. Acquiring buy-in from participating physician practices was viewed as very important. However, it was recognized early on that buy-in was needed on all levels. There was some concern among practice-based nurses, particularly at smaller practices, that there would be a duplication of effort. To obtain buy-in from the nurses, the CMP case managers spent time working with the practice-based nurses to educate them that the goal of the program was to augment and not to replicate their efforts.

With leadership support for integration within the MGH health system, the CMP was able to marshal a wide range of MGH internal resources to more fully develop particular aspects of their program that were tailored to the needs of the MGH patient population. Specific examples included the development of a CMP-specific mental health team comprised of MGH psychiatrists and a CMP social workers to screen for and treat depression among its participants; development of a shared planning protocol with MGH discharge planning case managers; and enlistment of an MGH pharmacist to review the appropriateness of medication regimens and assist patients with access to medications.

Another critical element of integration was the use of MGH's information technology (IT) system to support CMP operations. By gaining access to MGH's existing IT system and

MGH internal resources to make necessary modifications during early stages of implementation, the CMP was able to draw upon existing infrastructure and augment it to provide immediate decision management support for its case managers.

Further, MGH's IT systems span all care settings at MGH, including all MGH physician practice settings. And, according to CMP leadership, MGH patients are very loyal to MGH and receive the vast majority of their health care from the large network of MGH-affiliated providers. Thus, CMP case managers had access to real-time patient information across virtually their patients' entire continuum of care. Yet, the CMP went through several iterations of data system enhancements at considerable expense as it sought to increase usefulness of its systems for managing patient care and reducing documentation burden.

During our site visits, CMP leadership opined that creating a similar program may require a large setting like a teaching hospital where the information technology component and the related underlying infrastructure are in place before program implementation. CMP leadership felt that a care management program such as theirs may not work well in individual practices because of resource constraints.

As one specific example of the value of integration of the CMP with the MGH electronic medical record (EMR), MGH's EMR was modified to include a CMP icon to alert providers that a particular patient was a participant in the CMP. Because of the leadership effort to make the program both visible and integrated, the CMP icon generated the type of response from MGH providers that has eluded other Medicare chronic care management programs that we have evaluated. According to CMP leadership, the day the CMP launched the MGH ER notified the CMP Project Manager that dozens of Medicare beneficiaries with CMP icons were in their emergency room. A key focus of the MGH CMP and many other chronic care management programs is to prevent acute care hospitalizations yet many other programs that we have evaluated have been unable establish real-time notification systems with emergency rooms allowing case managers an opportunity to intervene prior to admission.

CMP integration with MGH's EMR also allowed case managers to receive weekly reports that showed which patients were scheduled for a physician visit so they could meet with patients to conduct patient interventions such as providing educational information. Case managers could also use this information to identify patients who missed their appointments so they could follow up and determine if the beneficiary needed additional support. Case managers also received timely notifications of patient admission to the emergency department or to an inpatient unit so that they could assist with transitions out of acute care and help patients avoid future exacerbations that would require acute care.

A third explanation may be elements of the management of the CMP itself. CMP leadership recognized prior to launch of its program that its population would require case managers with substantial experience in dealing with frail and medically complex patients. The CMP selected nurses with strong clinical skills, critical thinking abilities, and the ability to work independently. In addition, CMP leadership felt strongly that to be successful case managers had to learn quickly how to fit into their assigned practice setting in a way that would be helpful and valued by the physicians and their clinical and administrative staff members.

The CMP leadership organized a comprehensive orientation program to introduce the case managers to MGH, including patient resources available across the continuum of care at the hospital. Case managers met with various MGH staff members and spent time observing preceptor case managers to learn how to perform their jobs. Case managers also received training to conduct patient assessments, create comprehensive care plans, arrange for referrals to various services (e.g., transportation), and use the information systems available to support the CMP.

Due to the complexity of the CMP demonstration population, CMP leadership felt that constant and good communication between all staff within the program was essential. The CMP leadership implemented *Virtual Rounds,* regular e-mail reports that went to all staff, as a mechanism of providing feedback on a weekly basis. Case managers used *Virtual Rounds* to report on difficult patients and unnecessary admissions, and to describe both positive and negative events. *Virtual Rounds* were also used for case reviews with forms that staff filled out at the end of the week. These case reviews were then discussed with physicians in weekly face-to-face meetings. Common themes and issues from the *Virtual Rounds* were also presented at bi-monthly management meetings. The bi-monthly management meetings were used to review protocols, present resources, provide training, and identify issues and brainstorm solutions.

CMP leadership also emphasized team support and peer counseling by developing infrastructure that provided opportunities for mutual support among CMP case managers and peer counseling from the members of the mental health team as the emotional toll on staff of working with a highly frail and sick population are substantial. Such challenges include high mortality rates among program enrollees and challenges making a meaningful clinical or social impact in very advanced cases. In focus groups conducted during both of our site visits, case managers expressed strong support for CMP leadership and a strong sense of job satisfaction.

In evaluations of other Medicare chronic care management programs, we have observed other programs that exhibited *strong program leadership* and a strong sense of job satisfaction among the case managers, yet we have not observed the same degree of *integration* of the care management program into an integrated health system and its *IT system* as we do with MGH's CMP. And, MGH's CMP beneficiaries were sufficiently concentrated in the primary care practices making placement of *full-time case managers, in general, in the practices* economically feasible. It may be that all four elements are necessary to be successful reducing acute care utilization and the cost of care for chronically ill Medicare beneficiaries.

Yet, even with the level of observed practice and IT integration MGH's CMP was not successful reducing 90-day readmissions. This is surprising given the emphasis of MGH's CMP on managing care transitions across settings. In the CMP, case managers followed a protocol for every transition between care settings. The protocols include step by step daily workflow instructions for the following transitions of care: emergency department admissions, inpatient MGH admissions, post hospital discharges to home from MGH, post hospital discharges to other facilities, and post discharge from post acute facility to home. In addition, a post-episode assessment was completed within 24-72 hours of the patient's discharge from the MGH emergency department or inpatient unit, and other acute or post-acute care facilities, if known. In addition, the patient's primary care-based case manager interfaced with the MGH inpatient case manager during the admission and prior to discharge.

This level of effort to prevent readmissions and the disappointing results suggests that broadly reducing readmissions among chronically ill Medicare FFS beneficiaries may be a far bigger challenge than has been envisioned by MGH's CMP leadership or Federal policy makers. In Phase II, MGH's CMP has been granted a demonstration waiver to allow for direct admission to a skilled nursing facility (SNF) beneficiaries who meet specific clinical criteria. It will be important to examine if the SNF waiver is a tool that can be used to reduce readmissions of patients who become clinically unstable after discharge. In Phase II, it will also be important to explore in greater detail the reasons for readmission and the degree to which the readmission is clinically related to the prior admission. Our analyses showed that ambulatory care sensitive conditions (ACSCs) account for about one-quarter of reasons for admission. It is plausible that there is a sizeable subset of beneficiaries being admitted in short spans of time for acute care services not readily amenable to reduction through case management of care transitions (e.g., hip replacement, cataract extraction, etc.).

