May 2011

Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Montefiore Medical Center's Care Guidance program (CGP)

Revised Final Report

Prepared for

Lorraine Johnson, ScD
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
Kevin Smith, MS
RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0207964.025.000.001

RTI INTERNATIONAL

EVALUATION OF MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION: MONTEFIORE MEDICAL CENTER'S CARE GUIDANCE PROGRAM (CGP)

by Nancy McCall, ScD
Jerry Cromwell, PhD
Carol Urato, MA
Kevin Smith MS

Federal Project Officer: Lorraine Johnson, Sc.D.

RTI International*

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

May 2011

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 Montefiore Medical Center's (MMC's) Care Guidance Program (CGP) operated under the Center for Medicare & Medicaid Services' (CMS) Care Management for High Cost Beneficiaries (CMHCB) demonstration. MMC is an integrated delivery system that provides patient care, conducts research, and serves as a teaching hospital for the Albert Einstein College of Medicine. MMC provides a full continuum of health care services (emergency, inpatient, outpatient, and post-acute care) primarily to residents of the Bronx and Westchester County, New York. A corporate subsidiary of MMC, Care Management Organization, was established in 1996 as a managed services organization to contract with various IPAs to perform administrative functions and obtain and manage risk-bearing contracts. To fulfill its commitment to Montefiore's mission of improving health in the Bronx and Westchester, MMC's Care Management Organization supports provider-driven strategic initiatives derived from the assumption and management of risk, with medical management services, financial administration capabilities, information systems, and infrastructure to manage customer and provider relations. MMC's Care Management Organization developed the Care Guidance Program to help participants access the medical care and social services they need to maintain health and avoid unnecessary hospitalization, regardless of condition.

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 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 MMC's CGP 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 care management organizations, MMC's CGP 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 relative to all eligible beneficiaries assigned to its comparison group. CMS's policy interest is in the extent to which intervention performance may be extrapolated to new population cohorts of beneficiaries in different settings, not just to those who are most cooperative and compliant within a particular demonstration program. Thus, RTI's evaluation focuses upon measuring the overall ***effectiveness*** of MMC's CGP that includes all intervention beneficiaries, and not just those agreeing to participate. A narrower ***efficacy*** analysis would restrict eligibility to those who "actively participate" in the program. Under the intent-to-treat principle, all beneficiaries selected for the intervention serve as the intervention group regardless of whether they "actively participated." To conduct an ***efficacy*** analysis, would require drawing a separate comparison group matched on a set of characteristics of the "active participant" group. By contrast, our evaluation provides a population-based estimate of the

MMC CGP's ***effectiveness*** in engaging the ***full*** group of pre-identified FFS beneficiaries. Consequently, less than full participation will dilute the program's overall ***effectiveness***.

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:

1. **Implementation.** To what extent was MMC able to implement its CGP?
2. **Reach.** How well did the CGP engage its intended audiences?
3. **Effectiveness.** To what degree did the CGP 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' 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

MMC launched its program on June 1, 2006, or two years after the mid-point of the identification period for the intervention and comparison groups[1]. MMC worked with its CMS project officer and analysts from Actuarial Research Corporation (ARC) to develop a method for selecting the starting population for GCP. Inclusion criteria for eligibility in MMC's CGP included:

- Medicare FFS beneficiaries with a primary residence in one of 16 designated ZIP codes in Bronx, New York surrounding MMC, with a high level of disease severity as indicated by Hierarchical Condition Categories (HCC) scores of 1.8 or greater.[2]
- Two visits to MMC physicians between January 1, 2004 and December 31, 2004, or one visit to MMC physicians in the 12-month claims period with no visits to other physicians, or a plurality of visits to MMC inpatient facilities, or one visit to an MMC inpatient facility and no visits to other inpatient facilities.

---

1 The MMC CGP had the longest lag time between the identification period and program launch among the six CMHCB Demonstration programs.

2 MMC's original CMHCB proposal included a minimum HCC score of 2.0 for eligibility in the program; however, this criterion was reduced to a score of 1.8 or greater to obtain an intervention group of desired size.

- Absence of selected conditions as indicated by ICD-9 diagnosis codes and DRG codes obtained from claims data, including dementia, substance abuse, and schizophrenia, among others.

The original population was further restricted using the following exclusion criteria: age less than 45, receiving the Medicare hospice benefit, receiving the Medicare end-stage renal disease (ESRD) benefit, history of dialysis treatment, resident in a nursing home or skilled nursing facility (SNF), enrolled in a Medicare Advantage (MA) plan, Medicare as a secondary payer, or no Medicare Part A or Part B coverage as of May 1, 2006. Using these criteria, a total of 2,969 Medicare beneficiaries were assigned to the CGP's original intervention group.

Following the development of the original intervention population criteria, MMC worked with CMS and RTI to develop specifications to select the original comparison group. The comparison group was selected using the following eligibility criteria:

- Medicare FFS beneficiaries with a primary residence in 16 ZIP codes in Brooklyn and Manhattan surrounding 5 comparison hospitals with household income levels and proportions of Hispanic residents similar to the intervention ZIP codes with a high level of disease severity as indicated by HCC scores of 1.8 or greater.

- A plurality of visits to at least 1 of 19 physician group practices (identified by tax identification number), 1 visit to a comparison group practice and no visits to any other physicians, or a plurality of admissions to 1 of 5 inpatient facilities or 1 admission to a comparison hospital and no admissions to any other hospitals.

- Absence of selected conditions as indicated by ICD-9 diagnosis codes and DRG codes obtained from claims data, including dementia, substance abuse, and schizophrenia.

The exclusion criteria that were applied to the intervention group were also used to limit the comparison group (i.e., age less than 45, receiving the Medicare hospice benefit, receiving the Medicare ESRD benefit, history of dialysis, resident in a nursing home or SNF, enrolled in an MA plan, Medicare as a secondary payer, or lack Medicare Part A or Part B coverage as of May 1, 2006).

In order to ensure that the comparison group had Medicare costs similar to the intervention group, the comparison group members were randomly selected from each of five cost strata representing the cost quintiles observed in the intervention population. The number of comparison beneficiaries selected from each stratum was determined by the number of intervention beneficiaries in each stratum. The final initial comparison group size was 1,837 beneficiaries.

Using 2006 Medicare claims data, a refresh intervention population was selected. The refresh intervention group selection method replicated the procedures used to define the original population. The refresh inclusion, exclusion, and loyalty criteria were the same as those specified for the original intervention population with one exception—the list of CPT and Place of Service codes was expanded to exclude more residents of skilled nursing facilities (SNFs) and nursing homes. The final refresh intervention population size was 912 beneficiaries.

The refresh comparison group selection method replicated the procedures used to define the original refresh group. The refresh inclusion, exclusion, and loyalty criteria were the same as those specified for the original intervention and refresh populations with the addition of more exclusions of SNF and nursing home residents. Eligible refresh beneficiaries were randomly selected from five cost strata as determined by the number of intervention beneficiaries in each stratum. A total of 887 beneficiaries were selected for the comparison refresh population.

Of all the CGP original intervention group beneficiaries, 75% verbally consented to participate in its demonstration at some point during the intervention period, 21% refused to participate, and 4% were not contacted or were unable to be located. Of the refresh intervention beneficiaries, 81% consented to participate at some point during the 24-month period. The percent that refused to participate was lower (14%), the percent that were not contacted or were unable to be contacted increased modestly to 5%. The CGP ended May 31, 2009 or 36 months after initiation of the original population and 24 months after the start of the refresh population.

MMC negotiated a management fee of $120 for the original intervention group during the first year and $123.84 in years 2 and 3 ($0.00 fee in the last two months for all intervention beneficiaries and no fee for intervention beneficiaries in the self-directed program as of September 2008). Fees for the refresh intervention group were $123.84 for all 24 months ($0.00 fee in the last two months and no fee for intervention beneficiaries in the self-directed program as of September 2008). 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 with the noted exceptions. The net savings requirements for MMC's CGP are 5% for the original cohort and 2.5% for the refresh cohort.

E.2 Overview of MMC's CMHCB Demonstration Program

The CGP was a complex case management program designed for the frail elderly population and disabled adults that was supported with technology (e.g., sophisticated information systems and telemonitoring) and MMC's established relationships with providers and community organizations. The CGP used a holistic approach to address the full complement of medical, psychological, and socioeconomic problems of the target population. Each program participant received interventions tailored to his or her specific needs. By partnering with program participants, their families, caregivers, and the medical community, the CGP aimed to help physicians manage high-risk patients, reduce medication complications, emergency room visits, and avoidable hospitalizations, improve the quality of life for both the participants and their caregivers, and support participants in the community.

The core of the CGP consisted of one-on-one telephone calls between participants and care managers, who linked beneficiaries with needed medical and social services. The program provided the following specific services to participants: care coordination, clinical pharmacist review, link to community support services, nutritional monitoring and counseling, psychosocial support, life care planning, and disease management and telemonitoring. Beneficiaries could participate in any or all of the program elements during the demonstration program, depending on their needs and preferences throughout the period. Participants were assigned to a care team based on the location of their residence and their primary language. In this way, each beneficiary interacted with two or three people throughout the program period. Each care team used the

following basic strategy to support participants: assess participant problems and resources and develop care plan to address identified needs, implement and deliver interventions to address participant problems, and re-assess on a regular basis and adjust care plans based on changes in participant problems and resources.

The CGP helped physicians manage their high-risk patients by providing information about patient health status and supporting patients in their efforts to comply with provider treatment plans. The CGP provided physicians with a summary of results from its comprehensive baseline assessment and informed physicians about their patients' conditions between office visits. In addition, the program's clinical staff educated patients about their conditions and prescribed treatment plans and helped participants manage psychosocial stressors that may have impacted their ability to adhere to treatment plans. Rather than request extra services from physicians, CGP staff collaborated with physicians to care for their patients. As a result, the program did not provide financial incentives to participants' physicians.

The CGP was delivered by 28 staff members: 12 full-time clinical staff and 16 full and part-time administrative staff. MMC staffed its CGP with registered nurses (RNs), licensed practical nurses (LPNs), social workers with a Master's degree (MSWs), and patient educators with a bachelor's degree and community-based experience working with the elderly. Social workers and patient educators were included in the staffing model to address the high number of psychosocial issues faced by the participants. Project staff also included a physician and a geriatric psychiatrist. The CGP staff worked in teams with RNs functioning as team leaders and MSWs and LPNs focusing upon identification of and health coaching or care coordination for clinical and/or psychosocial issues. The team leaders were readily available to assist the MSWs and LPNs with more complex patients. Patient educators were well suited to linking participants with community resources, as well as working with the larger than expected proportion of the population that was "self-directed."

The CGP developed a comprehensive, 35-page assessment tool to enable care managers to identify the specific problems faced by each participant and implement interventions that effectively target these issues. MMC contracted with LifePlans, a long-term risk management company, to do 1,000 assessments, with additional assessments conducted by the CGP clinical staff. The decision to outsource risk assessment was driven by the need to assess a large number of beneficiaries as soon as possible upon engagement with the program. In contrast to the CGP staff, LifePlans tried to conduct the full assessment during one phone call, rather than two or more. MMC planned for LifePlans to conduct 60% of all assessments; however the organization succeeded in completing assessments for only 40% of participants. Further, the CGP staff found that they needed to conduct follow-up calls with participants to more fully understand the problems identified during LifePlans' assessments. Thus, the CGP changed its enrollment process whereby the assessment was split into 2 parts (clinical vs. psychosocial). With the introduction of the refresh cohort, demographic and nonclinical (pre-baseline) data collection was done by nonclinical enrollment staff, who received extensive training for this process. Pre-baseline questions and also certain baseline questions were completed by enrollers, rather than by care managers, to improve efficiency.

For each participant, the CGP information system developed a problem list based on responses to the assessment. Clinical staff developed a customized care plan by assigning at

least one intervention to each problem identified. Notably, the program focused on those problems that could benefit from an actionable intervention. The beneficiaries receiving case management support were stratified to help determine the level and types of services they received from the CGP. The CGP was designed to contact participants at least one time per month during the period when they were actively receiving interventions.

- **Facilitation of access to and coordination of care.** The CGP provided a variety of interventions to ensure participant access to primary care services, including referrals to primary care providers, discharge planning services, referrals to home care services, and fall prevention assessments and referrals.

- **Medication management.** The CGP provided medication management support to participants, including clinical pharmacist review of medication regimens, patient education about medication, support to address issues of noncompliance, information about the Medicare Part D benefit, and referrals to resources to assist with payments for medicine.

- **Facilitation of access to community support services.** The CGP staff helped participants access a variety of community services to help them maintain their independence and quality of life (e.g., Meals on Wheels, transportation services, community-based case management services, adult day care, senior centers, caregiver support programs, and extended in-home services for the elderly population).

- **Nutritional monitoring and counseling.** Clinical staff monitored patients for issues related to proper nutrition (e.g., malnutrition, obesity, and special diets to manage chronic conditions).

- **Psychosocial support.** The CGP provided support to participants to address social isolation and depression, issues common to the frail elderly population.

- **Life care planning.** Case managers provided education and support to help participants establish advanced directives, and consider options for palliative and hospice care as needed.

- **Disease management and telemonitoring.** MMC's CMO developed disease management programs for diabetes, heart failure (HF), and chronic obstructive pulmonary disease (COPD). In the CGP, beneficiaries with one or more of these conditions and no other major or complex medical needs received disease management services.

- **Reassess and adjust.** Routine reassessment of participant issues and adjustment of interventions delivered on a routine basis occurred at least once per quarter. Once CGP participants had no active problems, they were assigned to the status of medically stable, a "surveillance status," and received routine monitoring calls every 30 to 90 days.

Program changes. A number of changes to the CGP occurred over the course of the demonstration as relayed to RTI staff during the second site visit. Noteworthy changes included the following:

- **Changes in engagement strategies for the original population.** Several modifications were made to the process used to engage the original population: the CGP prioritized mailings and telephone outreach to the highest risk patients first, offered a range of degrees of program participation (e.g., self-directed and active case management) to optimize program acceptance, and staggered mailings and outreach. CGP leaders felt that two factors contributed the most to the higher rate of engagement for the refresh cohort: better program recognition in the community and the choice of engagement levels.

- **Optimizing staff resources**. The CGP began engaging nonclinical staff in the assessment process, ongoing monitoring and retention activities. The extent of social problems among participating beneficiaries was significant and greater than expected. This required staff knowledgeable about community resources and availability of services.

- **Development of varying levels of program engagement.** To maximize participation in the program, the CGP created different levels of program engagement ranging from self-directed to active case management and high utilization.

- **Simplification of the assessment and documentation processes.** MMC terminated their contract with Life Plans and began conducting assessments in-house. Moreover, they simplified the process for documenting assessments and interventions with automatically generated notes to summarize information efficiently.

- **Modifications to the disease management program.** The CGP made several changes to MMC's CMO's traditional disease management model as they realized the following about their participants: (1) routine screening and evidence-based guidelines for diabetes and HF are not always applicable to CGP beneficiaries because of their high levels of frailty and cognitive impairment; (2) functional decline, cognitive limitations, co-morbidities, life expectancy, and quality of life needed to be considered; and (3) beneficiaries in the program struggled with information overload. Thus, the program was modified so that education focused on early detection of worsening symptoms. Motivational interviewing techniques were used to help improve compliance with treatment.

- **Hiring a dedicated provider for palliative care program.** A palliative care initiative began in June 2008 (after the initial site visit) and included a funded nurse practitioner position to work with the hospital-based Palliative Care Team. This team was credentialed at MMC and at least two high-volume SNFs in the area. The plan was for the nurse practitioner to follow participants across all settings—from inpatient to SNF or home. The goal of the palliative care initiative was to increase completion of and follow-through on advance directives and to facilitate more timely referrals to hospice.

- **Creation of the High Utilizers Team (HUT).** One of the important changes implemented in the CGP was the introduction of the HUT in January 2008. The CGP generated high-risk reports using Medicare claims and MMC census data. Based on admissions, these monthly reports categorized participants into high, medium, and low risk. Based on the information from the reports, case managers developed intervention plans. Most interventions for these participants included referral to either the House Calls Program or disease management, or assignment to the HUT. The HUT consisted of three team members (2 licensed practical nurses and 1 social worker) who had smaller case loads than regular case managers. The HUT contacted participants weekly by phone and followed them closely if they were hospitalized or admitted to SNF.

- **Use of telemonitoring with participants.** The CGP also tested telemonitoring using the Health Buddy® device with a subset of participants. CGP staff recognized telemonitoring's potential for yielding effective interactions in the frail and elderly population.