REFERENCES

Anderson, G.: Chronic Conditions: Making the Case for Ongoing Care. Partnership for Solutions, Johns Hopkins University and the Robert Wood Johnson Foundation. Baltimore. Johns Hopkins University, Dec. 2002.

Brody, E. and McCall, N.: Evaluation of the Medicare Care Management for High Cost Beneficiaries Demonstration: Site Visit Report on RMS. Prepared for Centers for Medicare & Medicaid Services, Nov. 2006.

Centers for Disease Control and Prevention Diabetes Cost-Effectiveness Group: Cost-effectiveness of intensive glycemic control, intensified hypertension control, and serum cholesterol level reduction for type 2 diabetes. JAMA 287:2542-2551, 2002.

Centers for Medicare & Medicaid Services (CMS): Medicare to award contracts for demonstration projects to improve care for beneficiaries with high medical costs. Press release from CMS Office of Public Affairs, Jul. 1, 2005. Available at http://www.cms.hhs.gov/apps/media/press_releases.asp

Coleman, E.A., Parry, C., Chalmers, S., and Min, S.: The care transitions intervention. Results of a randomized controlled trial. Arch. Int. Med. 166:1822–1828, 2006.

Congressional Budget Office (CBO): An Analysis of the Literature on Disease Management Programs. Washington. U.S. Congress, Oct. 2004.

Congressional Budget Office (CBO): High-Cost Medicare Beneficiaries. Washington. U.S. Congress, May 2005.

Glasgow, R.E., Orleans, C.T., and Wagner, E.H.: Does the chronic care model serve also as a template for improving prevention? Milbank Q. 79(4):579-612, 2001.

Greene, W. (2003). *Econometric Analysis*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hays, R.D.: RAND-36 Health Status Inventory. San Antonio, Tex. The Psychological Corporation, 1998.

Hibbard, J.H., Mahoney, E.R., Stock, R., and Tusler, M.: Do increases in patient activation result in improved self-management behaviors? Health Serv. Res. 42(4):1443-1463, 2007.

Jencks, S.F., Huff, E., and Cuerdon, T.: Change in the quality of care delivered to Medicare beneficiaries. JAMA 289(3):305-312, 2003.

Jenkins, C.R., and Dillman, D.A.: Towards a theory of self-administered questionnaire design. In L. Lyberg, P. Biemer, M. Collins, et al. (eds.), Survey Measurement and Process Quality, 165-196. New York. Wiley, 1997.

Kazis, L.F., Lee, A., Spiro, A., et al.: Measurement comparisons of the Medical Outcomes Study and Veterans SF-36 Health Survey. Health Care Financ. Rev. 25:43-58, 2004.

Kroenke, K., Spitzer, R.L., and Williams, J.B.W.: The Patient Health Questionnaire-2: Validity of a two-item depression screener. Med. Care 41:1284-1292, 2003.

Lorig, K.: Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: A randomized trial. Med. Care 37(1):5-14, 1999.

McCall, N., Urato, C, Spain, P., Smith, K., and Bernard, S: Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Second Annual Report. Prepared for Centers for Medicare & Medicaid Services, Feb. 2008.

McGlynn, E.A., Asch, S.M., Adams, J., et al.: The quality of health care delivered to adults in the United States. N. Engl. J. Med. 348(26):2635-2645, 2003.

Medicare Payment Advisory Commission (MedPAC): Report to the Congress: New Approaches in Medicare. Washington. MedPAC, Jun. 2004.

Naylor, M.D., Brooten, D., Campbell, R., Jacobsen, B.S., Mezey, M.D, Pauly, M.V., and Schwartz, J.S.: Comprehensive discharge planning and home follow-up of hospitalized elders: A randomized clinical trial. JAMA 281:613-620, 1999.

Norris, S.L., Nichols, P.J., Caspersen, C.J., et al.: The effectiveness of disease and case management for people with diabetes: A systematic review. Am. J. Prev. Med. 22(4S):15-38, 2002.

Parchman, M.L., Noel, P.H., & Lee, S.: Primary care attributes, health care system hassles, and chronic illness. Med. Care 43(11):1123-1129, 2005.

Peikes, D., Brown, R., Chen, A., and Schore, J.: Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. JAMA 301(6):603-618, 2009.

Piantadosi, S.: Clinical Trials: A Methodologic Perspective. New York. John Wiley and Sons, 1997.

Plocher, D., and Wilson, W.: Buying an outcomes management system. Healthc. Inform., 19(10:71-74, Oct. 2002.

Rich, M., Beckman, V., Wittenberg, C., Leven, C., Freedland, K., and Carney, R.: A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure. N. Engl. J. Med. 333, 1190-1195, 1995.

Rosner, B. (2006). *Fundamentals of Biostatistics* (6th ed.). Belmont, CA: Thomson: Brooks/Cole.

Schore, J., Brown, R., and Cheh, V.: Case management for high-cost Medicare beneficiaries. Health Care Financ. Rev. 20(4):87-101, Sum. 1999.

Smith, K.W., Bernard, S.L., Khatutsky, G., and McCall, N.: Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Results of the Medicare Health Services Survey. Prepared for Centers for Medicare & Medicaid Services, Oct. 2008.

Spain, P. and McCall, N: Evaluation of the Medicare Care Management for High Cost Beneficiaries Demonstration: Follow-up Site Visit Report on VillageHealth's Key to Better Health (KTBH). Prepared for Centers for Medicare & Medicaid Services, Aug. 2008

Todd, W., and Nash, T.: Disease Management, a Systems Approach to Improving Patient Outcomes. New York. Jossey-Bass Publishers, 2001.

Toobert, D.J., Hampson, S.E., and Glasgow, R.E.: The summary of diabetes self-care activities measure: Results from 7 studies and a revised scale. Diabetes Care 23(7):943-950, 2000.

Van Der Ven, N.C., Weinger, K., Yi, J., Pouwer, F., Ader, H., Van Der Ploeg, H.M., and Snoek, F.J.: The confidence in diabetes self-care scale: Psychometric properties of a new measure of diabetes-specific self-efficacy in Dutch and US patients with type 1 diabetes. Diabetes Care 26(3):713-718, 2003.

Vladek, B.: You can't get there from here: Obstacles to improving care of the chronically ill. Health Aff. (Millwood) 20(6):175-179, 2001.

Wagner, E.: The changing face of chronic disease care. Accelerating Change Today for America's Health. Curing the System: Stories of Change in Chronic Illness Care. Washington. The National Coalition on Health Care and The Institute for Healthcare Improvement, 2002.

Wagner, E., Austin, B., Davis, C., et al.: Improving chronic illness care: Translating evidence into action. Health Aff. (Millwood) 20(6):64-78, 2001.

Wagner, E.H.: Chronic disease management: What will it take to improve care for chronic illness? Eff. Clin. Pract. 1(1):2-4, 1998.

Ware, J.E., Jr., Bayliss, M.S., Rogers, W.H., Kosinski, M., and Tarlov, A.R.: Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems: Results from the Medical Outcomes Study. JAMA 276:1039-1047, 1996.

APPENDIX A

SUPPLEMENT 2A
DETAILED SPECIFICATIONS FOR THE CONSTRUCTION OF CLINICAL ANALYTIC VARIABLES

1. Health Status Variables

a. Charlson Comorbidity Index SAS Code

Array all the diagnoses from the dataset and search for each of the codes in the Charlson categories. If any are found, the category has a value of 1, else 0. Add weighted categories to create Charlson score.

```
AMI=0;          Acute Myocardial Infarction;
CHF=0;          Congestive Heart Failure;
PVD=0;          Peripheral Vascular Disease;
CVD=0;          Cerebrovascular Disease;
dementia=0;     Dementia;
COPD=0;         Chronic Pulmonary disease;
conn_tissuedz=0; Connective Tissue disease;
ulcer=0;        Ulcer disease;
liverdz_mild=0; Mild liver disease;
diabetes=0;     Diabetes without complications;
hemiplegia=0;   Hemiplegia;
CRF=0;          Moderate or severe renal disease;
DMwcc=0;        Diabetes with complications;
neoplasia=0;    Neoplasia;
leukemia=0;     Leukemia;
lymphoma=0;     Lymphoma;
liverdz_modsev=0; Moderate or severe liver disease;
cancer_mets=0;  Metastatic solid tumor;
HIV=0;          HIV/AIDS

array diag(6) diag1 diag2 diag3 diag4 diag5 diag6;
do i = 1 to 6;
dg3 = substr(diag(i),1,3);
dg4 = substr(diag(i),1,4);

select;
when (dg3='410') AMI=1;
when (dg3='428') CHF=1;
when (dg3='441' or dg4 in ('4439' '7854' 'V434')) PVD=1;
when (dg3 in ('430' '431' '432' '433' '434' '435' '436' '437' '438')) CVD=1;
when (dg3='290') dementia=1;
when (dg3 in ('490' '491' '492' '493' '494' '495' '496' '500' '501' '502' '503' '504' '505') or
    dg4='5064') COPD=1;
when (dg3 in ('710' '714' '725')) conn_tissuedz=1;
when (dg3 in ('531' '532' '533' '534')) ulcer=1;
when (dg3 in (‘571’)) liverdz_mild=1;
when (dg3 in ('250','249') or dg4 in ('7915','9623') or
 &dx in ('V5867','99657')) diabetes=1;
when (dg3='342' or dg4='3441') hemiplegia=1;
```

```
when (dg3 in ('582' '583' '585' '586' '588')) chronic renal failure=1;
when (dg4 in ('2504' '2505' '2506')) diabetes with complications=1;
when (dg3 in ('200' '201' '202' '203' '204')) lymphoma=1;
when (dg3 in ('205' '206' '207' '208')) leukemia=1;
when (dg3 in ('140' '141' '142' '143' '144' '145' '146' '147' '148' '149' '150' '151' '152' '153'
    '154' '155' '156' '157' '158' '159' '160' '161' '162' '163' '164' '165' '170' '171' '172' '174'
    '175' '176' '179' '180' '181' '182' '183' '184' '185' '186' '187' '188' '189' '190' '191' '192'
    '193' '194' '195')) neoplasia=1;
when (dg4 in ('5722' '5723' '5724' '5728' '4560' '4561' '4562')) moderate to severe liver
    disease=1;
when (dg3 in ('196' '197' '198' '199')) metastisized cancer =1;
when (dg3 in ('042' '043' '044')) HIV=1;
otherwise;
end; end;

chscore=AMI + CHF + PVD + CVD + dementia + COPD + conn_tissuedz + ulcer +
  liverdz_mild + diabetes + 2*hemiplegia + 2*CRF + 2*DMwcc + 2*neoplasia +
  2*leukemia + 2*lymphoma + 3*liverdz_modsev + 6*cancer_mets + 6*HIV;
```

b. Chronic Conditions SAS code

```
%MACRO CHECKCC(DX);
DX4=SUBSTR(&DX,1,4);
DX3=SUBSTR(&DX,1,3);
DXL=SUBSTR(&DX,5,1);
IF DX4='4280' THEN CHF_CC=1;
IF (('41400'<=&DX<='41407') OR
 ('41000'<=&DX<='41092') OR
    DX4 in ('4142','4143','4148','4149') OR
 ('4110 '<=&DX<='41189') OR
 ('4130'<=DX4<='4139') OR DX3='412') THEN CAD_CC=1;
 IF (DX3 IN ('496','492','493','494') OR DX4='4912') THEN
 RESP_CC=1;
 IF DX4='2500' or DX4='2490' THEN DIABWO_CC=1;
 IF ('2501'<=DX4<='2509' or '2491'<=DX4<='2499' or
 DX4 in ('7915','9623') or &dx in ('V5867','99657')) THEN DIABC_CC=1;
 IF (DX3='401') THEN HYPER_CC=1;
 IF (DX3='424') THEN VALV_CC=1;
 IF (DX3='425') THEN CARD_CC=1;
 IF (DX3 IN ('584','586')) THEN RENFAIL_CC=1;
 IF (DX4='4439') THEN PVD_CC=1;
 IF (DX3='272') THEN LIPID_CC=1;
 IF (DX3 IN ('427','426')) THEN DYS_CC=1;
 IF (DX3='290') THEN DEM_CC=1;
 IF ((DX3 IN ('434','433') & DXL='1') OR DX3='431' OR
 &DX='V1259') THEN STROKE_CC=1;
 IF (DX4 IN ('2504','4039','5811','5818','5819','5829','5939','5996','7100',
```

```
'7531','7910') OR DX3 IN ('582','585') OR &DX='58381') THEN ACREN_CC=1;
IF DX4='7865' then CHPAIN_CC=1;
IF DX4 in ('5990','5999') THEN UTI_CC=1;
IF DX3='285' THEN ANEMIA_CC=1;
IF DX4='7807' THEN MALAISE_CC=1;
IF (&DX IN ('78002','78009','78093','78097','78039') OR DX4 IN ('7802','7804'))
THEN DIZZ_CC=1;
IF DX3='719' THEN JOINT_CC=1;
IF DX3='244' THEN THYROID_CC=1;

%MEND;

%LET CCDXLIST=%STR(CHF_CC CAD_CC RESP_CC DIABWO_CC DIABC_CC
    HYPER_CC VALV_CC CARD_CC ACREN_CC RENFAIL_CC PVD_CC
    LIPID_CC DYS_CC DEM_CC STROKE_CC CHPAIN_CC UTI_CC ANEMIA_CC
    MALAISE_CC DIZZ_CC JOINT_CC THYROID_CC);
```