E.3 Key Findings

In this section, we present key findings based upon the 36 months of the CGP's operations with its original population and 24 months with its refresh population. Our findings are based on the experience of approximately 6,000 ill Medicare beneficiaries split across 4 groups for analysis purposes (original and refresh intervention and comparison groups) limiting statistical power somewhat to detect differences. CMS required RTI to analyze the original and refresh populations separately to be consistent with the financial reconciliation. Doing so allowed us to quantify intervention effects over time as MMC's CGP matured. One drawback to separate analyses of each group is the smaller samples available for statistical testing. Only 2,891 and 896 intervention beneficiaries were available for analysis in the original and refresh intervention groups and 1,785 and 868 comparison beneficiaries in the corresponding original and refresh comparison groups. Wide variation in beneficiary costs over time make precise estimates of program success difficult with such small samples. Key findings presented below are based on the resulting statistical tests at standard 95% confidence levels. To better understand the statistical power underlying RTI's analyses, in subsequent chapters we present detailed statistics including confidence intervals for quality of care and acute care utilization measures and a detectable threshold for cost savings, or the rate of savings that would allow us to reject the null hypothesis of no savings.

Six key findings on participation, intensity of engagement in the CGP, beneficiary satisfaction and experience with care, clinical quality, health outcomes, and financial outcomes have important policy implications for CMS and future disease management or care coordination efforts among Medicare FFS beneficiaries. The CMHCB demonstration program holds MMC financially responsible for financial savings but does not hold MMC financially responsible for quality of care improvements.

Key Finding #1: The CGP did not preferentially engage beneficiaries who were at highest risk of acute clinical deterioration as measured by the concurrent HCC score.

Of the CGP's original intervention beneficiaries, 75% verbally consented to participate in the CMHCB demonstration at some point during the intervention period; 81% of the refresh population agreed to participate. In spite of this fairly high level of participation, we find that beneficiaries from the original population with high baseline HCC scores or medium or high baseline PBPM costs more likely to be participants. Demonstration period health status as measured by the concurrent HCC score had no impact after controlling for baseline health status characteristics and demographics. This suggests that the CGP was able to engage the historically sicker Medicare beneficiaries but did not preferentially engage those with acute clinical deterioration as measured by the concurrent HCC score. None of these measures were statistically significant for the refresh population. However, only 32% of the refresh population, or 386 beneficiaries, were in the reference group making it difficult to determine statistically significant differences.

Key Finding #2: The CGP did not substantially affect beneficiary reported experience with care, level of physical activity, and self-reported physical health.

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 the CGP's intervention: helpfulness of discussions with their health care team and quality of communication with their health care team. Because we used the same survey questions for both the intervention and comparison groups, we did not ask specifically about the helpfulness of discussions with staff of the CGP. Thus, demonstration programs with a telephonic care management approach might be at a disadvantage because beneficiaries might not consider the telephonic care managers as being their health care providers. 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, only one statistically significant positive group difference was detected—members of the CGP's intervention group reported fewer limitations in their activities of daily living than those in the comparison group. This difference, however, was not reflected in another measure of physical health-PHC scores. We did not detect any statistically significant intervention effects on any measures of beneficiary's satisfaction and experience with care, nor on any of the self-management outcomes for the CGP.

Key Finding #3: We did not detect improvement in 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 Medicare 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 did not detect systematic improvement in quality of care among the original or refresh intervention beneficiaries. We generally observed negative trends in rates, not positive, due to either improvement of the rate or less of a decline in the rate of receipt among the comparison beneficiaries. However, we would like to note that claims data are likely to produce underestimates of the rate of influenza vaccination as they do not capture flu vaccines that people receive in pharmacies, supermarkets, senior centers, or city-funded health care centers because services received in those settings do not result in Medicare claims. Although we do not have self-reported rates of influenza vaccines for the comparison group beneficiaries, MMC's CGP reported to RTI that their survey of participants revealed over three-quarters reported having received a flu vaccine during each of the three years of the demonstration.

Key Finding #4: We did not detect systematic reductions in acute care utilization as measured by rate of hospitalization, ER visits, or 90-day readmissions nor did we detect any difference in the use of the Medicare hospice benefit between the intervention and comparison groups.

During the course of CGP, we observed 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. Out of 30 acute care utilization comparisons, we observe one statistically significant positive intervention effect; the rate of all-cause ER visits grows at a slower rate during the last 12 months of the demonstration within the original intervention group than within the comparison group, or -275 per 1,000 beneficiaries ($p=0.02$).

We do not detect any other statistically significant positive intervention effects; however, we do detect one statistically significant negative intervention effect. During months 7 to 18 of the CGP demonstration period, the percent of original intervention beneficiaries with an all-cause hospitalization increased from 38% to 41% while the percent of comparison beneficiaries with an all-cause hospitalization declined from 41% to 40% ($p=0.04$).

We did not detect any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rates of the Medicare hospice benefit or in mean or median number of days of hospice.

Key Finding #5: We did not detect a difference in the rate of or time to death between the original and refresh intervention and comparison beneficiaries.

We did not detect a statistically significant differential rate of mortality between the intervention and comparison groups of the original or refresh population. Over the 36-month demonstration period for the original population, 29% of beneficiaries in the intervention and comparison groups died. During the 24-month demonstration period for the refresh population, 17% of beneficiaries in the intervention and comparison groups died.

Key Finding #6: Medicare cost growth in the intervention group was not statistically different from the rate of growth in the comparison group.

No statistically significant savings were detected for the intervention in the original population. Costs rose -$13 slower in the original intervention group (0.6 % of comparison costs), but savings needed to exceed 8.2% to be considered statistically significant. The CGP, overall, performed better with its refresh population as gross savings averaged -$124 (5.5% of comparison monthly costs). This savings level, however, did not achieve statistical significance. Intervention and comparison groups were somewhat unbalanced in baseline characteristics. Intervention beneficiaries were less likely to be non-White or enrolled in Medicaid. However, controlling for these baseline imbalances in characteristics had no effect on our overall final conclusion of no statistically significant savings. For differences in 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.

E.4 Conclusion

Based on extensive quantitative analysis of performance using statistical tests at standard 95% confidence levels, we did not detect improvement in key processes of care, beneficiary self-reported experience with care, self-management, and functional status, increase in use of the Medicare hospice benefit, or decrease in mortality. We did detect one statistically significant positive intervention effect; the rate of all-cause ER visits grew at a slower rate during the last 12 months of the demonstration within the original intervention group than within the comparison group. However, we also detected one statistically significant negative intervention effect. During months 7 to 18 of the CGP demonstration period, the percent of original intervention beneficiaries with an all-cause hospitalization increased from 38% to 41% while the percent of comparison beneficiaries with an all-cause hospitalization declined from 41% to 40%.

Although PBPM costs rose slower in the original and refresh intervention groups relative to the comparison groups, statistically significant savings were not detected in the *overall* samples. This may have been due to relatively small sample sizes and lack of statistical power. 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 2,900 original and 900 refresh intervention beneficiaries and 1,800 original and 900 refresh comparison beneficiaries, we had limited power to detect significant savings. Gross savings had to be 8.2% in the original intervention population and 14.9% in the refresh intervention population to be considered significant at the 95% confidence level.

What might explain the lack of *overall* program effectiveness? One explanation may be the targeting of beneficiaries at greatest risk of intensive, costly, service use (as distinct from the need for general care management). MMC and CMS agreed upon a predicted costly set of Medicare beneficiaries for their intervention and comparison groups using 2004 claims data. Mean per beneficiary per month base year claims costs (weighted by fraction of time eligible for the intervention) were approximately $1,600 in both groups, a figure considerably higher than in the general Medicare population. During the intervention period, the comparison group

exhibited both rapidly rising costs ($792 in the original and $688 in the refresh group) as well as sizable regression-to-the-mean (RtoM) effects. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by nearly $2,000. Over the same time period, beneficiaries with costs over $4,000 saw their average costs decline by over $3,500. The large churning of beneficiaries in both the intervention and comparison groups from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater is the potential for RtoM, the greater is the challenge to identify lower cost, lower utilizing beneficiaries initially to avoid expensive hospitalizations in the near future. The "low cost" beneficiary problem was exacerbated by the more than one-year lag between selection and start date. Many originally high cost beneficiaries two years prior to start date became much lower cost one year prior to start date.

A second explanation may be the CGP's beneficiary recruitment strategy. Given the program's monthly management fee (roughly $120 per month) and the population-based financial risk feature of this demonstration, less than full engagement of the intervention population required the CGP to have been extremely successful in reducing costs associated with the participating beneficiaries. The CGP was not successful in broadly reducing hospitalizations during the demonstration period. The lack of substantive improvements in acute care utilization broadly across their intervention population translated into limited financial savings.

And, a third explanation may be the model of intervention itself. Prior evaluations of Medicare care management programs that were primarily telephonic have not demonstrated savings sufficient to cover fees similar to the CGP's fee. A cornerstone of the CGP was health coaching interactions with care manager nurses. However, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments, and the most dominant form of contact was telephonic.

Furthermore, CGP nurse care managers were not part of the beneficiaries' primary health care teams, thereby hindering their interactions with the beneficiaries' primary providers, changing medical care plans, or mitigating deterioration in health status. And, not all intervention beneficiaries had primary care physicians at MMC, therefore the care managers had to interact with nonMMC providers with whom they had little or no prior relationship.

CHAPTER 1
INTRODUCTION TO THE MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION AND MONTEFIORE MEDICAL CENTER'S (MMC) CARE GUIDANCE PROGRAM (CGP)

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 Montefiore Medical Center's (MMC) 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. 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
2. Care Level Management (CLM)
3. Massachusetts General Hospital (MGH) and Massachusetts General Physicians Organization (MGPO) and its Care Management Program (CMP)
4. Montefiore Medical Center (MMC) and its Care Guidance Program (CGP)
5. VillageHealth (formerly known as RMS) and its Key to Better Health program (KTBH)
6. Texas Tech University Health Sciences Center (TTUHSC) and its Texas Senior Trails (TST) program

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% safety margin, there is also a shared savings provision with CMS according to the following percentages:

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 CMO.
3. Savings in the >10%-20% range will be shared equally between CMO (50%) and CMS (50%).
4. Savings of >20% will be shared between CMO (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' 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 20 of the intervention period. MMC's CGP 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 MMC's CGP operations with its original population and 24 months with its refresh population. We start by reporting on the degree to which CGP 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 the CGP and participating beneficiaries from encounter data provided to RTI from the CGP. We then report findings related to the effectiveness of the CGP to improve beneficiary and provider satisfaction, improve

functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings.

1.2 MMC's CMHCB Demonstration Program Design Features

1.2.1 MMC's Organizational Characteristics

Montefiore Medical Center was one of six organizations providing care management support as part of the CMHCB demonstration coordinated by CMS. MMC is an integrated delivery system that provides patient care, conducts research, and serves as a teaching hospital for the Albert Einstein College of Medicine. MMC provides a full continuum of health care services (emergency, inpatient, outpatient, and post-acute care) primarily to residents of the Bronx and Westchester County, New York. Through its affiliation with Albert Einstein, MMC employs 800 house staff, trains 700 medical students, and provides continuing medical educational opportunities for its staff of 2,000 physicians. MMC has a partnership with the Montefiore Independent Practice Association (IPA), which is the only entity in New York State that is eligible to enter risk arrangements with health plans. The IPA directly manages a population of 100,000 people and an additional 50,000 individuals through indirect risk arrangements, including approximately two-thirds of the Medicaid Advantage population in the Bronx.

A corporate subsidiary of MMC, Care Management Organization (CMO), was established in 1996 as a managed services organization to contract with various IPAs to perform administrative functions and obtain and manage risk-bearing contracts. To fulfill its commitment to Montefiore's mission of improving health in the Bronx and Westchester, MMC's CMO supports provider-driven strategic initiatives derived from the assumption and management of risk, with medical management services, financial administration capabilities, information systems, and infrastructure to manage customer and provider relations.

During its 10-year history of providing care management services, MMC's CMO has observed that chronic illnesses that carry the opportunity for exacerbation coupled with the social isolation characteristic of the CMHCB population often lead to increased use of health care resources and poor patient outcomes. Further, the specific chronic diseases do not result in unique sets of care coordination issues; rather, the factors that lead to hospitalization are often psychosocial in nature and are relevant to a variety of chronic conditions. For example, noncompliance with medication regimens may be related to inability to pay for medications, inability to travel to a pharmacy, or lack of knowledge about pharmacies that deliver products—all issues that are not associated with a specific disease. Consequently, MMC's CMO developed the Care Guidance program (CGP) to help participants access the medical care and social services they need to maintain health and avoid unnecessary hospitalization, regardless of condition.

1.2.2 Market Characteristics

MMC primarily serves residents of Bronx, New York; therefore, MMC selected this area as the target region for its CMHCB program. A densely populated urban area of 42 square miles, the Bronx is home to more than 1.3 million people. The population is racially diverse with a high proportion of Hispanic residents (48%) and African Americans (31%), groups that

are at high risk for many chronic conditions (e.g., diabetes). Sixty percent of this population speaks a primary language that is not English.

Almost 10% of the population is aged 65 years or older and 21% of households include an elderly person. At the time of the initial site visit in 2007, more than three quarters of the Bronx Medicare population received health care services through the FFS benefit. A significant proportion of the elderly population in the Bronx (25%) earns an income below the federal poverty level, which is a higher level of poverty than that of New York City or the nation overall. In addition, 30% of the Medicare beneficiaries in this area are dually eligible for Medicaid. This vulnerable population has insufficient access to primary care services as evidenced by the high rates of hospital admissions for ambulatory care sensitive conditions.

Overall, the supply of acute care beds in the Bronx is relatively low, whereas the number of nursing home beds available is very high. As a result, hospitals have an incentive to discharge patients as quickly as possible and place them in nursing homes for additional care when needed. In this market, Medicare FFS beneficiaries with chronic illnesses are at particularly high risk for multiple hospital and nursing home admissions which drives up health care costs.

A large number of community-based services are available to Bronx residents to address health care problems and psychosocial issues. However, this collection of resources is not centrally organized and different programs are subject to variable funding cycles, therefore Medicare beneficiaries may not know about available services and/or have difficulty accessing these services.

1.2.3 MMC Intervention and Comparison Populations

Intervention population. MMC worked with its CMS project officer and analysts from ARC to develop a method for selecting the starting population for its CMHCB demonstration program. Inclusion criteria for eligibility in the CGP included:

- Medicare FFS beneficiaries with a primary residence in one of 16 designated ZIP codes in Bronx, New York surrounding MMC, with a high level of disease severity as indicated by Hierarchical Condition Categories (HCC) scores of 1.8 or greater.[3]

- Two visits to MMC physicians between January 1, 2004 and December 31, 2004, or one visit to MMC physicians in the 12-month claims period with no visits to other physicians, or a plurality of visits to MMC inpatient facilities, or one visit to an MMC inpatient facility and no visits to other inpatient facilities.

- Absence of selected conditions as indicated by ICD-9 diagnosis codes and DRG codes obtained from claims data, including dementia, substance abuse, and schizophrenia, among others.

The population was further restricted using the following exclusion criteria:

---

[3] MMC's original CMHCB proposal included a minimum HCC score of 2.0 for eligibility in the program; however, this criterion was reduced to a score of 1.8 or greater to obtain an intervention group of desired size.

- age less than 45,
- receiving the Medicare hospice benefit,
- receiving the Medicare end-stage renal disease (ESRD) benefit,
- history of dialysis treatment,
- enrolled in a Medicare Advantage (MA) plan,
- Medicare as a secondary payer, or
- no Medicare Part A or Part B coverage as of May 1, 2006.

Using these criteria, a total of 2,969 Medicare beneficiaries were assigned to the Care Guidance Program's original intervention group.

Beneficiaries who elected the ESRD benefit following the launch of the program became ineligible for the CGP intervention. Most of these beneficiaries entered an ESRD-specific case management program that provided services similar to those of the CGP; therefore continued participation in the CMHCB demonstration would be duplicative and potentially confusing to patients. MMC decided to retain individuals who elected hospice during the intervention period since the CGP included counseling participants regarding options for care at the end of life. Such interventions provided an opportunity for CGP staff to help participants receive appropriate care and potentially impact costs associated with care during this period.

Using 2006 Medicare claims data, a refresh intervention population was selected. The refresh intervention group selection method replicated the procedures used to define the original population. The refresh inclusion, exclusion, and loyalty criteria were the same as those specified for the original intervention population with one exception--the list of CPT and Place of Service codes was expanded to exclude more residents of SNFs and nursing homes. The final refresh intervention population size was 912 beneficiaries.

Comparison population. Following the development of the intervention group criteria, MMC 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 this CMHCB demonstration program. The comparison group was selected using the following eligibility criteria:

- Medicare FFS beneficiaries with a primary residence in 16 ZIP codes in Brooklyn and Manhattan surrounding 5 comparison hospitals with household income levels and proportions of Hispanic residents similar to the intervention ZIP codes with a high level of disease severity as indicated by HCC scores of 1.8 or greater.
- A plurality of visits to at least 1 of 19 physician group practices (identified by tax identification number), 1 visit to a comparison group practice and no visits to any other physicians, or a plurality of admissions to 1 of 5 inpatient facilities or 1 admission to a comparison hospital and no admissions to any other hospitals.