c. Ambulatory Care Sensitive Conditions (ACSCs).

```
%LET ACSCLIST = %STR(ALL DIAB CELL ASTHMA COPD CHF DHYD PNEU
    SEPT STROKE UTI);
%macro chkdx(diag);
dx3=substr(&diag,1,3);
dx4=substr(&diag,1,4);
all=1;
if dx3 in ('250','249') or dx4 in ('7915','9623') or
 &diag in ('V5867','99657') then diab=1;
if dx3 in ('681','682') then cell=1;
if dx3 in ('493') then asthma=1;
if dx3 in ('491','492','494','496') then copd=1;
if dx3='428' or &diag in ('40201','40211','40291','40401','40411','40491',
 '39891','40403','40413','40493','78550','78551') then chf=1;
if dx4='2765' then dhyd=1;
if dx3 in ('481','482','483','485','486') then pneu=1;
if dx3='038' then sept=1;
if dx3 in ('434','436') then stroke=1;
if dx4 in ('5990','5999') then uti=1;
```

2. Hospitalization, Emergency Room and Readmission Analytic Variables

To report descriptive statistics on the rates of ACSCs by location of service using claims files to create of rates of ACSCs by location of service: 1) inpatient; 2) hospital outpatient department or physician’s office; and) ER/observation bed stays. For example, we will be examining the number of inpatient cellulitis admissions per 1,000 beneficiaries, the number of physician office/OPD visits per 1,000 beneficiaries, and the number of ER visits per 1,000 beneficiaries in the baseline, and the last 12 months of the intervention period.

A. Hospitalizations: Step 1 Combine transfer records as follows:

1. If the admission date (ADMSN_DT) or discharge date (DSCHRGDT) is missing on the claim, or equal to "0," set them equal to "from" (FROM_DT) and "through" (THRU_DT) dates, respectively.

2. Combine multiple claims that represent pieces of stays or transfers between hospitals, or separately administered units of a single hospital, into a single record representing an admission. Some records in the Inpatient claims file that look like new admissions are actually transfers between or within facilities. This process uses all claims; do not exclude claims for periods if ineligibility until after the transfers have been processed.

    a. Create a claim type variable as CLMB_TYP = FAC_TYPE || TYPESRVC

    b. Sort the data by HICNO FROM_DT THRU_DT

    c. Designate the first record for each HICNO in the reference period as a new admission.

    d. If the length between reference record discharge date and next admission date is more than one day, the next admission record is considered a new admission.

    e. If the discharge status code of the reference record is not equal to 30, 02, 05, 61, or 62 and the status code of the record previous to the reference record is not equal to 30, 02, 05, 61, or 62, then the reference record is considered a new admission. The definition of the discharge status codes are:
    30: Still a patient
    02: Discharged/transferred to other short term general hospital for inpatient care
    05: Discharged/transferred to skilled nursing facility (SNF)
    61: Discharged/transferred within this institution to a hospital-based Medicare-approved swing bed (1/1/02)
    62: Discharged to another IRF or IRF unit (1/1/02)

    f. If the discharge status code of the record previous to the reference record is equal to 30, 02, 05, 61, or 62 and the difference between the reference record's admission date and the record previous to the reference record's admission date is less than or equal to 1 day, then the reference record is considered a transfer.

    g. If the discharge status code of the reference record is equal to 30, 02, 05, 61, or 62 and the discharge status code of the record previous to the reference record is not equal to 30, 02, 05, 61, or 62, then the reference record is considered a new admission.

    h. The length of stay is calculated, as described for the row 2 measure below. If the length of stay is negative, the record is removed.

    i. The system counts each unique admission falling within the reference period.

j. Note that admission dates that fall within the reference period are counted even if the discharge date falls outside of the reference period. Also note that, in some cases, the system will be missing the later pieces of a stay that commences within the period, especially when hospitals "split-bill" at calendar year-end, but the admission will still be counted in the reference period.

B. Step 2: Create Causes of Hospitalization Analytic Variables: All cause and 10 ACSCs

(1) All cause hospitalizations:
Select if PDGNS_CD = any diagnosis code

(2) Heart failure hospitalization:
Select if PDGNS_CD =
428
40201
40211
40291
40401
40411
40491
39891
40403
40413
40493
78550
78551

(3) Diabetes hospitalization:
Select if PDGNS_CD =
250
249
7915
9623
V5867
99657

(4) Cellulitis:
Select if PDGNS_CD =
681
682

(5) Asthma hospitalization:
Select if PDGNS_CD =
493

(6) COPD and Chronic Bronchitis
Select if PDGNS_CD =
491
492
494
496

(7) Dehydration
Select if PDGNS_CD =
2765

(8) Bacterial Pneumonia
Select if PDGNS_CD =
481
482

	483
	485
	486
(9) Septicemia	
Select if PDGNS_CD =	038
(10) Ischemic Stroke	
Select if PDGNS_CD =	434
	436
(11) UTI	
Select if PDGNS_CD =	5990
	5999

C. Emergency Room Visits, including observation stays

Calculate the number of beneficiary visits to a hospital's outpatient emergency room (ER) **or** for an observation stay during the reference period. Restrict the measure to ER and observation visits identified on the Outpatient (OPD) claims file. Keep records with a revenue center line item (REV_CNTR) equal to 045X or 0981 (emergency room care) unless the HCPCS for the line item equals 70000 through 79999 or 80000 through 89999 (thus excluding claims where only radiological or pathology/laboratory services were provided) for revenue code dates (REV_DT) that fall within the reference period. Keep records with a revenue center line item (REV_CNTR) equal to 0762 (treatment of observation room-observation room) for revenue code dates (REV_DT) that fall within the reference period. This will capture ER claims for beneficiaries that were not subsequently admitted to the hospital.

To capture ER visits that led to a hospitalization, claims are identified in the MedPAR (inpatient) file. Keep records with revenue center code values of 0450-0459, 0981, and 0762. The diagnostic emergency room details are on the inpatient claim.

Count each of the 10 types of ACSC visits for a unique beneficiary on a unique date. If a beneficiary has more than one visit on the same day, count them insofar as they are of different types. That is, no one can have more than one "all cause" visits on a given day; no one can have more than one CHF visit on a given day. A person can have a CHF visit and a CAD visit on the same day, however. Visit type is the same as for hospitalizations.

D. 30-day Hospital Readmissions

Each admission within the reference period is eligible to be a readmission; that is, a single beneficiary can be counted more than once if she/he had more than one hospital admission during the period. Calculate all measures after handling transfers, as described in the hospital admission specifications. After identifying unique hospital admissions in the reference period, calculate the number of days between the admission date and the most immediate previous discharge date, if any, from a short-stay acute-care inpatient hospital department, for any reason, as identified in the Inpatient claims file. Flag as a 90-day readmit, if admission date is less than or equal to 90 days from date of discharge. The intervention period examined admissions during the period from 15 months through 3 months prior to the end of the demonstration and included readmissions through the end of the demonstration period. We constructed: all cause readmission rates for all

hospitalizations and same cause readmission rates for the ten ambulatory care sensitive conditions.

a. All cause readmissions after all cause hospitalizations

b. Same cause readmissions for the 10 ACSCs.

3. Guideline Concordant Care

Quality of Care Variables

1) Rate of influenza shots during influenza season (September through February) for adults – all beneficiaries (AMA, NQF endorsed measure – for patients $\geq$ 50 years but we will evaluate for all beneficiaries).

- ***Denominator***: All beneficiaries with at least one day of eligibility in both baseline and the demo period(s). (Note: we are not excluding those with egg allergies or known adverse reaction to influenza vaccine in the past for simplification.)

- ***Numerator:*** Beneficiaries who receive a test between September 1 and February 28 (or 29th if a leap year (2004, 2008, 2012)) for the baseline or demo periods.

    i. For the MGH original population, the dates would be as follows:
    Baseline: 9/1/05 – 2/28/06
    Demo Period 1: 2/1/07-2/28/07; 9/1/07 – 1/31/08
    Demo Period 2: 9/1/08 – 2/28/09
    For the MGH refresh population, the dates would be as follows:
    Baseline: 9/1/06 – 2/28/07
    Demo Period 1: 9/1/08 – 2/28/09

    ii. CPT Codes to define receipt of influenza vaccine in either physician claims or OPD file: 90656, 90658, 90660, 90661, 90662, 90663, G0008

2) Rate of annual HbA1c testing – beneficiaries with diabetes in baseline (Alliance, NQF endorsed measure – exclusive of CPT II or LOINC codes for identification of test being performed).

- ***Denominator***: All beneficiaries with diabetes identified in the baseline period and at least one day of eligibility in both baseline and the demo period.

- ***Numerator:*** Beneficiaries who have a claim for a test as defined by CPT codes in the physician and OPD file: 83036, 83037.

3) Rate of annual low-density lipoprotein cholesterol (LDL-C) testing – beneficiaries with diabetes or ischemic vascular disease (Alliance, NQF endorsed for diabetes and NCQA, NQF endorsed for ischemic vascular disease – exclusive of CPT II or LOINC codes for identification of test being performed).

- ***Denominator A***: All beneficiaries with diabetes identified in the baseline period and at least one day of eligibility in both baseline and the demo periods.

- ***Denominator B***: All beneficiaries with ischemic vascular disease identified in the baseline period and at least one day of eligibility in both baseline and the demo periods.

- ***Numerator:*** Beneficiaries who have a claim for a test as defined by CPT codes in the physician and OPD file: 80061, 83715, 83700, 83716, 83701, 83704, 83721.

SUPPLEMENT 4A
PARTICIPATION TABLES

Supplement Table 4A-1
Characteristics of MGH's CMP intervention and comparison populations: Original population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Total number of beneficiaries	2,584	2,719	—	—
Full time equivalent	2,137	2,153	—	—
Beneficiary characteristics				
Aged-in (vs. disabled)	87.6	82.5	5.0	**
In Medicaid (vs. not in Medicaid)	3.7	6.8	-3.2	**
Male (vs. female)	47.5	47.1	0.4	N/S
Urban (vs. rural)	100.0	100.0	0.0	N/S
Age				
Mean	75.0	73.6	1.4	**
<65	12.6	17.5	-5.0	**
65-69	11.2	11.6	-0.3	N/S
70-74	18.4	16.0	2.4	*
75-79	22.4	20.4	1.9	N/S
80-84	19.2	19.2	0.0	N/S
85+	16.3	15.3	1.1	N/S
Race				
White	91.6	81.6	10.0	**
African American	4.6	13.0	-8.4	**
Other	3.9	5.2	-1.4	*
Unknown	0.0	0.2	-0.2	*
Health status				
Recalculated HCC score				
Mean	2.0	2.0	-0.1	N/S
Low: $\geq$ 1.35 and < 2.00	35.3	35.2	0.1	N/S
Medium: $\geq$ 2.00 and $\leq$ 3.10	35.0	33.1	1.9	N/S
High: > 3.10	29.6	31.6	-2.0	N/S
Baseline PBPM low	34.8	34.1	0.7	N/S
Baseline PBPM medium	34.7	35.6	-0.9	N/S
Baseline PBPM high	30.6	30.4	0.2	N/S
Charlson comorbidity index—mean	3.3	3.2	0.2	*

(continued)

Supplement Table 4A-1 (continued)
Characteristics of MGH's CMP intervention and comparison populations: Original population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	P[3]
Chronic conditions				
HF	18.1	17.9	0.2	N/S
Coronary artery disease	31.2	32.0	-0.9	N/S
Other respiratory disease	22.5	20.9	1.6	N/S
Diabetes without complications	32.2	39.3	-7.1	**
Diabetes with complications	17.3	22.4	-5.2	**
Essential hypertension	60.4	53.6	6.7	**
Valve disorders	7.5	6.4	1.2	N/S
Cardiomyopathy	3.9	5.7	-1.8	**
Acute & chronic renal disease	14.1	20.6	-6.6	**
Renal failure	10.1	7.6	2.5	**
Peripheral vascular disease	6.3	6.4	-0.1	N/S
Lipid metabolism disorders	42.0	35.0	7.0	**
Cardiac dysrhythmias & conduction disorders	25.1	25.9	-0.8	N/S
Dementias	1.2	0.8	0.4	N/S
Strokes	4.4	3.4	1.0	N/S
Chest pain	8.4	10.3	-1.9	*
Urinary tract infection	11.9	9.6	2.3	**
Anemia	17.7	18.1	-0.4	N/S
Malaise & fatigue (including CFS)	9.5	5.7	3.8	**
Dizziness, syncope, convulsions	12.6	11.1	1.5	N/S
Disorders of joint	16.0	12.0	4.0	**
Hypothyroidism	10.8	9.3	1.5	N/S

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; I = intervention population; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

[3] * denotes statistical significance at the 5% level; ** denotes statistical significance at the 1% level.

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-3.sas 27JUN2010.