- Absence of selected conditions as indicated by ICD-9 diagnosis codes and DRG codes obtained from claims data, including dementia, substance abuse, and schizophrenia.

The exclusion criteria that were applied to the intervention group were also used to limit the comparison group (i.e., age less than 45, receiving the Medicare hospice benefit, receiving the Medicare ESRD benefit, history of dialysis, enrolled in an MA plan, Medicare as a secondary payer, or lack Medicare Part A or Part B coverage as of May 1, 2006).

In order to ensure that the comparison group had Medicare costs similar to the intervention group, the comparison group members were randomly selected from each of five cost strata representing the cost quintiles observed in the intervention population. The number of comparison beneficiaries selected from each stratum was determined by the number of intervention beneficiaries in each stratum. The final initial comparison group size was 1,837 beneficiaries.

The refresh comparison group selection method replicated the procedures used to define the original refresh group. The refresh inclusion, exclusion, and loyalty criteria were the same as those specified for the original intervention and refresh populations with the addition of more CPT codes to exclude additional SNF and nursing home residents. Eligible refresh beneficiaries were randomly selected from five cost strata as determined by the number of intervention beneficiaries in each stratum. A total of 887 beneficiaries were selected for the comparison refresh population.

1.2.4 Care Guidance Program Operations

The CGP was launched on June 1, 2006. MMC negotiated a management fee of $120 for the original intervention group during the first year and $123.84 in years 2 and 3 ($0.00 fee in the last two months for all intervention beneficiaries and no fee for intervention beneficiaries in the self-directed program as of September 2008). Fees for the refresh intervention group were $123.84 for all 24 months ($0.00 fee in the last two months and no fee for intervention beneficiaries in the self-directed program as of September 2008). At the end of the 3-year period, MMC was obligated to achieve a 5% savings in Medicare payments among the intervention group (regardless of participation in the CGP) compared to the comparison group, and to cover program fees collected. In addition, CMO had the opportunity to share a portion of any savings beyond 5% that were achieved. Specifically, 100% of savings between 5% and 10% were to be paid to CMO, savings between 10% and 20% were to be shared equally among CMS and CMO, and savings over 20% would be shared with 70% going to CMO and 30% retained by CMS.

The CMHCB program allowed MMC to fulfill the following goals: (1) apply lessons learned about chronic disease management from managing risk contracts, (2) demonstrate the value of its Care Guidance Program based on cost savings to Medicare and improved patient outcomes, and (3) expand upon the services that it provides to the Bronx community.

The CMHCB demonstration provided MMC with the opportunity to develop the CGP, which built on the organization's knowledge of the population and resources in the Bronx, as well as its experience coordinating the care of Medicare beneficiaries participating in managed

care plans. Participation in the CMHCB demonstration gave MMC's CMO the opportunity to demonstrate the value of the care coordination services provided by the CGP, in particular, the program's scalability and effectiveness for MMC patients as well as patients who received care outside of the MMC system.

The demonstration presented MMC with an additional way to serve its community and conduct research, both key elements of the organization's mission. From a financially strategic perspective, the CGP had the potential to increase MMC's market share and build patient loyalty to MMC by helping patients access needed services in a timely way to avoid acute exacerbations of disease.

1.2.5 Overview of the CGP Demonstration Program

RTI conducted two site visits to the Montefiore Medical Center in Yonkers, NY, where the corporate offices of MMC's CMO are located. The first site visit was conducted in January 2007, seven months after the launch of their CMHCB demonstration program. The site visit, one of several evaluation components, was designed to focus on implementation: understanding the services offered by the CGP and reporting early experiences with program implementation and engagement of eligible beneficiaries, providers, and CMS. The second site visit, 25 months into the demonstration, focused on CGP staff's impressions and interpretation of its 25-month experience in working on the demonstration program. During the follow-up visit, RTI staff met with MMC's CMO senior management, key program staff, and physicians affiliated with the CGP. The protocol to conduct the follow-up interviews included a range of questions related to:

- Program implementation,
- Program monitoring/outcomes to date, and
- Implementation experience/lessons learned to date.

The description of the CGP and its activities in this report reflects MMC'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 CGP participants as well as the clinical protocols/analytic tools to support CGP nurse care managers and other health professionals who deliver these services. Second, we discuss program changes and enhancement activities that occurred as the program evolved.

Participant Support Services. The CGP was a complex case management program designed for the frail elderly population and disabled adults that was supported with technology (e.g., sophisticated information systems and telemonitoring) and MMC's established relationships with providers and community organizations. The CGP used a holistic approach to address the full complement of medical, psychological, and socioeconomic problems of the target population. Each program participant received interventions tailored to his or her specific needs. By partnering with program participants, their families, caregivers, and the medical community, the CGP aimed to help physicians manage high-risk patients, reduce medication complications, emergency room visits, and avoidable hospitalizations, improve the quality of life for both the participants and their caregivers, and support participants in the community.

The core of the CGP consisted of one-on-one telephone calls between participants and care managers, who linked beneficiaries with needed medical and social services. The program provided the following specific services to participants: care coordination, clinical pharmacist review, link to community support services, nutritional monitoring and counseling, psychosocial support, life care planning, and disease management and telemonitoring. Beneficiaries could participate in any or all of the program elements during the demonstration program, depending on their needs throughout the period. Participants were assigned to a care team based on the location of their residence and their primary language. In this way, each beneficiary interacted with two or three people throughout the program period. Each care team used the following basic strategy to support participants: assess participant problems and resources and develop care plan to address identified needs, implement and deliver interventions to address participant problems, and re-assess on a regular basis and adjust care plans based on changes in participant problems and resources.

Physician Support Services. The CGP helped physicians manage their high-risk patients by providing information about patient health status and supporting patients in their efforts to comply with provider treatment plans. The CGP provided physicians with a summary of results from its comprehensive baseline assessment and informed physicians about their patients' conditions between office visits. In addition, the program's clinical staff educated patients about their conditions and prescribed treatment plans and helped participants manage psychosocial stressors that may have impacted their ability to adhere to treatment plans. Rather than request extra services from physicians, the CGP collaborated with physicians to care for their patients. As a result, the program did not provide financial incentives to participants' physicians.

Staffing. MMC's strategy to staffing its CGP addressed the high intensity of work required during the first months of the program to enroll beneficiaries and to deliver the program using resources as efficiently as possible. MMC enlisted the support of temporary staff with customer service experience to conduct outreach telephone calls to enroll beneficiaries in the CGP. The program was delivered by 28 staff members: 12 full-time clinical staff and 16 full and part-time administrative staff. MMC staffed its CGP with registered nurses (RNs), licensed practical nurses (LPNs), social workers with a Master's degree (MSWs), and patient educators with a bachelor's degree and community-based experience working with the elderly. Social workers and patient educators were included in the staffing model to address the high number of psychosocial issues faced by the participants. Project staff also included a physician and a geriatric psychiatrist. Patient educators were well suited to linking participants with community resources, as well as working with the larger than expected proportion of the population that was "self-directed." The CGP provided support to these individuals through weekly meetings with the clinical staff and through supervision provided by RN team leaders. The clinical staff was divided into three teams, each led by a registered nurse with support from patient educators. Two teams were assigned to beneficiaries in specific geographic areas defined by ZIP code, while the third team supported all Spanish-speaking beneficiaries.

Beneficiary outreach/engagement. MMC used a variety of strategies to conduct outreach to the beneficiary population to solicit participation in the CGP, including development of a brand and logo, mailings, outbound telephone calls, and in-person interaction with hospitalized patients. CGP leadership staff conducted focus groups with beneficiaries to inform

the schedule of marketing activities. The outreach process was conducted in 3 waves organized alphabetically by last name to help staff manage the work of contacting the 2,969 beneficiaries eligible for the program. Because all eligible beneficiaries were at high risk for acute medical events as indicated by having an HCC score of 1.8 or higher, MMC chose not to use level of health risk to prioritize the waves of outreach activities.

Provide outreach/participation. Under the leadership of one of the CGP's medical directors (Dr. Reynolds), a strategy was developed to inform providers about the program and enlist their support using CGP staff's knowledge of the MMC system. The CGP used CMS claims data and MMC information systems to match beneficiaries to their primary care and specialty providers and stratified physicians according to patient volume to prioritize outreach efforts. They conducted a mail campaign and in-person presentations to introduce providers to the CGP. Overall, physicians were aware of the benefits of interdisciplinary care and welcomed the help provided by the CGP, since they were not able to provide all desired patient support services in their practices.

Risk Assessment. The CGP developed a comprehensive, 35-page assessment tool to enable care managers to identify the specific problems faced by each participant and implement interventions that effectively target these issues. They contracted with LifePlans, a long-term risk management company, to do 1,000 assessments, with additional assessments conducted by CGP clinical staff. The decision to outsource risk assessment was driven by the need to assess a large number of beneficiaries as soon as possible upon engagement with the program. In contrast to CGP staff, LifePlans tried to conduct the full assessment during one phone call, rather than two or more. MMC planned for LifePlans to conduct 60% of all assessments; however the organization succeeded in completing assessments for only 40% of participants. Further, CGP staff found that they needed to conduct follow-up calls with participants to validate the problems identified during LifePlans assessments.

Delivery of Interventions. For each participant, the CGP information system developed a problem list based on responses to the assessment. Clinical staff developed a customized care plan by assigning at least one intervention to each problem identified. Notably, the program focused on those problems that could benefit from an actionable intervention. The beneficiaries receiving case management support were stratified to help determine the level and types of services they received from the CGP. The CGP was designed to contact participants at least one time per month during the period when they were actively receiving interventions.

- **Facilitation of access to and coordination of care.** The CGP provided a variety of interventions to ensure participant access to primary care services, including referrals to primary care providers, discharge planning services, referrals to home care services, and fall prevention assessments and referrals.

- **Medication management.** The CGP provided medication management support to participants, including clinical pharmacist review of medication regimens, patient education about medication, support to address issues of noncompliance, information about the Medicare Part D benefit, and referrals to resources to assist with payments for medicine.

- **Facilitation of access to community support services.** CGP staff helped participants access a variety of community services to help them maintain their independence and quality of life (e.g., Meals on Wheels, transportation services, community-based case management services, adult day care, senior centers, caregiver support programs, and extended in-home services for the elderly population).

- **Nutritional monitoring and counseling.** Clinical staff monitored patients for issues related to proper nutrition (e.g., malnutrition, obesity, and special diets to manage chronic conditions).

- **Psychosocial support.** The CGP provided support to participants to address social isolation and depression, issues common to the frail elderly population.

- **Life care planning.** Case managers provided education and support to help participants establish advanced directives, and consider options for palliative and hospice care as needed.

- **Disease management and telemonitoring.** MMC's CMO developed disease management programs for diabetes, heart failure (HF), and chronic obstructive pulmonary disease (COPD). The CGP provided disease management services to beneficiaries with one or more of these conditions and no other major or complex medical needs.

- **Reassess and adjust.** Routine reassessment of participant issues and adjustment of interventions delivered on a routine basis occurred at least once per quarter. CGP staff collected updated clinical information, reviewed problems that had not yet been resolved and interventions conducted to date, and identified new problems and interventions required. Once participants had no active problems, they were assigned to the status of medically stable, a "surveillance status," and received routine monitoring calls every 30 to 90 days.

Program changes. A number of changes to the CGP occurred over the course of the demonstration as relayed to RTI staff during the second site visit. Noteworthy changes included the following:

- **Changes in engagement strategies for the original population.** Several modifications were made to the process used to engage the original population: CGP staff prioritized mailings and telephone outreach to the highest risk patients first, offered a range of degrees of program participation (e.g., self-directed and active case management) to optimize program acceptance, and staggered mailings and outreach. CGP leaders felt that two factors contributed the most to the higher rate of engagement for the refresh cohort: better program recognition in the community and the choice of engagement levels.

- **Optimizing staff resources**. The CGP began engaging nonclinical staff in the assessment process, ongoing monitoring, and retention activities. The extent of social problems among participating beneficiaries was significant and greater than expected.

This required staff knowledgeable about community resources and availability of services.

- **Development of varying levels of program engagement.** To maximize participation in the program, the CGP created varying levels of program engagement ranging from self-directed to active case management and high utilization.

- **Simplification of the assessment and documentation processes.** The CGP terminated their contract with Life Plans and began conducting assessments in-house. Moreover, they simplified the process for documentation of assessments and interventions with automatically generated notes to efficiently summarize information.

- **Modifications to the disease management program.** The CGP made several changes to MMC's CMO's traditional disease management model as they realized the following about their participants: (1) routine screening and evidence-based guidelines for diabetes and HF were not always applicable to CGP beneficiaries because of their high levels of frailty and cognitive impairment; (2) functional decline, cognitive limitations, co-morbidities, life expectancy, and quality of life needed to be considered; and (3) beneficiaries in the program struggled with information overload. Thus, the disease management program was modified so that education focused on early detection of worsening symptoms. Motivational interviewing techniques were used to help improve compliance with treatment.

- **Hiring a dedicated provider for palliative care program.** A palliative care initiative began in June 2008 (after the initial site visit) and included a funded nurse practitioner position to work with the hospital-based palliative care team. This team was credentialed at MMC and at least two high-volume SNFs in the area. The plan was for the nurse practitioner to follow participants across all settings—from inpatient to SNF or home. The goal of the palliative care initiative was to increase completion of and follow-through on advance directives and to facilitate more timely referrals to hospice.

- **Creation of the High Utilizers Team (HUT).** One of the important changes implemented in the CGP was the introduction of the HUT in January 2008. The CGP generated high-risk reports using Medicare claims and MMC census data. Based on admissions, these monthly reports categorized participants into high, medium, and low risk. Based on the information from the reports, case managers developed intervention plans. Most interventions for these participants included referral to either the house calls program or disease management, or assignment to the HUT. The HUT consisted of three team members (two licensed practical nurses and a social worker) who had smaller case loads than regular case managers. The HUT contacted participants weekly by phone and followed them closely if they were hospitalized or admitted to SNF.

- **Use of telemonitoring with participants.** The CGP also trialed telemonitoring using the Health Buddy® device with a subset of participants. CGP staff recognized the

potential of telemonitoring for yielding effective interactions in the frail and elderly population.

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 MMC's CGP demonstration period and provider satisfaction with the CGP culled from interviews with physicians during the site visit. In ***Chapter 4***, we provide the results of our analyses of participation levels in the CGP and level of intervention with participating beneficiaries (i.e., the number of reported interventions and/or assessments). 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. Supplements to ***Chapters 2, 4, and 7*** are available from the CMS Project Officer upon request.

This page intentionally left blank

CHAPTER 2
EVALUATION DESIGN AND DATA

2.1 Overview of Evaluation Design

2.1.1 Gaps in Quality of Care for the 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 CMHCB demonstration has been designed to address current failings of the health care system for chronically ill Medicare FFS beneficiaries.

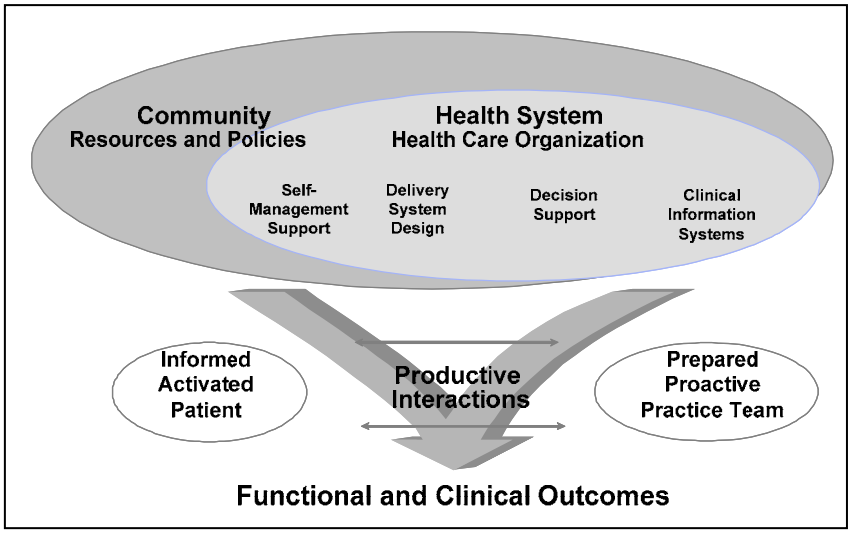
2.1.2 Emerging Approaches to Chronic Care

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).[4] 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

[4] 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.