Supplement Table 4A-2
Characteristics of MGH's CMP intervention and comparison populations: Refresh population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Total number of beneficiaries	775	766	—	—
Full time equivalent	695	660	—	—
Beneficiary characteristics				
Aged-in (vs. disabled)	86.4	82.3	4.1	*
In Medicaid (vs. not in Medicaid)	4.5	6.2	-1.8	N/S
Male (vs. female)	45.1	46.8	-1.7	N/S
Urban (vs. rural)	100.0	100.0	0.0	N/S
Age				
Mean	74.8	74.6	0.2	N/S
<65	13.5	17.2	-3.8	*
65-69	12.2	9.9	2.3	N/S
70-74	15.5	13.1	2.4	N/S
75-79	20.3	19.7	0.6	N/S
80-84	20.1	20.9	-0.8	N/S
85+	18.5	19.3	-0.7	N/S
Race				
White	92.1	82.4	9.7	**
African American	3.7	12.6	-8.9	**
Other	4.2	4.6	-0.4	N/S
Unknown	0.1	0.5	-0.4	N/S
Health Status				
Recalculated HCC score				
Mean	1.7	2.0	-0.3	**
Low: $\geq$ 1.35 and < 2.00	35.9	33.1	2.8	N/S
Medium: $\geq$ 2.00 and $\leq$ 3.10	37.4	31.7	5.7	*
High: > 3.10	26.8	35.2	-8.5	**
Baseline PBPM low	34.2	34.5	-0.3	N/S
Baseline PBPM medium	35.0	33.9	1.1	N/S
Baseline PBPM high	30.8	31.6	-0.8	N/S
Charlson comorbidity index—mean	3.1	3.4	-0.3	*

(continued)

Supplement Table 4A-2 (continued)
Characteristics of MGH's CMP intervention and comparison populations: Refresh population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Chronic conditions				
HF	16.8	21.0	-4.2	*
Coronary artery disease	30.2	33.2	-3.0	N/S
Other respiratory disease	20.7	22.5	-1.8	N/S
Diabetes without complications	28.2	37.3	-9.1	**
Diabetes with complications	15.2	22.8	-7.7	**
Essential hypertension	64.9	60.1	4.8	N/S
Valve disorders	8.3	7.9	0.4	N/S
Cardiomyopathy	3.6	8.6	-5.0	**
Acute & chronic renal disease	16.9	21.7	-4.8	*
Renal failure	9.6	9.5	0.1	N/S
Peripheral vascular disease	7.0	7.3	-0.3	N/S
Lipid metabolism disorders	47.1	46.2	0.9	N/S
Cardiac dysrhythmias & conduction disorders	25.1	25.3	-0.2	N/S
Dementias	1.2	1.4	-0.2	N/S
Strokes	4.0	4.1	-0.1	N/S
Chest pain	10.7	12.4	-1.8	N/S
Urinary tract infection	10.8	10.9	-0.2	N/S
Anemia	21.7	21.8	0.0	N/S
Malaise & fatigue (including CFS)	10.3	8.8	1.6	N/S
Dizziness, syncope, convulsions	12.8	12.8	0.1	N/S
Disorders of joint	18.3	14.6	3.7	N/S
Hypothyroidism	13.8	10.4	3.4	*

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; I = intervention population; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

[3] * denotes statistical significance at the 5% level; ** denotes statistical significance at the 1% level.

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-3.sas 27JUN2010.

Supplement Table 4A-3
Characteristics of MGH's CMP intervention population by participation status: Original population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Total number of beneficiaries	2,280	1,985	304	—	—
Full time equivalent	1,949	1,803	188	—	—
Beneficiary characteristics					
Aged-in (vs. disabled)	87.8	87.7	85.1	2.7	N/S
In Medicaid (vs. not in Medicaid)	3.5	3.4	5.1	-1.6	N/S
Male (vs. female)	47.7	47.5	45.1	2.6	N/S
Urban (vs. rural)	100.0	100.0	100.0	0.0	N/S
Age					
Mean	75.0	75.0	74.9	0.0	N/S
<65	12.3	12.4	14.9	-2.6	N/S
65-69	11.5	11.5	8.3	3.2	N/S
70-74	18.5	18.6	17.1	1.4	N/S
75-79	22.3	22.2	22.8	-0.5	N/S
80-84	19.3	19.4	17.3	2.0	N/S
85+	16.0	15.8	19.5	-3.5	N/S
Race					
White	91.5	91.7	92.7	-1.2	N/S
African American	4.7	4.6	3.4	1.3	N/S
Other	3.8	3.7	4.0	-0.1	N/S
Unknown	0.0	0.0	0.0	0.0	N/S
Health status					
Recalculated HCC score					
Mean	2.0	2.0	1.8	0.2	**
Low: $\geq$ 1.35 and < 2.00	34.9	34.8	40.4	-5.5	N/S
Medium: $\geq$ 2.00 and $\leq$ 3.10	34.5	34.8	40.2	-5.6	N/S
High: > 3.10	30.6	30.4	19.5	11.2	**
Baseline PBPM low	34.4	35.0	38.4	-4.0	N/S
Baseline PBPM medium	34.8	35.1	33.8	1.0	N/S
Baseline PBPM high	30.8	30.0	27.8	3.0	N/S
Charlson comorbidity index—mean	3.3	3.3	3.3	0.1	N/S

(continued)

Supplement Table 4A-3 (continued)
Characteristics of MGH's CMP intervention population by participation status: Original population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	P[3]
Chronic conditions					
HF	18.5	18.4	13.7	4.8	N/S
Coronary artery disease	31.4	31.4	29.2	2.1	N/S
Other respiratory disease	22.8	22.6	20.1	2.7	N/S
Diabetes without complications	32.2	32.3	32.3	0.0	N/S
Diabetes with complications	17.3	17.3	16.8	0.5	N/S
Essential hypertension	60.6	60.7	57.8	2.8	N/S
Valve disorders	7.5	7.1	8.2	-0.8	N/S
Cardiomyopathy	3.9	3.8	4.0	-0.1	N/S
Acute & chronic renal disease	14.1	13.6	13.6	0.5	N/S
Renal failure	10.1	9.8	10.1	0.0	N/S
Peripheral vascular disease	6.2	6.3	7.7	-1.5	N/S
Lipid metabolism disorders	42.4	42.7	38.1	4.3	N/S
Cardiac dysrhythmias & conduction disorders	25.5	25.1	20.0	5.5	N/S
Dementias	1.1	1.1	1.8	-0.7	N/S
Strokes	4.2	4.1	6.2	-2.0	N/S
Chest pain	8.4	8.3	8.1	0.3	N/S
Urinary tract infection	11.8	11.8	13.1	-1.3	N/S
Anemia	17.5	17.3	19.6	-2.1	N/S
Malaise & fatigue (including CFS)	9.5	9.6	9.8	-0.3	N/S
Dizziness, syncope, convulsions	12.7	12.6	11.1	1.6	N/S
Disorders of joint	15.9	15.8	17.8	-1.9	N/S
Hypothyroidism	11.1	11.2	7.3	3.8	N/S

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; P = participating; NP = nonparticipating; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

[3] * denotes statistical significance at the 5% level; ** denotes statistical significance at the 1% level.

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-4.sas 27JUN2010.