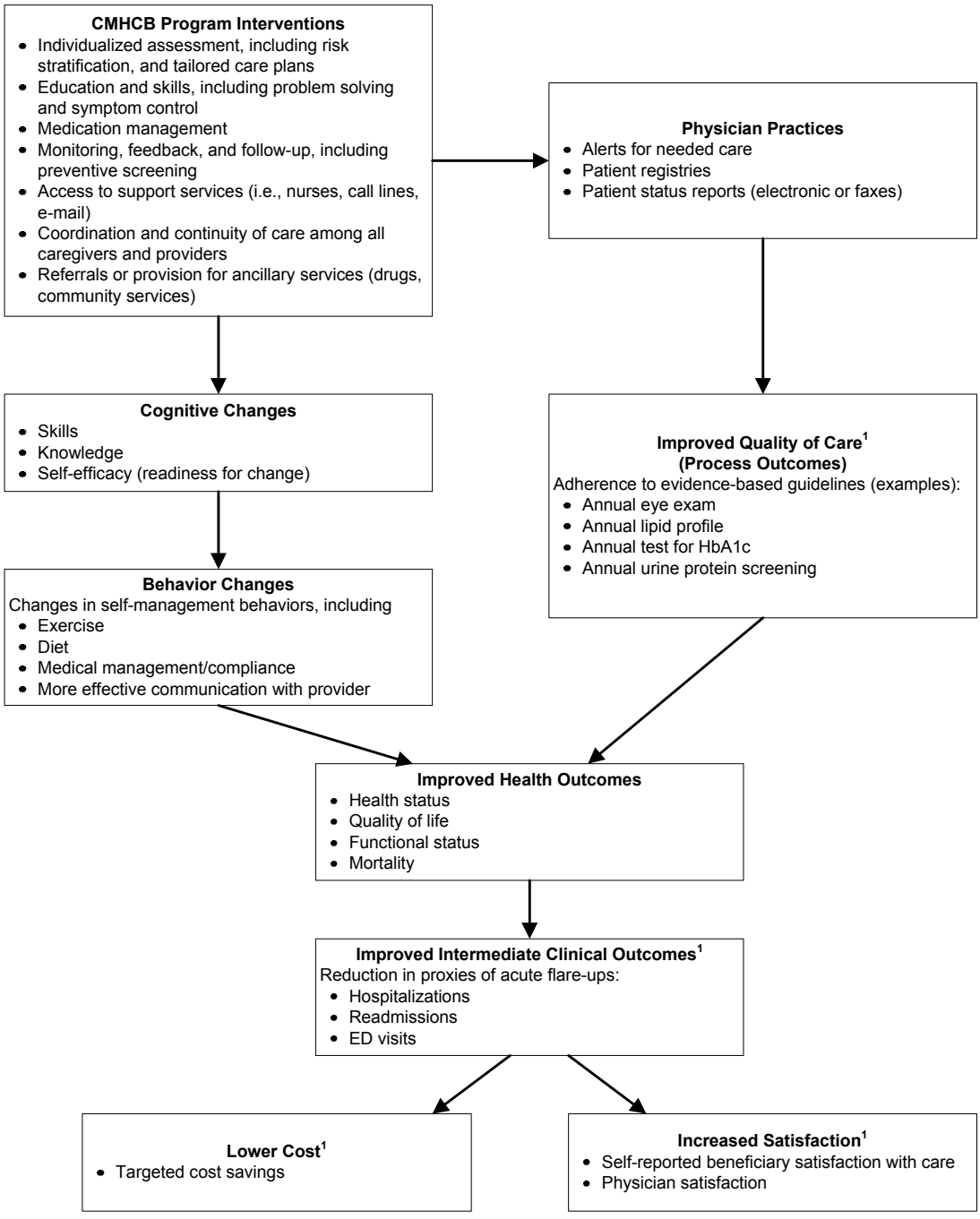
- *Monitoring, feedback, and follow-up*. Activities in this domain include ongoing biomonitoring of beneficiaries by placing scales or other equipment in their homes or by having the beneficiaries self-report their weights, blood sugars, or other measures. When data on preventive services, screenings, 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 care 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 CGP was able to engage its randomized 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. The CGP's implementation experience was 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 MMC able to implement its CGP?				
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 the CGP?	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 the CGP engage physicians and physician practices in their programs?	Yes	No	No	No
REACH: How well did the CGP 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 the CGP 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 the CGP?	No	Yes	Yes	No
4. To what extent were the intended audiences exposed to the CGP 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 the CGP intervention versus a low level of intervention?	No	Yes	Yes	No
EFFECTIVENESS: To what degree was the CGP 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 the CGP 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 the CGP 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 the CGP result in greater engagement in health behaviors?	No	No	No	Yes
3. Did the CGP 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 the CGP improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?	No	No	Yes	No
2. Did the CGP improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?	No	No	Yes	No
3. Did the CGP 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 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 the CGP 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: CGP = Care Guidance Program; CMS = Centers for Medicare & Medicaid Services; ER = emergency room.

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 prespecified 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

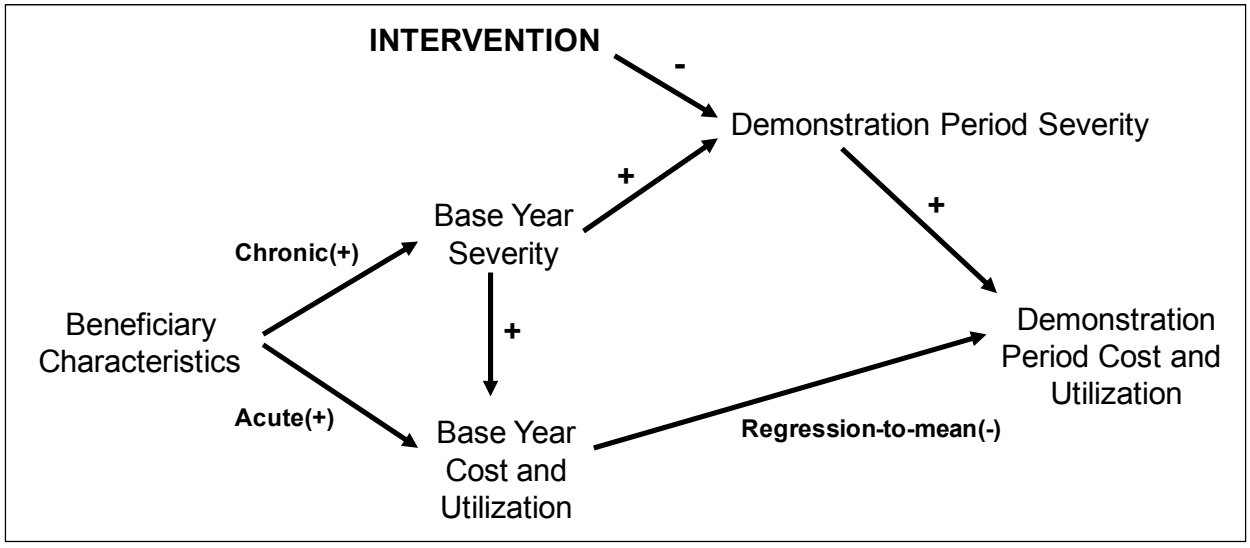
ε = a regression error term.

This model uses three sets of variables in analysis of covariance (ANCOVA) format to capture differences between intervention and comparison beneficiaries. The β_1 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, β_1 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 β_1 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 the CGP. As noted in Chapter 1, we also conducted a survey of CGP beneficiaries to assess their satisfaction with the CMHCB demonstration program and semi-structured interviews with a small number of physicians to assess their awareness of and satisfaction with the CGP. 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 MMC's CGP and was submitted to ARC. This file was updated quarterly and logged status changes among the intervention groups by CGP staff. 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 CGP beneficiary was assigned—intervention or comparison—for both the original and refresh populations.

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

1. ARC provided RTI with an EDB file for the CGP comprised of all original and refresh beneficiaries. RTI used this file to determine daily eligibility based on CGP 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 the CGP 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.*
2. RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (May 1, 2006) for the CGP original population.
3. RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (May 10, 2007) for the CGP 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 the CGP 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, the CGP sent RTI beneficiary-level intervention files that provided information on intervention and assessment activities. Intervention activities included referrals and mailings, while assessments included comprehensive baseline, disease-specific baseline, PHQ-9 for depression, and use of the Health Buddy® device. 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 the CGP

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 the CGP. 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: CGP = Care Guidance Program; ESRD = end-stage renal disease; MA = Medicare Advantage; GHO = Group Health Organization.

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

Table 2-3
Analysis periods used in the CGP 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				
6/1/06	5/31/09	36	6/1/05	5/31/06
Refresh Population				
6/1/07	5/31/09	24	6/1/06	5/31/07

NOTES: CGP = Care Guidance 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 selection, May 1, 2006 for the original population and May 10, 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 the CGP was active in the demonstration. The numerator is the number of days the beneficiary is eligible during that time period. MMC participated in the demonstration for the full 36 months, so the number of days in the denominator for each original population beneficiary in the CGP is 1,096 (the CGP end date minus the CGP start date + 1). If a beneficiary died 420 days into the intervention period, the eligibility fraction for the participation analysis 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 the CGP. 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

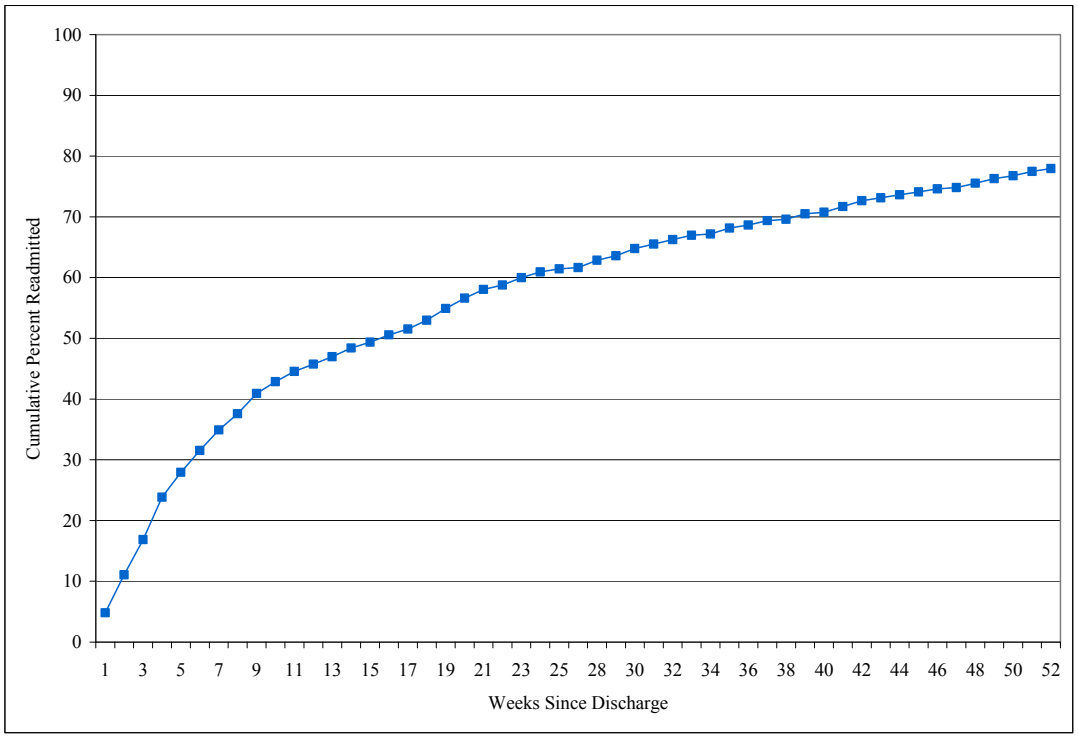
   - emergency room visits, including observation bed stays.

   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 the CGP's 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[5].

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



We examined readmissions following admissions that occurred during two 12-month periods for the original population and one 12-month period for the refresh

5 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.

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 per beneficiary per month (PBPM) cost 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 FFS population, influenza vaccination. Second, we selected several measures that are specific to beneficiaries with diabetes and heart failure as these populations are prevalent in the CGP population. 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 – diabetes
- Rate of low-density lipoprotein cholesterol (LDL-C) testing – diabetes
- Rate of low-density lipoprotein cholesterol (LDL-C) testing – 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 in the CGP. 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 the CGP in case the services occurred after exposure to the CMHCB demonstration intervention and during the intervention period. One could envision that the CGP 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 MMC, we constructed counts of the number of months participants received interventions, the total number of interventions per beneficiary, and the average number of interventions per quarter that beneficiaries were participants. We also constructed similar measures for total number of assessments and by assessment category.

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 promoting 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[6].

The CGP at MMC included four major components: (1) facilitating access to and coordination of care by providing access to MMC medical health care professionals, facilitating communication between providers and the care guidance team, providing care for depression when appropriate, and providing money for transportation to health care providers; (2) implementing chronic care management by ensuring that care delivered adheres to guidelines developed by the CGP, providing education materials for participants and family/caregivers, and providing telemonitoring equipment to monitor weight, other biomarkers, and symptoms; (3) implementing community-based palliative care by ensuring timely referral to palliative care; and (4) implementing medication noncompliance/polypharmacy review whereby a pharmacist supports patients and providers and MMC partners with a large community pharmacy to deliver prescriptions and support enrollment in the New York state drug benefit.

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 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).

---

6 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 CGP 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 CGP intervention on the beneficiaries' assessment of satisfaction that their full health care team is helping them to cope with their chronic conditions.

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. Because we used the same survey questions for both the intervention and comparison groups, we did not ask specifically about the helpfulness of discussions with staff of the CGP. Thus, demonstration programs with a telephonic care management approach might be at a disadvantage because beneficiaries might not consider the telephonic care managers as being their health care providers. 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 CGP 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 response propensity analysis to identify factors that influenced whether a beneficiary responded to the survey and included the above listed characteristics among others. The response propensity analysis results were used to derive survey sample weights so that analysis results would be representative of the original intervention and comparison beneficiaries.

We then conducted a series of statistical analyses to explore intervention-comparison differences and CMHCB 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 of the CGP'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 = control or 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 the MMC program.

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 an MA 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 CGP'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 the CGP was 62.8%. 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 the CGP

This section presents the results of the Medicare Health Services Survey data analysis for the CGP. 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 the CGP.

Table 3-1
Medicare Health Services Survey: Estimated intervention effects for experience and satisfaction with care, CGP
(N = 508)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Helping to cope with a chronic condition (1 to 5)	3.54	3.58	0.09	N/S
Number of helpful discussion topics (0 to 5)	2.05	2.44	-0.11	N/S
Discussing treatment choices (1 to 4)	3.16	3.01	0.18	N/S
Communicating with providers (0 to 100)	75.8	74.7	3.30	N/S
Getting answers to questions quickly (0 to 100)	65.6	60.7	5.20	N/S
Multimorbidity Hassles score (0 to 24)	2.98	3.57	-0.60	N/S

NOTES: CGP = Care Guidance 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

Overall experience and satisfaction with care—The average score for the key satisfaction outcome item that assessed how well the health care team helped beneficiaries cope with their illness was 3.5 for the intervention group, or about midway between "very good" and "good" ratings. The average score for the comparison group was about 3.6. Over fifty four percent of CGP beneficiaries rated their experience as "excellent" or "very good" and about 27% 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. Across the six measures of experience and satisfaction with care, we observe no statistically significant positive intervention effects for the CGP.

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 the CGP.

Table 3-2
Medicare Health Services Survey: Estimated intervention effects, self-management, CGP (N = 508)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Percent receiving help setting goals	61.0	69.5	0.30	N/S
Percent receiving help making a care plan	55.3	60.5	2.80	N/S
Self-efficacy ratings				
Take all medications (1 to 5)	4.37	4.17	0.17	N/S
Plan meals and snacks (1 to 5)	4.03	3.78	0.23	N/S
Exercise 2 or 3 times weekly (1 to 5)	3.47	3.15	0.29	N/S
Self-care activities				
Prescribed medications taken (mean # of days)	6.72	6.61	0.04	N/S
Followed healthy eating plan (mean # of days)	5.32	5.22	0.09	N/S
30 minutes of continuous physical activity (mean # of days)	3.26	2.92	0.56	N/S

NOTES: CGP = Care Guidance 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

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 CGP beneficiaries, in the intervention group 61% received help setting goals and 55.3% received help making a care plan. In the comparison group, 69.5% and 60.5%, respectively, received assistance on these self-management activities. The ANCOVA results reveal these differences are not statistically significant: the CGP did not have a higher proportion of intervention beneficiaries who report receiving help setting goals for self-care management, nor was the CGP associated with an increase in the number of intervention beneficiaries receiving help in making care plans.

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, CGP 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, on the scale of 1 to 5, CGP beneficiaries in the intervention group rated their confidence in taking medications as often as prescribed as 4.4, compared to 4.2 in the comparison group. Confidence in planning meals and snacks was rated 4.0 and 3.8, respectively, and confidence in exercising was rated as

3.5 and 3.2, respectively. However, none of the three ANCOVA effects were statistically significant.

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 CGP beneficiaries said they take their medications as prescribed ranged from 6.6 to 6.7; the mean number of days that CGP beneficiaries reported following a healthy eating plan was between 5.2 to 5.3, and the mean number of days CGP beneficiaries reported exercising was 2.9-3.3 days out of 7. ANCOVA analysis revealed that there were no statistically significant differences in the ratings for any of the three self-care activities between intervention and comparison group.

Physical and mental function—***Table 3-3*** displays the mental and physical functioning outcomes for the CGP. On average, CGP respondents had the mean PHC score for the intervention group of 30, slightly higher when compared to 28.4 for the comparison group but not statistically significantly so according to the ANCOVA adjustment. The mean MHC score for the intervention group was 37.3 and the PHQ-2 score of 2.26, compared to 35.9 and 2.55 for the comparison group, also not statistically significant intervention effects.

Table 3-3
Medicare Health Services Survey: Estimated intervention effects, physical and mental health function, CGP
(N = 508)

Outcome	Intervention mean	Comparison mean	ANCOVA-adjusted intervention effect	Stat. sig.
PHC score (physical health, mean =50, std=10)	30.0	28.4	1.8	N/S
MHC score (mental health, mean =50, std=10)	37.3	35.9	1.3	N/S
PHQ-2 score (depression, 0 to 6)	2.26	2.55	-0.07	N/S
Number of ADLs difficult to do (0 to 6)	2.40	3.15	-0.61	**
Number ADLs receiving help (0 to 6)	1.49	2.09	-0.23	N/S

NOTES: CGP = Care Guidance 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

Activities of daily living—On average, respondents in the CGP intervention group reported being limited on 2.4 ADLs compared to 3.2 ADLs for the comparison group, a difference that is statistically significant. CGP beneficiaries in the intervention group also reported receiving help with an average of 1.5 ADLs, compared to 2.1 ADLs in the comparison group, the intervention effect that is consistent in direction but not statistically significant. In summary, ANCOVA results indicate that there were fewer reported ADL limitations in the intervention group.

3.1.4 Conclusions

The CGP 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. The CGP included four major components: (1) facilitating access to and coordination of care, (2) implementing chronic care management, (3) implementing community-based palliative care, and (4) implementing medication noncompliance/polypharmacy review. The CGP demonstration hypothesized that 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.

Among the 19 outcomes covered by the survey, only one statistically significant positive group difference was detected—members of the CGP intervention group reported fewer limitations in their activities of daily living than those in the comparison group. This difference, however, was not reflected in another measure of physical health-PHC scores. We did not detect any statistically significant intervention effects on any measures of beneficiary's satisfaction and experience with care, nor on any of the self-management outcomes for the CGP.

3.2 Physician Satisfaction

RTI made two site visits to meet with CGP staff during the demonstration period. During the visits, RTI evaluators spoke in person with a limited number of physicians affiliated with beneficiaries assigned to the intervention group of the CGP. Physicians participating in a focus group reported a great rapport with CGP staff and appreciated the support that case management brought for their practices. They saw this program as a preventive medicine practice; working with patients to keep them at home and working to see improvements in health status and quality of life, the physicians felt that, with time, the CGP efforts would pay off. In particular, physicians welcomed CGP staff's help with taking care of the multitude of social issues that their patients encountered—something that physicians did not have time for in their practice. If, for example, an elderly homebound patient qualified for a free air-conditioning unit and required a letter from a doctor to obtain the unit, a CGP case manager would make sure that the letter was drafted, signed, and sent. Although this was a small intervention, it was one that made a huge difference in a patient's life.

Physicians also welcomed reports from case managers about what happened in patients' homes—with time constraints during office visits, physicians did not have the time to discuss some of these issues. They especially appreciated e-mails and feedback they received about

hospitalizations, medication issues, risks of falls or actual falls that patients did not report, depression, lack of advance directions, and so forth. The acceptance of the program was more favorable in the Montefiore system. Outside of MMC, the reception was more mixed, and physicians required more time to get to know the program and appreciate it.

CHAPTER 4
PARTICIPATION RATES IN MMC'S GCP AND LEVEL OF INTERVENTION

4.1 Introduction

Our participation analysis is designed to critically evaluate the level of engagement by the CGP in this population-based demonstration program 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 CGP 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 the CGP engage the 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 CGP 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 CGP?

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

- What beneficiary characteristics predict a high level of CGP 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. During our January 2007 site visit, CGP staff reported that they had engaged 56% of their intervention population, but they also reported that 11% had become ineligible (Brody and Bernard, 2007). As of July 2008, CGP staff reported that 625 refresh population beneficiaries had agreed to participate in the CGP which equates to an approximately 70% participation rate (Khatutsky, McCall, and Bernard, 2009). 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 the CGP and its assigned beneficiaries. The CGP intervention had assessment elements (care coordination, clinical pharmacist review, link to community support services, nutritional monitoring and counseling, psychosocial support, life care planning, and disease management and telemonitoring).

Therefore, we examine the number of telephonic and assessment contacts between CGP staff and their participants. For each participating beneficiary, the CGP provided RTI with a count of the number of interventions by type: provider or program referrals, the CGP interventions, pharmacist review, nutritionist review/referral, and patient education/mailing (general, disease management, and life planning). The CGP also provided information on assessments (e.g., comprehensive baseline, falls risk inventory, routine follow-up, flu/pneumovax, PHQ-9, Health Buddy enrollment/disenrollment).

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 the CGP demonstration. 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 the CGP) and the percentage of beneficiaries participating for more than 75% of their eligible months.[7] These latter two sets of numbers provide an estimate of the number of beneficiaries with whom CGP 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 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

[7] 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, gets a new primary payer (i.e., Medicare becomes secondary payer), receives hospice care, or dies.

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 enrollment database (EDB) and determined as of the date of assignment for the original population (May 1, 2006) and the refresh population (May 10, 2007):

- male, a dichotomous variable, set at 1 for males;
- non-White, a dichotomous variable, set at 1 for beneficiaries whose race code is not White.
- 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 3 (medium) and 4 or greater than (high); Charlson score of less than 3 is the reference group for the original population. For the refresh population, baseline Charlson scores of 2, 3 or 4 were medium and 5 or greater were in the high group. The reference group was a score of less than 2.
- baseline costs PBPM 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 CGP's original demonstration program was greater than or equal to $355.83 and less than $1,407 (medium) and $1,407 or greater (high); PBPM cost less than $355.83 is the reference group for the original population. For the refresh population, baseline PBPM costs greater than or equal to $245 and less than $1,439 were assigned to the

medium group and $1,439 or greater to the high category; PBPM cost less than $245 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 0.631 but less than 1.629 (medium) and greater than or equal to 1.629 (high); concurrent HCC score less than or equal to 0.631 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 0.606 but less than 1.493 and values greater than or equal to 1.493 were assigned to the high category; a concurrent HCC score less than or equal to 0.606 is the reference group.

4.2.2 Level of Intervention Analysis Methods

The CGP provided RTI with the number and nature of interventions and assessments for participating beneficiaries at the beneficiary level for the full CMHCB demonstration. We use these data to develop estimates of the level of intervention provided to CGP participants. The core of the CGP consisted of one-on-one telephone calls between participants and care managers, who linked beneficiaries with needed medical and social services. The CGP collected telephone contact data; however, early in the demonstration project CMS decided not to have it reported on a member level on a regular basis. Total number of contacts per month was reported to CMS. By the time we conducted our analysis, individual-level data on telephone contacts was no longer available as the CGP had been required to destroy all data files at the conclusion of their demonstration project. Therefore this important information was not available to us for analysis. Intervention information included in our analyses are primarily referrals and patient education with mailings.

Using the encounter data submitted by the CGP, we constructed counts of the number of interventions and assessments with participants. We also report the mean and median number of total contacts (as defined by interventions and assessments) 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 the CGP and 0 for beneficiaries who had a low level of contact. Beneficiaries who had a medium level of contact with the CGP 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 these same time periods.

We included these two additional demonstration period intervention variables because CGP 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.

We report levels of intervention with the original and refresh intervention populations during months 7 to the end of the demonstration. Because beneficiaries could have intermittent periods of eligibility and participation, we restricted inclusion in this analysis to beneficiaries who were eligible for and participating in the CGP for each month during this time period. This is the subset of beneficiaries with whom CGP staff would have had the maximum opportunity to intervene. Beneficiaries who died during this period but were fully eligible and participating up to their deaths were also included. The number of intervention beneficiaries that met these criteria was 1,083 for the original population and 534 for the refresh population.

4.3 Findings

4.3.1 Participation Rates for the CGP 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 CGP is 36 months for the original population and 24 months for the refresh.

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 here are 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 the CGP had a total of 36 months (or 1,096 days) of possible enrollment. If they 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 in the tables below.

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

The ratio of FTEs to the total number of eligible beneficiaries in the original intervention population is 0.79 for the entire intervention period (months 1-36) compared with a higher ratio (0.92) for each individual year of the demonstration. These differences in ratios illustrate the effect of subdividing 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 the CGP if they joined a Medicare Advantage (MA) plan, lost Medicare Part A or B eligibility, Medicare became a secondary payer, had ESRD, or they moved out of the service area.

Forty-two percent of the original intervention and 47% 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 CGP nonparticipant group was eligible only 67% of all possible days—much lower than the 83% of days for participants. Also, the participant group had a higher rate of beneficiaries being fully eligible for the entire intervention period (61%) compared with 50% for the nonparticipant group.

Table 4-2 displays eligibility data for the refresh population, which is about one-half the size of the original population. The ratio of total number of beneficiaries to FTEs was lower for the full 24 months (0.87) compared to the two 12-month periods (0.93) for the intervention population. This held true for the comparison population as well. However, the percent of beneficiaries that were fully eligible for the full refresh time period is higher among participants (74%) than nonparticipants (61%) or the comparison group (65%), but the difference narrows by the last 12-months of the demonstration (84%, 86%, and 81%, respectively).

Table 4-1
Number of Medicare FFS beneficiaries eligible for and participating in the CGP: Original population

Characteristics	Months 1-36	Months 1-12	Months 13-24	Months 25-36
Intervention group				
Number eligible[1]	2,891	2,889	2,462	2,084
Full time equivalent[2]	2,277	2,656	2,257	1,918
Number fully eligible	1,677	2,423	2,032	1,729
Participants				
Number eligible	2,159	2,023	1,837	1,553
Full time equivalent	1,783	1,925	1,695	1,441
Number fully eligible	1,311	1,774	1,525	1,296
Participants > 75%				
Number eligible	1,575	1,158	1,701	1,509
Full time equivalent	1,374	1,119	1,571	1,409
Number fully eligible	1,038	1,046	1,414	1,272
Non-participants				
Number eligible	732	866	625	531
Full time equivalent	494	731	562	476
Number fully eligible	366	649	507	433
Comparison group				
Number eligible	1,785	1,778	1,509	1,258
Full time equivalent	1,376	1,629	1,357	1,141
Number fully eligible	945	1,444	1,183	1,019

NOTES:

FFS = fee-for-service; CGP = Care Guidance Program; CMHCB = Care Management for High Cost Beneficiaries.

[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/montefiore/final/tables/tableHB-1.sas 19MAY2010.

Table 4-2
Number of Medicare FFS beneficiaries eligible for and participating in the CGP: Refresh population

Characteristics	Months 1-24	Months 1-12	Months 13-24
Intervention group			
Number eligible[1]	896	896	779
Full time equivalent[2]	778	832	723
Number fully eligible	643	760	656
Participants			
Number eligible	725	710	646
Full time equivalent	654	683	600
Number fully eligible	539	626	542
Participants > 75%			
Number eligible	617	418	617
Full time equivalent	577	408	576
Number fully eligible	490	377	526
Non-participants			
Number eligible	171	186	133
Full time equivalent	124	149	124
Number fully eligible	104	134	114
Comparison group			
Number eligible	868	867	727
Full time equivalent	724	784	664
Number fully eligible	565	682	590

NOTES:

FFS = fee-for-service; CGP = Care Guidance Program; CMHCB = Care Management for High Cost Beneficiaries.

[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/montefiore/final/tables/tableHB-1.sas 19MAY2010.

Tables 4-3 and 4-4 present participation rates for the CGP's 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 effects the ability of CGP staff to contact beneficiaries and, ultimately, for the program to have any impact on utilization and costs.

Participation rates for the CGP's original population. Of all the CGP's original intervention group beneficiaries, 75% verbally consented to participate in its demonstration at some point during the intervention period. We previously reported (Brody and Bernard, 2007) that, as of January 2007, 56% consented in the initial 7-month engagement period and we observe an increase in the CGP's enrollment over the entire intervention period. Only 49% of beneficiaries were continuous participants (***Table 4-3***), which equates to 65% of participants. Among CGP beneficiaries, 21% refused to participate. The percent not contacted or unable to be located was 4%.

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 CMHCB demonstration. Of CGP's intervention beneficiaries, 54% participated for more than 75% of their eligible months, which is lower 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 the CGP refresh population. The criteria for selection of the intervention and comparison refresh populations were similar to the criteria used to select the initial populations with one noted exception. MMC expanded the list of CPT and Place of Service codes to exclude more residents of SNFs and nursing homes. With the selection criterion change, there was improvement in their participation rate ***(Table 4-4)***. Overall, 81% of the refresh intervention beneficiaries consented to participate at some point during the 24-month period. Of those, 63% were continuous participants, which equates to 78% of participants. The percent that refused to participate was modestly lower (14%), and the percent that were not contacted or were unable to be contacted was modestly higher at 5%.

Table 4-3
Participation in the CGP: Original population

Characteristics	Statistic
Number of intervention months	36
Participation rate (entire demonstration period)	75%
Length of participation	
Continuous participation after engagement	49%
After initial participation, became a continuous non-participant	22%
Intermittent participation	4%
Nonparticipation (never agreed)	25%
Refused to participate when contacted	21%
Not contacted/unable to be contacted	4%
Beneficiaries participating more than 75% of months	54%
Number of participants in selected months[1]	
Month 6	1,498
Month 12	1,711
Month 24	1,536
Month 36 (last month)	1,303

NOTES: CGP = Care Guidance Program; CMHCB = Care Management for High Cost Beneficiaries.

[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/montefiore/final/tables/tableHB-2.sas 19MAY2010.

Table 4-4
Participation in the CGP: Refresh population

Characteristics	Statistic
Number of intervention months	24
Participation rate (entire demonstration period)	81%
Length of participation	
Continuous participation after engagement	63%
After initial participation, became a continuous nonparticipant	15%
Intermittent Participation	3%
Nonparticipation (never agreed)	19%
Refused to participate when contacted	14%
Not contacted/unable to be contacted	5%
Beneficiaries Participating more than 75% of months	69%
Number of participants in selected months[1]	
Month 6	658
Month 12	629
Month 24 (last month)	547

NOTES: CGP = Care Guidance Program; CMHCB = Care Management for High Cost Beneficiaries.

[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/montefiore/final/tables/tableHB-2.sas 19MAY2010.

4.3.2 Characteristics of the CGP's Intervention and Comparison Populations

In addition to evaluating the level of initial engagement by the CGP, 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 program 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 the CPG's populations—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 the CPG demonstration program 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 higher percentages of the disabled, men, beneficiaries under the age of 65 and ages 80-84, and higher rates comorbidities such as heart failure, and coronary artery disease. Intervention beneficiaries also had lower rates of Medicaid beneficiaries, non-White beneficiaries, high HCC scores, and other comorbid conditions – e.g., diabetes without complications, and renal failure. The refresh population has fewer statistically significant differences in the beneficiary characteristics – primarily in higher percentages of coronary artery disease among the intervention beneficiaries. This is due to the very low number of beneficiaries in the never participated category. ***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 the CGP's Original and Refresh Populations

In order to better understand the characteristics that most strongly predicted 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 the CPG more than 75% of their eligible months. Model 1 reflects the initial recruitment emphasis by the CGP, or characteristics

of beneficiaries with whom the CGP had the longest potential period of intervention. Model 3 reflects characteristics of the beneficiaries who demonstrated the greatest willingness or ability to participate in the CGP. 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 for the original population comprises characteristics associated with 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 0.03 or less, with the full Model 3 exhibiting pseudo R-squares of 0.03 for the refresh population. ***Supplement 4A*** contains tables that present the odds ratios and level of significance for Models 1 and 2.

Model 3a shows that non-White beneficiaries are more likely to participate (***Table 4-5***). Model 3b for the original population (***Table 4-5***), introduces baseline and demonstration period health status measures. Non-White beneficiaries continue to be more likely to participate. In addition, beneficiaries with high baseline HCC scores and medium and high baseline PBPMs were more likely to participate than those with a low baseline HCC score or low baseline PBPMs, holding other factors constant. These measures are proxies for poorer health status.

There are a few statistically significant results for the refresh population (***Table 4-6***). There are no differences in Model 3A and in Model 3B, beneficiaries that died were less likely to participate than those that remained alive, holding other factors constant. Only 32% of the refresh population, or 386 beneficiaries, were in the reference group making it difficult to determine statistically significant differences.