Supplement Table 4A-4
Characteristics of MGH's CMP intervention population by participation status: Refresh population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Total number of beneficiaries	648	574	127	—	—
Full time equivalent	598	557	97	—	—
Beneficiary characteristics					
Aged-in (vs. disabled)	86.4	85.8	86.4	0.0	N/S
In Medicaid (vs. not in Medicaid)	4.5	4.6	4.4	0.1	N/S
Male (vs. female)	43.8	42.9	53.4	-9.7	N/S
Urban (vs. rural)	100.0	100.0	100.0	0.0	N/S
Age					
Mean	74.9	74.6	74.5	0.4	N/S
<65	13.4	14.0	13.6	-0.2	N/S
65-69	12.0	12.1	13.2	-1.2	N/S
70-74	15.1	15.4	18.0	-2.9	N/S
75-79	20.6	20.8	18.2	2.3	N/S
80-84	20.2	20.3	20.0	0.1	N/S
85+	18.8	17.4	17.0	1.8	N/S
Race					
White	91.8	92.2	93.7	-1.9	N/S
African American	3.7	3.5	3.3	0.4	N/S
Other	4.4	4.3	3.0	1.4	N/S
Unknown	0.1	0.0	0.0	0.1	N/S
Health status					
Recalculated HCC score					
Mean	1.7	1.7	1.8	0.0	N/S
Low: $\geq$ 1.35 and < 2.00	35.0	35.5	41.0	-5.9	N/S
Medium: $\geq$ 2.00 and $\leq$ 3.10	38.7	38.9	29.5	9.2	N/S
High: > 3.10	26.3	25.5	29.5	-3.2	N/S
Baseline PBPM low	34.6	35.0	32.0	2.6	N/S
Baseline PBPM medium	34.0	34.4	41.2	-7.2	N/S
Baseline PBPM high	31.4	30.6	26.9	4.6	N/S
Charlson comorbidity index—mean	3.1	3.0	3.2	-0.1	N/S

(continued)

Supplement Table 4A-4 (continued)
Characteristics of MGH's CMP intervention population by participation status: Refresh population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Chronic conditions					
HF	16.3	15.1	19.7	-3.4	N/S
Coronary artery disease	30.6	30.9	28.0	2.6	N/S
Other respiratory disease	21.5	20.2	15.6	5.9	N/S
Diabetes without complications	26.8	27.4	36.5	-9.7	N/S
Diabetes with complications	15.5	15.8	13.2	2.3	*
Essential hypertension	65.2	65.4	62.5	2.7	N/S
Valve disorders	8.5	8.1	6.7	1.8	N/S
Cardiomyopathy	3.3	2.9	5.4	-2.1	N/S
Acute & chronic renal disease	16.7	16.1	17.9	-1.2	N/S
Renal failure	9.7	9.5	9.1	0.6	N/S
Peripheral vascular disease	7.5	7.3	3.5	4.0	N/S
Lipid metabolism disorders	48.6	49.1	38.2	10.4	*
Cardiac dysrhythmias & conduction disorders	25.1	24.2	24.9	0.3	N/S
Dementias	1.1	1.1	1.4	-0.3	N/S
Strokes	4.3	4.3	2.0	2.3	N/S
Chest pain	10.0	10.1	14.7	-4.7	N/S
Urinary tract infection	10.3	10.3	14.0	-3.8	N/S
Anemia	22.4	22.0	17.8	4.5	N/S
Malaise & fatigue (including CFS)	10.3	9.7	10.5	-0.3	N/S
Dizziness, syncope, convulsions	12.7	12.3	13.6	-0.9	N/S
Disorders of joint	18.9	18.6	14.5	4.4	N/S
Hypothyroidism	13.0	13.3	18.6	-5.7	N/S

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; P = participating; NP = nonparticipating; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the program.

[3] * denotes statistical significance at the 5% level; ** denotes statistical significance at the 1% level.

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mgh/final/tables/tableHB-4.sas 27JUN2010.

Supplement Table 4A-5
Participation rates during the first 6 months of MGH's CMP by beneficiary characteristics, baseline characteristics, and intervention period health status: Original and refresh populations

Characteristics	Original (%)	Refresh (%)
Overall participation rate[1,2]	86	84
Beneficiary characteristics		
Male	86	71
Female	86	86
White	86	84
African American/other/unknown	85	82
Age < 65 years	83	84
Age 65-74	87	83
Age 75-84	87	83
Age 85 + years	85	85
Medicaid	82	82
Non-Medicaid	86	84
Baseline characteristics		
Baseline HCC score low	85	82
Baseline HCC score high	87	80
Low baseline PBPM	85	85
High baseline PBPM	85	84
Baseline Charlson score low	86	87
Baseline Charlson score high	85	85
Demonstration period health status		
Died	84	83
Alive	87	84
Institutionalized	51	59
Not institutionalized	87	84
Concurrent HCC score low	86	84
Concurrent HCC score high	86	83
Number of participants	2,175	632
Number of total beneficiaries	2,523	756

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] Counts of beneficiaries are adjusted for CMHCB program eligibility during the first six months the Care Management Organization (CMO) was active in the program.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: partab2.sas 27JUNE2010.

Supplement Table 4A-6
Logistic regression modeling results comparing beneficiaries that participated at least one eligible month in the first 6 months of MGH's CMP to all other intervention beneficiaries: original population[1,2]

Characteristics	Model 1A OR	p^3	Model 1B OR	p^3
Intercept	7.03	**	6.70	**
Beneficiary characteristics				
Male	0.95	N/S	0.92	N/S
African American/other/unknown	0.94	N/S	0.92	N/S
Age < 65 years	0.75	N/S	0.70	N/S
Age 75-84	0.98	N/S	0.99	N/S
Age 85 + years	0.86	N/S	0.88	N/S
Medicaid	0.72	N/S	0.72	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.27	N/S
Baseline HCC score high	N/I	N/I	1.78	**
Medium baseline PBPM	N/I	N/I	1.15	N/S
High baseline PBPM	N/I	N/I	0.86	N/S
Baseline Charlson score medium	N/I	N/I	0.92	N/S
Baseline Charlson score high	N/I	N/I	0.69	N/S
Demonstration period health status				
Died	N/I	N/I	0.79	N/S
Concurrent HCC score medium	N/I	N/I	1.06	N/S
Concurrent HCC score high	N/I	N/I	1.01	N/S
Number of cases	2,584	N/A	2,584	N/A
Chi-square (p<)	4.94	N/S	20.38	N/S
Pseudo R-square	0.00	N/A	0.01	N/A

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the first 6 months the Care Management Organization (CMO) was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $336.70. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .811 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data. Program: bene02 27JUNE2010, partab3a and partab4a 28JUNE2010.