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

Characteristics	Model 3A OR	p^3	Model 3B OR	p^3
Intercept	1.38	**	1.30	N/S
Beneficiary characteristics				
Male	0.87	N/S	0.89	N/S
Non-White	1.41	**	1.34	**
Age < 65 years	1.49	N/S	1.39	N/S
Age 75-84	1.03	N/S	1.05	N/S
Age 85 + years	1.10	N/S	1.15	N/S
Medicaid	0.87	N/S	0.85	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.17	N/S
Baseline HCC score high	N/I	N/I	1.63	**
Medium baseline PBPM	N/I	N/I	1.25	*
High baseline PBPM	N/I	N/I	1.34	*
Baseline Charlson score medium	N/I	N/I	0.86	N/S
Baseline Charlson score high	N/I	N/I	0.81	N/S
Demonstration period health status				
Died	N/I	N/I	0.58	**
Concurrent HCC score medium	N/I	N/I	0.87	N/S
Concurrent HCC score high	N/I	N/I	0.85	N/S
Number of cases	2,891	N/A	2,891	N/A
Chi-square (p<)	22.73	**	66.73	**
Pseudo R-square	0.01	N/A	0.02	N/A

NOTES: CGP = Care Guidance 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 (CMO) 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 < $355.83. The baseline Charlson score reference group is < 3. The concurrent HCC score reference group is 0.631 or less.

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

Program: bene02 25MAY2010, partab3b and partab4b 01JUNE2010.

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

Characteristics	Model 3A OR	p^3	Model 3B OR	p^3
Intercept	2.67	**	2.00	**
Beneficiary characteristics				
Male	0.88	N/S	0.83	N/S
Non-White	1.42	N/S	1.37	N/S
Age < 65 years	0.99	N/S	0.99	N/S
Age 75-84	0.99	N/S	0.98	N/S
Age 85 + years	1.21	N/S	1.13	N/S
Medicaid	0.56	N/S	0.64	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.50	N/S
Baseline HCC score high	N/I	N/I	1.22	N/S
Medium baseline PBPM	N/I	N/I	1.22	N/S
High baseline PBPM	N/I	N/I	0.96	N/S
Baseline Charlson score medium	N/I	N/I	1.17	N/S
Baseline Charlson score high	N/I	N/I	1.08	N/S
Demonstration period health status				
Died	N/I	N/I	0.48	**
Concurrent HCC score medium	N/I	N/I	1.16	N/S
Concurrent HCC score high	N/I	N/I	1.11	N/S
Number of cases	896	N/A	896	N/A
Chi-square (p<)	7.51	N/S	23.22	N/S
Pseudo R-square	0.01	N/A	0.03	N/A

NOTES: CGP = Care Guidance 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 (CMO) 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 <$245. The baseline Charlson score reference group is <2. The concurrent HCC score reference group is 0.606 or less.

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

Program: bene02 25MAY2010, partab3b and partab4b 01JUNE2010.

4.3.4 Level of Intervention

In this section, we report the frequency of interaction between the CGP and intervention beneficiaries for a subset of original intervention population beneficiaries who were fully eligible and participating for months 7 through the end of the CGP. However, all interactions for the full demonstration time period (36 months for the original and 24 months for the refresh populations) were included for this subset of beneficiaries. We also examine whether there is evidence of selective targeting of beneficiaries for interventions based upon level of perceived need as determined by beneficiary demographic, health status, baseline costliness, and acute care utilization during the demonstration period. During the first site visit, CGP staff stated that all eligible beneficiaries were at high risk for acute medical events as indicated by having an HCC score of 1.8 or higher. Thus, they chose not to use level of health risk to prioritize the waves of outreach activities (Brody and Bernard, 2007). By the second site visit, CGP staff reported that they had implemented a number of major changes, including creating a High Utilization Team (HUT) that targets the sickest and most frail beneficiaries, continuing outreach to nonparticipating beneficiaries from the original cohort, recruiting self-directed beneficiaries into active case management, optimizing staff resources by using nonclinical staff for baseline assessments and patient enrollment, and hiring a dedicated palliative care provider. In addition, the CGP attempted to reduce unnecessary admissions by conducting discharge assessments to facilitate further outpatient follow-up visits, evaluating the need for home care services, and making sure all needed services and medications were actually received not just prescribed (Khatutsky, McCall, and Bernard, 2009). Thus, we expect to see a pattern of higher levels of interventions for beneficiaries in poorer health status or higher users of hospitalization services.

Descriptive statistics were performed on beneficiaries participating in the CGP to determine the breadth and depth of interventions related to care management. The data represent beneficiaries who were fully eligible and participating (unless they died) for months 7 through 36 for the original population and 7-24 for the refresh population. A total of 1,083 unique beneficiaries met these criteria for the original population. Observations were weighted by the fraction of eligible days, accounting for fewer interventions due to loss of program eligibility primarily due to death, which resulted in 1,005 full-time equivalent beneficiaries. The refresh population has 534 beneficiaries, or 517 full-time equivalent beneficiaries.

RTI was provided information on numbers of CGP interventions that participants received during the demonstration period. Interventions were defined as assessments as well as intervention actions taken by or strategies determined by CGP staff to resolve "problems" identified during assessment or re-assessment, or as requested by the participant. These two types of interventions are presented separately in order to describe the different types of interventions that the CGP provided. This is followed by tables of total interventions, defined as assessments plus intervention actions, to provide the breadth and depth of the program.

Tables 4-7 and 4-8 provide a detailed description of the type of intervention and number of interventions during this time period for the subset of eligible beneficiaries. ***Table 4-7*** gives a broad sense of the types of assessments that were conducted with participants. The majority of assessments were routine follow-up/post discharge assessments (about 50%). Thirteen percent of assessments were determined to be for self-directed participants. These were participants that decline to answer the assessment survey, usually because they did not want to take the time. As

a result, these beneficiaries are referred to as "self-directed" and were to receive periodic telephone calls from a member of the enrollment staff every 90 days to determine whether major issues had emerged. If a self-directed participant was hospitalized, the CGP medical director was to visit the patient in the hospital and a care manager was assigned to contact the beneficiary by telephone following discharge. Over the 36-month demonstration period, over 10,000 assessments were conducted for the 1,083 participants in our analysis. This demonstrates that CGP staff repeatedly worked with their participants to personalize their care plan.

Among the refresh population, a much higher percentage of assessments were for self-directed participants (27%). The lower number of total assessments is driven by the self-directed participants and by the fact that the refresh population in this analysis is about ½ that of the original population and the time period was only 24 months.

Table 4-7
Frequency distribution of the CGP's assessments[1,2]

Assessment type	Original Frequency	Percent	Refresh Frequency	Percent
Comprehensive Baseline	619	5.8	153	4.3
Falls Risk Inventory	621	5.8	162	4.6
Routine Follow-Up/Post Discharge	5,354	50.4	1,353	38.4
Flu/Pneumovax	2,083	19.6	731	20.8
PHQ-9	115	1.1	17	0.5
Self-directed	1,379	13.0	965	27.4
Diabetes DM Baseline/DM Enrollment	78	0.7	18	0.5
CHF DM Baseline	44	0.4	18	0.5
Health Buddy® Enrollment/Disenrollment	312	2.9	100	2.8
Undefined	11	0.1	3	0.1
Total assessments	10,616	100.0	3,520	100.0

NOTES: CGP = Care Guidance Program

[1] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

[2] Includes only assessments with an interview status of accepted and excludes pending assessments.

Data Sources: RTI analysis of 2007-2008 Medicare enrollment, eligibility, and CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab1 05AUG2010.

Table 4-8 shows the types of intervention actions that were implemented based on the responses to assessments or that were determined by CGP staff to be in need of some action. For each participant, the CGP information system developed a problem list based on responses to the assessment. Clinical staff developed a customized care plan by assigning at least one intervention to each problem identified. Notably, the program focused on those problems that could benefit from an actionable intervention. For instance, if a beneficiary had functional difficulties, the CGP system would identify this issue as a problem only if the participant needed additional assistance in the home; conversely, if the beneficiary had sufficient home-based support, the functional difficulties were not labeled as a problem. Patient education/mailings for life planning were the most common action in the original and refresh populations, 49% and 43%, respectively. Referrals to providers (e.g., primary care physicians, dentists, physical/occupational therapy, and home care) comprised about 18% of all interventions for both the original and refresh populations.

Table 4-8
Frequency distribution of the CGP intervention actions[1,2]

Intervention type	Original Frequency	Percent	Refresh Frequency	Percent
Provider Referrals	651	18.3	142	17.9
Program Referrals	117	3.3	24	3.0
Mailing (only)	23	0.6	6	0.8
Home Visit	1	0.0	1	0.1
CGP Interventions	24	0.7	9	1.1
Other Referrals	2	0.1	0	0.0
Pharmacist Review/(internal)	357	10.0	75	9.4
Nutritionist Review/Referral	296	8.3	112	14.1
Patient Education/Mailing – General	147	4.1	48	6.0
Patient Education/Mailing – DM only	205	5.8	40	5.0
Patient Education/Mailing – Life Planning only	1,734	48.7	338	42.5
Total intervention actions	3,557	100.0	795	100.0

NOTES: CGP = Care Guidance Program

[1] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

[2] Includes only intervention actions with a valid intervention status (status value was not missing or did not indicate that the intervention was not met, deactivated, or postponed) and excludes intervention actions defined as mailings only.

Data Sources: RTI analysis of 2007-2008 Medicare enrollment, eligibility, and MMC's CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab1 05AUG2010.

Table 4-9 displays the number and percent of the participants in our analysis that received each type of assessment. Nearly 90% of the 1,005 FTE participants in our original population had a flu vaccination assessment while 71% received a routine follow-up/post discharge assessment. Nearly 30% of participants were determined to be self-directed. In the refresh population, 55% of the 517 FTE participants were self-directed. Thus, the percent of participants completing the comprehensive baseline and falls risk inventory assessments was about one-half that of the original population.

Table 4-9
Number and percent of participants[1] receiving assessments[2] by type in the CGP

Assessment Type	Original FTE beneficiaries	Percent	Refresh FTE beneficiaries	Percent
Comprehensive Baseline	616	61.3	153	29.6
Falls Risk Inventory	618	61.5	162	31.3
Routine Follow-Up/Post Discharge	711	70.7	247	47.8
Flu/Pneumovax	882	87.8	414	80.1
PHQ-9	65	6.5	12	2.3
Self-directed	280	27.9	283	54.7
Diabetes DM Baseline/DM Enrollment	78	7.8	18	3.5
CHF DM Baseline	43	4.3	18	3.5
Health Buddy® Enrollment/Disenrollment	232	23.1	78	15.1
Undefined	11	1.1	3	0.6

NOTES: CGP = Care Guidance Program; FTE = full time equivalent.

[1] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population

[2] Includes only assessments with an interview status of accepted and excludes pending assessments.

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

Data Sources: RTI analysis of 2006-2009 Medicare enrollment, eligibility, and MMC'S CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab2 05AUG2010; enctab4a 17AUG2010.

Table 4-10 displays the number and percent of participants receiving an intervention action based on assessments or CGP staff determination. Nearly one-third of the 1,005 FTE original population beneficiaries received a provider referral, pharmacist review and/or patient education and mailing in reference to disease management. Among the refresh population, the percentage of beneficiaries receiving intervention actions was about one-half that of the original population.

Table 4-10
Number and percent of participants[1] receiving intervention actions[2] by type in the CGP

Action Type	Original FTE beneficiaries	Percent	Refresh FTE beneficiaries	Percent
Provider Referrals	320	31.8	70	13.5
Program Referrals	66	6.6	18	3.5
Home Visit	1	0.1	1	0.2
CGP Interventions	17	1.7	7	1.4
Other Referrals	1	0.1	0	0.0
Pharmacist Review/(internal)	314	31.2	73	14.1
Nutritionist Review/Referral	183	18.2	67	13.0
Patient Education/Mailing – General	107	10.6	25	4.8
Patient Education/Mailing – DM only	330	32.8	13	2.5
Patient Education/Mailing – Life Planning only	48	4.8	92	17.8

NOTES: CGP = Care Guidance Program; FTE = full time equivalent.

[1] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population

[2] Includes only intervention actions with a valid intervention status (status value was not missing or did not indicate that the intervention was not met, deactivated, or postponed) and excludes intervention actions defined as mailings only.

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

Data Sources: RTI analysis of 2006-2009 Medicare enrollment, eligibility, and MMC'S CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab2 05AUG2010.

Table 4-11 displays the overall distribution of care management-related interventions for the original and refresh populations. A total of 1,083 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 interventions due to attrition because of death, which resulted in 1,005 full-time equivalent beneficiaries. A high percentage (86%) of these beneficiaries had at least one assessment during the full demonstration period (months 1-36). The mean number of assessments among beneficiaries that had an assessment was 12 (an average of 1 per quarter) and the median was 1 assessment. Fifty-seven percent of participants had an intervention action. The mean number of interventions over the 36-month demonstration period for those original participants that had at least 1 intervention action was 6 with a median of 4 intervention actions. Intervention actions were focused on identified individual needs and took time to implement and complete. Total interventions are defined as assessments and intervention actions. On average, the 1,083 original population participants had 14 interventions over the course of the demonstration.

The refresh population had a total of 534 unique refresh population beneficiaries met the selection criteria (517 full-time equivalents). A slightly lower percentage of beneficiaries had an assessment (82%) and a much lower percentage of participants had an intervention action (32%) during the 24-month refresh period. However, those that did have an intervention action had on average 5 over the 24-month refresh demonstration period. Average total interventions for the 534 participants were about 1 per quarter (8 interventions over a 24-month period).

Table 4-12 displays the frequency of total interventions by baseline HCC score for both the original and refresh populations. Participants could have a combination of assessments and intervention actions at any time during the 36 (or 24) 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. There is evidence that CGP staff made a focused effort to determine and address the needs of their higher acuity beneficiaries. Thirty percent of the original population's beneficiaries received 20 or more interventions in contrast with 24% of beneficiaries with medium HCC sores and 19% of beneficiaries with a low baseline HCC score. Seventy-one percent of the highest risk beneficiaries had 10 or more interventions compared to 68% for beneficiaries with a medium baseline HCC score and 64% for beneficiaries with a low baseline risk score.

These differences are even more pronounced for the refresh population. Fourteen percent of participants with a high baseline risk score received 20 or more interventions compared to 5% for those in the low risk category. Participants in the high risk category had a higher percentage of beneficiaries that had 10 or more interventions compared to those categorized as medium risk at baseline during the 24-month refresh demonstration period (39% compared to 30%). Only 23% of beneficiaries with low baseline risk had 10 or more interventions. The main difference between the original and refresh distribution of interventions by HCC risk is the broadening of the differences between the high and low risk categories indicating even more targeting of interventions for the refresh population.

Table 4-11
Distribution of number of interventions[1,2] with participants3 in the CGP

Statistic	Original Population	Refresh Population
Number of beneficiaries[3]	1,083	534
FTE beneficiaries[4]	1,005	517
Number of beneficiaries with an assessment[4]	866	426
Mean number of assessments[5]	12	8
Median number of assessments[5]	11	7
Number of beneficiaries with an action[4]	573	166
Mean number of intervention actions[6]	6	5
Median number of intervention actions[6]	4	3
Mean number of total interventions[3]	14	8
Median number of total interventions[3]	13	7

NOTES: CGP = Care Guidance Program; FTE = full time equivalent.

[1] Includes interventions at any point during the demonstration period.

[2] Includes only intervention actions with a valid intervention status (status value was not missing or did not indicate that the intervention was not met, deactivated, or postponed) and excludes intervention actions defined as mailings only.

[3] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

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

[5] Reported for beneficiaries with an assessment.

[6] Reported for beneficiaries with an action.

Data Sources: RTI analysis of 2006-2009 Medicare enrollment, eligibility, and CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab2 11AUG2010.

Table 4-12
Frequency of CGP interventions by HCC score:
Original and refresh intervention populations

Total (both assessment and intervention action)	HCC Score High (>3.1) N = 276 Frequency	%	HCC Score Medium (2-<3.1) N = 404 Frequency	%	HCC Score Low (<2) N = 325 Frequency	%
Original						
0	42	15.1	48	11.9	49	15.2
1	3	1.1	2	0.6	2	0.7
2-4	7	2.7	15	3.7	11	3.4
5-9	29	10.4	65	16.0	53	16.4
10-19	111	40.3	178	44.1	146	45.0
20+	84	30.4	96	23.7	63	19.4
Refresh	N = 130 Frequency	%	N = 196 Frequency	%	N = 191 Frequency	%
0	7	5.7	25	13.0	58	30.4
1	1	0.9	1	0.5	0	0.0
2-4	15	11.5	17	8.4	14	7.3
5-9	56	42.9	93	47.7	76	39.8
10-19	33	25.2	48	24.2	34	17.8
20+	18	13.9	12	6.1	9	4.7

NOTES: CGP = Care Guidance Program; HCC =Hierarchical Condition Category; N = number of beneficiaries.

[1] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

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

Data Sources: RTI analysis of 2006-2009 Medicare enrollment, eligibility, and CGP encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/mmc/ enctab2 11AUG2010.

To examine more directly the targeting strategy of the CGP, a multivariate logistic regression model was estimated with the number of total interventions as the dependent variable. The model estimates the likelihood of a participant receiving a high number of interventions. The medium intervention group was omitted, thus comparing the high intervention group to the low intervention group. ***Tables 4-13*** (original population) ***and 4-14*** (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 the last 18 months for the refresh population, and their number of interventions categorized either as low or high. Low interventions are defined as 8 or fewer interventions for the original population and less than 6 interventions for the refresh population. Original population participants with 16 or more interventions are captured in the high intervention category, while 9 or more interventions are defined as high intervention 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 intervention rate are 1.29 times greater than those for a beneficiary age 65 and older, adjusting for any baseline difference in other beneficiary and baseline characteristics and demonstration period health status.

For the original population, medium and high baseline PBPM costs were statistically significant indicators of being in the high intervention category (***Table 4-14***). No demonstration period health status variables were found to be strong predictors of a high level of intervention. Beneficiaries who died or were institutionalized during the demonstration, or were ages 85 or older were less likely to be in the high intervention category. The explanatory power of the studied beneficiary characteristics was low, suggesting that there is not a strong set of variables that predict likelihood of a beneficiary being in the high intervention group. The pseudo R-square for this model was 13%.

For the refresh population, the only statistically significant result was that beneficiaries that died were less likely to be in the high intervention category (***Table 4-15***). However, the number of beneficiaries included in this analysis is very small (360).

Table 4-13
Logistic regression modeling results comparing the likelihood of being in the CGP high intervention[1] category relative to the low intervention category: Original intervention population

Characteristics	Odds ratio[2,3]	*P*[4]
Intercept	0.71	N/S
Beneficiary characteristics		
Male	2.52	**
Non-White	1.41	N/S
Age <65	1.29	N/S
Age 75-84	0.79	N/S
Age 85+ years	0.45	**
Medicaid	0.87	N/S
Baseline characteristics		
Baseline HCC score medium	1.20	N/S
Baseline HCC score high	1.24	N/S
Medium base PBPM costs	1.75	*
High base PBPM costs	2.04	*
Baseline Charlson score medium	0.69	N/S
Baseline Charlson score high	0.72	N/S
Demonstration period health status		
Died	0.09	**
Institutionalized	0.08	*
Concurrent HCC score medium	1.21	N/S
Concurrent HCC score high	1.24	N/S
One hospitalization	1.29	N/S
Multiple hospitalizations	1.28	N/S
Number of cases	729	N/A
Chi-square (p<)	105.25	**
Pseudo R2	0.13	N/A

NOTES: CGP = Care Guidance Program; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Defined as both intervention actions and assessments.

[2] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

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

[4] * 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 cost reference group is < $355.83. The baseline Charlson score reference group is <3. The concurrent HCC score reference group is 0.631 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/MMC/enctab3 18AUG2010.

Table 4-14
Logistic regression modeling results comparing the likelihood of being in the CGP high intervention[1] category relative to the low intervention category: Refresh intervention population

Characteristics	Odds ratio[2,3]	*P*[4]
Intercept	0.28	**
Beneficiary characteristics		
Male	1.12	N/S
Non-White	1.26	N/S
Age <65	0.77	N/S
Age 75-84	1.15	N/S
Age 85+ years	1.32	N/S
Medicaid	0.85	N/S
Baseline characteristics		
Baseline HCC score medium	1.35	N/S
Baseline HCC score high	2.22	N/S
Medium base PBPM costs	1.35	N/S
High base PBPM costs	1.55	N/S
Baseline Charlson score medium	1.98	N/S
Baseline Charlson score high	1.28	N/S
Demonstration period health status		
Died	0.23	*
Institutionalized	0.08	N/S
Concurrent HCC score medium	1.12	N/S
Concurrent HCC score high	1.16	N/S
One hospitalization	1.51	N/S
Multiple hospitalizations	1.54	N/S
Number of cases	360	N/A
Chi-square (p<)	42.89	**
Pseudo R2	0.11	N/A

NOTES: CGP = Care Guidance Program; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Defined as both intervention actions and assessments.

[2] Beneficiaries had to be fully eligible and full participants for months 7-36 for the original population and 7-24 for the refresh population.

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

[4] * 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 cost reference group is <$245. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.606 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/MMC/enctab3 18AUG2010.

4.4 Summary

The CGP was successful in recruiting 75% of its original population beneficiaries and 81% of its refresh population beneficiaries. Despite the high levels of participation we found statistically significant differences between participants and nonparticipants in the original intervention population. Participants tended to be in poor health as defined by having higher percentages of the disabled, beneficiaries with high baseline HCC risk scores and PBPM costs, and beneficiaries with several comorbid conditions such as heart failure, diabetes, and hypertension. These differences were not found for the refresh population. In multivariate modeling of factors that predict likelihood of participation, proxies for poorer health status such as high baseline HCC scores and medium and high baseline PBPM costs indicated a higher likelihood to participate than those with a low baseline HCC score or low baseline PBPM costs, holding other factors constant. The substantially smaller sample size for the refresh population limited our ability to detect participation factors.

The core of the CGP consisted of one-on-one telephone calls between participants and care managers, who linked beneficiaries with needed medical and social services. The program provided the following specific services to participants: care coordination, clinical pharmacist review, link to community support services, nutritional monitoring and counseling, psychosocial support, life care planning, and disease management and telemonitoring. Beneficiaries could participate in any or all of the program elements during the demonstration program, depending on their needs throughout the period. Nearly every participating original population beneficiary received at least one assessment during the demonstration and nearly 60% had an intervention action based on results from the assessment or because of perceived need determined by a CGP staff member. However, the average number of interventions was high for this chronically ill complex population, 14 interventions on average. When examining the rate of intervention by baseline health status measured by the HCC risk score, we found evidence that CGP staff made a focused effort to intervene with 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 interventions during the 36-month period compared to the low HCC risk score group (71% compared to 64%) with 30% of beneficiaries in the high HCC risk score group receiving 20 or more interventions compared with 19% of beneficiaries in the low HCC risk score group. We observe a lower percentage of high HCC risk score refresh beneficiaries with no interventions. We also observe more targeting by HCC risk score for the refresh population.

A major focus of the CGP was to develop customized care plans by assigning at least one intervention to each problem identified through the assessments. Notably, the program focused on those problems that could benefit from an actionable intervention. These findings suggest that CGP staff members were successful in their effort to conduct assessments and follow-through with intervention actions as needed for beneficiaries who were at the highest risk of health events as determined by baseline HCC risk scores.

CHAPTER 5
CLINICAL QUALITY PERFORMANCE

5.1 Introduction

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

- *Clinical Quality of Care:* Did the CGP 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 the CGP 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 Medicare 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.

Under an intent-to-treat (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.

Although the CMHCB demonstration program does not hold MMC financially responsible for quality of care improvements, the CGP does focus upon improvement in rate of compliance with evidence-based care guidelines. This analysis will provide CGP staff with additional information on intervention population performance against the comparison population.

5.2 Method

We created the process-of-care measures for the 12-month period immediately prior to the go-live date for the CGP 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 IVD.

Table 5-1
Number of beneficiaries included in analyses of guideline concordant care and acute care utilization for the CGP

Statistics	All	Diabetes	Ischemic vascular disease
Original beneficiaries			
Months 7-18			
Intervention			
Total number of beneficiaries	2,663	1,099	1,143
Full time equivalents[1]	2,654	1,094	1,139
Comparison			
Total number of beneficiaries	1,651	746	665
Full time equivalents[1]	1,638	740	662
Months 27-38			
Intervention			
Total number of beneficiaries	2,084	849	885
Full time equivalents[1]	2,079	847	883
Comparison			
Total number of beneficiaries	1,258	555	496
Full time equivalents[1]	1,248	550	495
Refresh beneficiaries			
Months 15-26			
Intervention			
Total number of beneficiaries	779	242	294
Full time equivalents[1]	774	241	292
Comparison			
Total number of beneficiaries	727	246	240
Full time equivalents[1]	717	243	236

NOTES: CGP = Care Guidance Program

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

SOURCE: RTI analysis of 2005-2009 Medicare enrollment and eligibility data; Computer runs: gcc01, gcc02, gcctab, gcc_rob, gcctabx, gcctab1 20MAY2010.

Medicare claims for the full baseline and intervention period were included regardless of beneficiary eligibility for the CGP (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 the CGP 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 the CGP—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. The CGP 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 CGP eligibility criteria in the baseline and demonstration periods. STATA SVY was used to fit the model with robust variance estimation.

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 D-in-D 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 CGP 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.[8] 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 CGP original population beneficiaries are reported 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 comparison group and across the two demonstration evaluation periods are relatively high ranging from 86% for LDL-C testing for beneficiaries with ischemic vascular disease to 91% for LDL-C testing for beneficiaries with diabetes. However, the baseline rate for influenza vaccine was only 50%. Claims data are likely to produce underestimates of the rate of influenza vaccination as they do not capture flu vaccines that people receive in pharmacies, supermarkets, senior centers, or city-funded health care centers because services received in those settings do not result in Medicare claims. We observe systematically lower rates of baseline compliance across all four measures for the original comparison population.

Over the course of the two demonstration periods for the original comparison population, we generally observe stable or increasing rates of receipt. In contrast, we generally observe declines in the rates of all measures for the original intervention population with the exception of influenza vaccination where we observe modest increases. Of the eight measures evaluated for the original population, we detect five statistically significant negative difference-in-differences rates; whereby the intervention group’s rates generally declined while the comparison group’s generally increased.

For the refresh population, we generally observe similar levels of compliance with the evidence-based care guidelines at baseline as noted for the original population. With the exception of influenza vaccination, rates decline over the last 12 months for both groups. For influenza vaccination, rates increase over the last 12 months for both groups. Although we do not have self-reported rates of influenza vaccines for the comparison group beneficiaries, MMC’s CGP reported to RTI that their survey of participants revealed over three-quarters reported having received a flu vaccine during each of the three years of the demonstration. We

---

[8] 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 fee-for-service (FFS) beneficiaries.

do not detect any statistically significant differences in the rates between the intervention and comparison beneficiaries.

Table 5-2
Comparison of rates of guideline concordant care for the first and last 12 months of the CGP demonstration period with rates for a 1-year period prior to the start of the CGP: 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	50	38	52	44	-4.40	0.83	0.04	0.69	1.00
Beneficiaries with diabetes									
HbA1c test	88	83	87	87	-5.57	0.63	0.02	0.43	0.94
LDL-C test	90	78	84	82	-9.24	0.48	0.00	0.33	0.70
Beneficiaries with IVD[2]									
LDL-C test	86	75	81	75	-5.34	0.67	0.02	0.48	0.95
Months 25-36									
All beneficiaries									
Influenza vaccine	52	39	58	43	2.10	1.08	0.44	0.88	1.33
Beneficiaries with diabetes									
HbA1c test	89	84	86	85	-3.93	0.71	0.13	0.45	1.11
LDL-C test	91	80	82	81	-9.87	0.44	0.00	0.29	0.66
Beneficiaries with IVD[2]									
LDL-C test	88	78	81	74	-3.15	0.72	0.11	0.48	1.08
Refresh population									
Months 13-24									
All beneficiaries									
Influenza vaccine	41	30	52	46	-5.35	0.77	0.09	0.57	1.04
Beneficiaries with diabetes									
HbA1c test	88	90	87	83	6.03	1.70	0.18	0.78	3.69
LDL-C test	87	85	86	83	1.11	1.07	0.85	0.52	2.22
Beneficiaries with IVD[2]									
LDL-C test	86	80	83	79	-1.51	0.87	0.67	0.45	1.66

NOTES: CGP = Care Guidance Program; CMHCB = Medicare Care Management for High Cost Beneficiaries; 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 (CMO) 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.

[3] The calculated differences for ESRD beneficiaries is a simple intervention minus comparison rate. T-tests are used to determine statistical significance.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: gcc01, gcc02, gcctab, gcc_rob, gcctabx, gcctab1 20MAY2010.

Table 5-3 displays the percentages of the CGP's original and refresh populations who did or did not receive one of the process-of-care measures during the baseline period and 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 period; and
- became compliant, meaning noncompliant in the baseline period but compliant in the intervention period.

Table 5-3
Percentage of comparison and intervention beneficiaries meeting process-of-care standards in the baseline year and last 12 months of the CGP: 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	74%	80%	67%	78%	63%	74%	24%	36%
Never compliant	5	5	7	5	11	7	42	31
Became noncompliant	9	9	12	13	14	13	15	15
Became compliant	11	6	14	5	12	6	20	18
Refresh population	C	I	C	I	C	I	C	I
Always compliant	78	80	73	77	68	75	19	30
Never compliant	5	6	6	3	8	5	44	37
Became noncompliant	12	8	11	11	13	12	10	11
Became compliant	5	6	10	9	11	9	27	22

NOTES: CGP = Care Guidance Program; CMHCB = Medicare Care Management for High Cost Beneficiaries; 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 CMO 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, gcctab, gcc_rob, gcctabx, gcctab3.sas 20MAY2010.

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, the intervention beneficiaries were far more likely to always be compliant in receipt of all four measures than the comparison beneficiaries, ranging from 6 percentage points higher for HbA1c testing to 12 percentage points higher for influenza vaccination. It is not surprising that we see lower rates of never compliant for three of the four measures among the intervention beneficiaries. There are not many differences in the rates of beneficiaries that became noncompliant. Given the higher baseline rates of compliance for the intervention group, it is not surprising to see smaller percentages of comparison beneficiaries becoming compliant during the last 12 months of the demonstration. A similar but less marked pattern is observed among the refresh beneficiaries.

5.4 Summary of Findings and Conclusion

In this chapter, we report on RTI's assessment of the effect of the CGP on quality of care. Specifically, we report findings for the key research question: did the CGP improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?

A review of baseline rates suggested a pre-demonstration difference in patterns of adherence to evidence-based care recommendations between the intervention and comparison groups. It was most notable for the original population. However, the data do not allow us to determine if these differential baseline rates are because of differences in beneficiary or provider behavior. In general, we observe higher rates of baseline compliance among the intervention beneficiaries.

Over the course of the demonstration, the CGP was expected to increase rates of adherence to evidence-based care. Within the original intervention population, we generally observed negative trends in the difference-in-differences rates, not positive, due to either improvement of the rate or less of a decline in the rate of receipt among the comparison beneficiaries. We did not detect statistically significant differences within the refresh population. We also observe one-tenth to just under one-half of intervention beneficiaries in both the original and refresh populations were not compliant during the last year of the demonstration.

CHAPTER 6
HEALTH OUTCOMES

6.1 Introduction

RTI's analysis of health outcomes focuses on answering the following two evaluation questions:

- Did the CGP program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and emergency room (ER) utilization?
- Did the CGP improve health outcomes by decreasing mortality?

In this chapter, we present analyses related to intermediate clinical health outcomes by examining changes in the rate of hospitalizations, ER visits, and readmissions during months 7-18 and the last 12 months of the CGP relative to a 12-month baseline period 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 Method

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 CGP launch 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 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 (D-in-D) 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. A positive intervention effect for the acute care utilization measures occurs if the intervention group's mean rate decreased more or increased less than the comparison group's mean rate during the demonstration period. A negative intervention effect occurs if the intervention group's mean rate declined less or grew more than the comparison group's mean rate during the demonstration period.

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 CGP 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 D-in-D of the rate of hospitalizations and ER visits, and the IRR'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 CGP go-live date. 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 conditions.

The analyses included beneficiaries who had at least 1 day of eligibility in both the baseline and demonstration periods in which index admissions were identified. Only claims that occurred during periods of eligibility were included in the admission and readmission estimates. Estimates of admission rates were weighted by the fraction of days eligible in the 12-month baseline or demonstration periods. 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 D-in-D rate between the groups. If the CGP reduced admissions and readmissions, we expect to observe negative D-in-D rates, reflecting greater reductions or smaller increases in the intervention group relative to the comparison group.

Logistic regression was used to estimate the likelihood of having an admission; a negative binomial generalized linear model was used for estimates of readmission rates. STATA SVY was used to fit the model with robust variance estimation. Regressions were weighted by the eligibility fractions described above. We report the odds ratio from the logistic regressions and the incidence rate ratio from the negative binomial regressions of the D-in-D test along with the associated *p* value and 95% confidence interval. ORs and IRRs less than 1.0 are associated with a negative D-in-D coefficient, indicating that the CGP reduced admissions or readmissions for the intervention group relative to the comparison or slowed the growth in rates.

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[9]. Both a logistic model of the likelihood of death and a Cox proportional hazard model of survival were estimated testing the relationship of a number of independent variables with the 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:

- intervention status, set at 1 for beneficiaries in the intervention group;
- male, a dichotomous variable, set at 1 for males;

[9] We attempted to model mortality with the inclusion of an indicator for being in the intervention group and completing at least one Health Buddy® survey; however sample size was not adequate. Only 164 original intervention and 38 refresh intervention beneficiaries used the Health Buddy® device in the CGP.