Supplement Table 4A-7
Logistic regression modeling results comparing beneficiaries that participated at least one eligible month in the first 6 months of MGH's CMP to all other intervention beneficiaries: refresh population[1,2]

Characteristics	Model 1A OR	p^3	Model 1B OR	p^3
Intercept	5.77	**	6.74	**
Beneficiary characteristics				
Male	0.70	N/S	0.70	N/S
African American/other/unknown	0.89	**	0.89	N/S
Age < 65 years	1.17	N/S	1.20	N/S
Age 75-84	1.05	N/S	1.01	N/S
Age 85 + years	1.14	N/S	1.10	N/S
Medicaid	0.89	N/S	0.87	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	2.07	*
Baseline HCC score high	N/I	N/I	1.03	N/S
Medium baseline PBPM	N/I	N/I	0.73	N/S
High baseline PBPM	N/I	N/I	0.94	N/S
Baseline Charlson score medium	N/I	N/I	0.54	*
Baseline Charlson score high	N/I	N/I	0.90	N/S
Demonstration period health status				
Died	N/I	N/I	0.89	N/S
Concurrent HCC score medium	N/I	N/I	1.04	N/S
Concurrent HCC score high	N/I	N/I	1.01	N/S
Number of cases	775	N/A	775	N/A
Chi-square (p<)	3.77	N/S	19.17	N/S
Pseudo R-square	0.00	N/A	0.02	N/A

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the first 6 months the Care Management Organization (CMO) was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $324. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .694 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: bene02 27JUNE2010, partab3a and partab4a 28JUNE2010.

Supplement Table 4A-8
Logistic regression modeling results comparing beneficiaries that participated at least one eligible month in MGH's CMP to all other intervention beneficiaries: original population[1,2]

Characteristics	Model 2 OR	p[3]
Intercept	11.99	**
Beneficiary characteristics		
Male	1.06	N/S
African American/other/unknown	1.25	N/S
Age < 65 years	0.65	N/S
Age 75-84	0.86	N/S
Age 85 + years	0.64	N/S
Medicaid	0.63	N/S
Baseline characteristics		
Baseline HCC score medium	1.16	N/S
Baseline HCC score high	2.75	**
Medium baseline PBPM	1.03	N/S
High baseline PBPM	0.94	N/S
Baseline Charlson score medium	0.90	N/S
Baseline Charlson score high	0.55	*
Demonstration period health status		
Died	1.04	N/S
Concurrent HCC score medium	0.92	N/S
Concurrent HCC score high	0.86	N/S
Number of cases	2,584	N/A
Chi-square (p<)	24.05	N/S
Pseudo R-square	0.01	N/A

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $336.70. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .811 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.

Program: bene02 27JUNE2010, partab1 28JUNE2010.

Supplement Table 4A-9
Logistic regression modeling results comparing beneficiaries that participated at least one eligible month in MGH's CMP to all other intervention beneficiaries: refresh population[1,2]

Characteristics	Model 2 OR	p[3]
Intercept	7.68	**
Beneficiary characteristics		
Male	0.69	N/S
African American/other/unknown	1.32	**
Age < 65 years	1.11	N/S
Age 75-84	1.17	N/S
Age 85 + years	1.12	N/S
Medicaid	0.97	N/S
Baseline characteristics		
Baseline HCC score medium	2.12	*
Baseline HCC score high	1.20	N/S
Medium baseline PBPM	0.81	N/S
High baseline PBPM	1.09	N/S
Baseline Charlson score medium	0.43	**
Baseline Charlson score high	0.82	N/S
Demonstration period health status		
Died	1.28	N/S
Concurrent HCC score medium	0.93	N/S
Concurrent HCC score high	0.99	N/S
Number of cases	775	N/A
Chi-square (p<)	20.60	N/S
Pseudo R-square	0.03	N/A

NOTES: MGH = Massachusetts General Hospital; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Numbers reported for the intervention periods include only persons who have some baseline eligibility.

[2] The regressions are adjusted for CMHCB program eligibility during the entire period the Care Management Organization (CMO) was active in the demonstration.

[3] * denotes statistical significance at the 5% level;** denotes statistical significance at the 1% level.

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $324. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .694 or less.

Data Sources: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data.
Program: bene02 27JUNE2010, partab1 28JUNE2010

SUPPLEMENT 7A
REGRESSION-TO-THE-MEAN

Regression-to-the-mean (RtoM) cannot be quantified simply by tracking the change in mean PBPM costs because of secular changes in costs of a particular group. RtoM more specifically refers to low (high) initial costs gravitating to the mean cost over time which could be rising or falling due to other factors. It would be possible to observe a rising PBM mean cost still with significant RtoM. Unbiased random sampling of a chronically ill population should have most of the positive and negative changes in beneficiary PBPM costs cancelling out, leaving the secular growth trend. A "biased" sample of high cost chronically ill, by contrast, should produce more declines in costs than increases and a lower (negative?) cost trend.

To estimate the impact of RtoM, we specify the following equation:

$$\Delta PBPM_{tp} = PBPM_{tp} - PBPM_{bp} = \alpha + \rho[PBPM_{bp} - PBPM_b^*] + \beta Status_p + \varepsilon_{tp} \qquad (7.1a)$$

- $\Delta PBPM_{tp}$ = the change in PBPM cost between the base period (b) and current period (t) for the p-th patient.
- $PBPM_{tp}$, $PBPM_{bp}$ = the p-th patient's average PBPM cost in the current and base periods, respectively.
- $PBPM_b^*$ = the mean PBPM cost for all patients in the base period.
- $Status_p$ = 1 if patient in the intervention group; 0 otherwise.

The growth in a beneficiary's PBPM cost from base to demonstration period is assumed to have a secular component, α, for the control group and $\alpha + \beta$ for the intervention group. Regression to the mean is captured by ρ. Beneficiaries with greater than average base year PBPM costs should exhibit lower PBPM costs in the demonstration period while those with below-average PBPM costs should exhibit growth in their PBPM costs, after adjusting for the secular trend in Medicare spending. Therefore, we assume that $\rho < 0$ and we should observe a compression in PBPM costs towards the secular mean rate over time. No regression to the mean would result in an estimate of $\rho = 0$. Solving equation 7.1a for $PBPM_{tp}$ gives

$$PBPM_{tp} = (\alpha - \beta Status_p - \rho PBPM_b^*) + (1 + \rho)PBPM_{bp} \qquad (7.2)$$

or

$$PBPM_{tp} = \gamma_p + \theta PBPM_{bp} \qquad (7.3)$$

where $\gamma_p = (\alpha - \beta Status_p - \rho PBPM_b^*)$ = the overall mean secular growth in PBPM costs that varies only by which study group to patient is in, and $\theta = (1 + \rho)$, or $\rho = \theta - 1$.

The ANCOVA regression specification is represented by equation 7.3. The intervention effect, β, can also be separated out of γ_p. The smaller the estimate of θ, the greater is the regression to the mean. For example, if the estimate of $\theta = 0.20$, then $\rho = 1 - .20 = -.80$, implying very substantial regression to the mean. Relative to secular growth, a $100 higher base year PBPM cost versus the mean would lower current period costs by $80 and vice-versa for a beneficiary with a base period PBPM cost of $100 less than average. At $PBPM_{bp}$ = $500 for the control group, the expected current period $PBPM_{tp}$ = $1,320, an increase of $820. At $PBPM_{bp}$ = $2,500, the current period $PBPM_{tp}$ = $1,720, a $780 decrease.