- Medicaid, set at 1 for beneficiaries who was enrolled in Medicaid for one or more months during the baseline period;

- 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

- 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 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 CGP original demonstration program was greater than or equal to $355.83 and less than $1,407 (medium) and $1,407 or greater (high); PBPM cost less than $355.83 is the reference group for the original population. For the CGP refresh population, baseline PBPM costs greater than or equal to $245 and less than $1,439 were assigned to the medium group and $1,439 or greater to the high category; PBPM cost less than $245 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 the CGP demonstration periods are presented in ***Table 6-1***. Rates of hospitalization and ER visits are presented for all causes and then for the 10 ACSCs. Next to the columns of the utilization rates are the D-in-D rates of change observed between the baseline period and the demonstration intervention period. Negative D-in-D 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 D-in-D 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. 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 the CGP with rates of utilization for a 1-year period prior to the start of the CGP: 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	712	780	971	990	49	1.07	0.36	0.92	1.25
10 ACSCs[5]	240	255	332	362	-14	0.98	0.85	0.77	1.24
ED/Obs visits									
All-cause	933	848	1,260	1,209	-35	0.95	0.52	0.80	1.12
10 ACSCs	210	194	341	348	-22	0.91	0.53	0.68	1.22
Months 25-36									
Hospitalizations									
All-cause	602	651	923	1,001	-29	1.00	0.97	0.84	1.19
10 ACSCs	183	193	329	360	-22	0.96	0.78	0.73	1.27
ED/Obs visits									
All-cause	827	791	1,173	1,411	-275	0.79	0.02	0.66	0.96
10 ACSCs	162	176	324	396	-58	0.89	0.49	0.64	1.24

NOTES: CGP = Care Guidance Program; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ED/Obs = emergency room visits, including observation bed stays; CMO = Care Management Organization.

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

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

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 acsctab acsc acsctab1 20MAY2010.

Not unexpectedly, the baseline rates of hospitalization and ER visits were high in the GCP original intervention and comparison populations. The baseline rate of all-cause hospitalization was 712 per 1,000 original intervention group beneficiaries (***Table 6-1***). And, the baseline rate of all-cause ER visits was 933 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-third of all-cause hospitalizations and all-cause ER visits. Thus, Medicare FFS beneficiaries in the CGP were being treated in acute care settings quite frequently for prevalent chronic medical conditions, such as HF, diabetes, and COPD, as well as prevalent acute medical conditions such as pneumonia.

The rate of all-cause and ACSC hospitalization increased similarly in the original intervention and the comparison groups between the baseline and both demonstration periods. The trend in D-in-D rates is negative for the hospitalization rates with the exception of the all-cause hospitalization rate during months 7-18 of the demonstration, indicating that the intervention rates increased modestly less than the comparison group's rates with none of the findings being statistically significant. In months 7-18, the D-in-D rate for all-cause hospitalizations is 49 per 1,000 beneficiaries higher in the intervention group than the comparison group (p-value of 0.99).

The rate of all-cause and ACSC ER visits increased similarly in the CGP's original intervention and the comparison groups between the baseline and both demonstration periods. The trend in D-in-D rates is negative for all of the ER visit measures, indicating once again that the intervention rates increased less than the comparison group's rates. Most notable, during the last 12 months of the demonstration we observe a 33% lower rate of growth in the intervention group's all-cause ER utilization which is a statistically significant difference (P=0.02).

Rates of hospitalization and ER visits per 1,000 refresh population beneficiaries for the year prior to go-live and months 13-24 of the CGP are presented in ***Table 6-2***. We observe roughly similar levels of baseline rates of use among 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 CGP demonstration period, and in a manner similar to the original population. We detect no statistically significant differential rates of hospitalizations or ER usage—either all-cause or ACSC—during the demonstration period relative to the baseline period.

Table 6-2
Comparison of rates of utilization for the last 12 months of the CGP with rates of utilization for a 1-year period prior to the start of the GCP: 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	700	707	927	920	15	1.02	0.88	0.8	1.3
10 ACSCs[5]	185	216	288	333	-14	1.01	0.95	0.68	1.51
ED/Obs visits									
All-cause	833	710	1,248	1,296	-171	0.82	0.12	0.64	1.06
10 ACSCs	162	172	290	354	-53	0.87	0.54	0.56	1.35

NOTES: CGP = Care Guidance Program; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ED/Obs = emergency room visits, including observation bed stays; CMO = Care Management Organization.

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

[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 last 12 months the CMO 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 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 UTI.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 acsctab acsc acsctab1 20MAY2010.

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 readmissions per 1,000 beneficiaries. Data are displayed for all-cause and ACSC admissions and readmissions. In general, we observe a pattern of increasing percentage of both intervention and comparison beneficiaries being hospitalized or having a readmission over the course of the demonstration. We detect no statistically significant reductions in percentage of beneficiaries with an admission or readmission among the original intervention beneficiaries during the early stage of the demonstration (months 7-18), nor during the last 12 months of the demonstration (***Table 6-3***). We do detect a statistically insignificant but a sizeable 18% lower rate of growth in rate of all-cause readmission among the intervention beneficiaries during months 22-33. Given that we observe no decline in the percentage of beneficiaries with all-cause readmissions, the trend of declining all-cause readmission rates implies that the CGP was more successful at reducing readmissions for beneficiaries with frequent readmissions than for beneficiaries with less frequent readmissions relative to the comparison group.

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. We do not detect any statistically significant moderation of the growth in the percentage of beneficiaries admitted or readmitted or the readmission rates within the intervention refresh population in comparison with the secular changes over time in the comparison group.

Table 6-3
Number of beneficiaries included in analyses of readmissions for the GCP

Counts of beneficiaries	Intervention	Comparison
Original beneficiaries		
Months 7-18		
Total number of beneficiaries	2,663	1,651
Full time equivalents[1]	2,654	1,638
Months 22-33		
Total number of beneficiaries	2,159	1,316
Full time equivalents[1]	2,154	1,306
Refresh beneficiaries		
Months 10-21		
Total number of beneficiaries	805	758
Full time equivalents[1]	799	747

NOTES: CGP = Care Guidance Program.

[1] Full Time Equivalent for the intervention group during the baseline period is the total number of beneficiaries weighed 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 readmtab1 20MAY2010.

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

Utilization	Baseline[1,2,3] I	Baseline[1,2,3] C	Demo period[1,2,3] I	Demo period[1,2,3] C	D-in-D	OR/IRR[4]	*p*	Low CI	High CI
Months 7-18									
Hospitalizations									
Percent with an admission	38	41	41	40	5	1.21	0.04	1.01	1.45
Percent with ACSC[5] admission	15	17	18	20	-0	1.01	0.91	0.80	1.28
All-cause 90-day readmission									
Percent with readmission	35	37	43	45	0	1.03	0.86	0.77	1.36
Readmission rate / 1,000	711	724	1,015	1,055	-28	0.98	0.87	0.75	1.27
ACSC same-cause 90-day readmission									
Percent with readmission	14	13	15	16	-2	0.86	0.61	0.47	1.56
Readmission rate / 1,000	246	176	251	235	-54	0.76	0.39	0.41	1.41
Months 22-33									
Hospitalizations									
Percent with an admission	34	38	40	42	2	1.10	0.38	0.89	1.35
Percent with ACSC admission	13	15	18	21	-1	0.98	0.91	0.75	1.29
All-cause 90-day readmission									
Percent with readmission	32	33	43	43	1	1.06	0.72	0.76	1.48
Readmission rate / 1,000	647	606	971	1,045	-115	0.87	0.36	0.65	1.17
ACSC same-cause 90-day readmission									
Percent with readmission	12	10	17	12	3	1.21	0.61	0.57	2.57
Readmission rate / 1,000	193	124	288	219	0	0.85	0.68	0.38	1.88

NOTES: CGP = Care Guidance 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 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 UTI.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: readm01 readm02 readmtab1 20MAY2010

Table 6-5
Change in 90-day readmission1 rates between the year prior to the GCP and months 10-21 of the demonstration: Refresh population

Utilization	Baseline[1,2,3] I	Baseline[1,2,3] C	Demo period[1,2,3] I	Demo period[1,2,3] C	D-in-D	OR/IRR[4]	p	Low CI	High CI
Months 10-21									
Hospitalizations									
Percent with an admission	38	39	39	42	-2	0.92	0.60	0.69	1.24
Percent with ACSC[5] admission	13	16	17	20	0	1.04	0.83	0.71	1.54
All-cause 90-day readmission									
Percent with readmission	31	39	44	49	3	1.16	0.53	0.73	1.85
Readmission rate / 1,000	682	730	1,092	1,020	119	1.14	0.52	0.76	1.73
ACSC same-cause 90-day readmission									
Percent with readmission	9	15	14	14	7	2.01	0.21	0.68	5.99
Readmission rate / 1,000	162	269	188	207	87	1.50	0.50	0.46	4.93

NOTES: CGP = Care Guidance 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 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 UTI.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: readm01 readm02 readmtab1 20MAY2010.

6.3.3 Mortality

Mortality rates during the CGP for the original and refresh intervention and comparison populations are displayed in ***Table 6-6***. During the first 6 months of the original and refresh demonstration periods, we detect no statistically significant differences in the rate of mortality between intervention and comparison beneficiaries: 5.3% of original intervention beneficiaries died versus 4.5% of the original comparison beneficiaries and 4.8% of refresh intervention beneficiaries died versus 4.2% of the refresh comparison beneficiaries. Over the 36-month demonstration period for the original population, 29% of beneficiaries in the intervention and comparison groups died. During the 24-month demonstration period for the refresh population, 17% of beneficiaries in the intervention and comparison groups died. Thus, we detect no statistically significant difference in mortality rates for the either the original or refresh populations using either time period of reference.

We further explored mortality in the original refresh populations by estimating a multivariate Cox proportional hazard model of survival for the full 36- or 24-month demonstration period for the original and refresh groups, respectively. The LifeTest procedure reveals that there is no statistically significant difference in time of survival between the intervention and comparison beneficiaries in either the original or refresh populations.

Table 6-6
Mortality rates during the GCP: Original and refresh populations

Description	Intervention number of deaths	Percent	Comparison number of deaths	Percent	Difference	P value
Original population (6 months)	154	5.3%	80	4.5%	0.8%	0.20
Refresh population (6 months)	43	4.8%	36	4.2%	0.6%	0.51
Original population (36 months)	838	29.0%	518	29.0%	-0.03	0.98
Refresh population (24 months)	149	16.6%	146	16.8%	-0.19	0.91

NOTES: CGP = Care Guidance Program.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: mortality.sas 17AUG2010.

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 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 6-7
Cox Proportional Hazard Survival Models for the CGP: Original Population

Original	Model 1 Hazard Ratio	Model 1 p value	Model 2 Hazard Ratio	Model 2 p value
Intervention	0.999	0.9825	1.001	0.9894
Male	N/I	N/I	1.197	0.0013
Medicaid	N/I	N/I	0.962	0.6384
Age less than 65	N/I	N/I	0.696	0.0142
Age 75 to 84	N/I	N/I	1.376	<.0001
Age 85 and older	N/I	N/I	2.528	<.0001
Baseline medium HCC risk score	N/I	N/I	0.929	0.3539
Baseline high risk HCC score	N/I	N/I	1.794	<.0001
Baseline medium PBPM	N/I	N/I	1.127	0.1603
Baseline high PBPM	N/I	N/I	2.015	<.0001

NOTES: CGP = Care Guidance Program; PBPM = per beneficiary per month; HCC = Hierarchical Conditions Category

Program: Dietab3f; August 17, 2010

Table 6-8
Cox Proportional Hazard Survival Models for the CGP: Refresh Population

Original	Model 1 Hazard Ratio	Model 1 p value	Model 2 Hazard Ratio	Model 2 p value
Intervention	0.944	0.992	1.028	0.8149
Male	N/I	N/I	1.278	0.0428
Medicaid	N/I	N/I	1.107	0.6015
Age less than 65	N/I	N/I	0.857	0.5242
Age 75 to 84	N/I	N/I	1.178	0.2856
Age 85 and older	N/I	N/I	2.115	<.0001
Baseline medium HCC risk score	N/I	N/I	0.543	0.0020
Baseline high risk HCC score	N/I	N/I	1.445	0.0478
Baseline medium PBPM	N/I	N/I	1.261	0.2231
Baseline high PBPM	N/I	N/I	2.414	<.0001

NOTES: CGP = Care Guidance Program; PBPM = per beneficiary per month; HCC = Hierarchical Conditions Category

Program: Dietab3f; August 17, 2010

In Model 1, we observe that the intervention variable has a hazard ratio of 0.999 implying no survival advantage to the intervention group (***Table 6-7***). Similarly, we do not see a statistically significant difference in the unadjusted intervention mortality rate for the refresh population with an intervention hazard ratio of 0.944 (***Table 6-8***). In Model 2, we continue to observe no survival advantage among the original or refresh intervention beneficiaries when baseline covariates are added; adjusting for any imbalances between intervention and comparison groups at baseline.

6.3.4 Hospice

A focus of the CGP was encouraging informed 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 3% to 6% (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; median number of hospice days range from 29 to 47.

Table 6-9
Rates of hospice use among intervention and comparison beneficiaries during the CGP

	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	2,891	1,785	6%	6%	-.2	0.80	86	97	-11	0.46	37	35	2	0.79
Refresh population	896	868	3%	4%	-1.0	0.27	106	101	5	0.91	47	29	18	0.49

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and intervention data; Computer runs: hsp01 18MAY2010, hospicetab1 19May2010, hsptest 26MAY2010.

6.4 Conclusions

RTI's analysis of health outcomes focuses on measuring effectiveness of the CGP by answering the following evaluation questions:

- Did the CGP improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?
- Did the CGP improve health outcomes by decreasing mortality?

During the course of the CGP, we observed 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. Out of 30 acute care utilization comparisons, we detect one statistically significant positive intervention effect; the rate of all-cause ER visits grows at a slow rate during the last 12 months of the demonstration within the original intervention group than within the comparison group. The overall net effect is a D-in-D difference in the all-cause ER rate of -275 per 1,000 beneficiaries ($p=0.02$).

We detect no other statistically significant positive intervention effects; however, we do detect one statistically significant negative intervention effect. During months 7 to 18 of the CGP demonstration period, the percent of original intervention beneficiaries with an all-cause hospitalization increased from 38% to 41% while the percent of comparison beneficiaries with an all-cause hospitalization declined from 41% to 40% ($p=0.04$).

We do not detect a statistically significant differential rate of mortality between the intervention and comparison groups of the original or refresh population. Over the 36-month demonstration period for the original population, 29% of beneficiaries in the intervention and comparison groups died. During the 24-month demonstration period for the refresh population, 17% of beneficiaries in the intervention and comparison groups died. We did not detect 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.

This page intentionally left blank

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 the CGP was in operation (or 24 months for the refresh population). The evaluation questions we address are:

- What were the Medicare per beneficiary per month (PBPM) costs 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 MMC in meeting budget neutrality?
- How balanced were the intervention and comparison group populations prior to the demonstration's start date? How important were any 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 that will be conducted for financial reconciliation by Actuarial Research Corporation (ARC) under contract to CMS. ARC will determine savings based on the demonstration's terms and conditions negotiated between CMS and MMC. 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 MMC's CGP 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 populations that might affect t-tests of mean differences in rates of PBPM cost growth. 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. 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 an MA plan). Claims represent utilization anywhere in the United States, not just the target area of the CGP. 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 basis, or the ratio of eligible Medicare costs to eligible months. The baseline period is defined as 365 days (or 1 year) prior to CGP'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[10]; and

[10] For example, if a beneficiary joined a managed care plan for a few months then returned to FFS Medicare, any claims for plan services were excluded.

- only claims for eligible services; end-stage renal disease [ESRD] and hospice 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 the original CGP population (2,891 eligible beneficiaries in both the base and intervention period) stratified by beneficiaries' number of eligible days in the demonstration period (1,095 maximum). Six beneficiaries were eligible for 10 days or less with average PBPM costs of $13,812. Beneficiaries eligible for a year or more averaged PBPM costs of $1,844. Beneficiaries with very truncated eligibility averaged monthly costs 7.5 times greater than those with much longer eligibility. About 3% of the original intervention population was eligible less than 3 months. (See ***Section 7.3.2*** for statistics on PBPM variation.) Maximum intervention period PBPM costs were $80,189.

Table 7-1
The CGP PBPM mean costs by eligible days, intervention group, demonstration period: Original population

Eligible days[1]	N (%)	PBPM Cost	Range
< 10	6 (0.2%)	$13,812	$0-35,706
11-30	16 (0.6)	11,538	0-49,977
31-60	23 (0.8)	11,237	60-47,877
61-90	40 (1.4)	9,520	0-41,399
91-365	261 (9.0)	7,971	0-80,189
366-730	270 (9.3)	5,481	0-46,137
731-1,095	276 (9.6)	3,740	0-23,903
>1,095	1,999 (69.2)	1,844	0-38,902
Mean	2,891	3,177	0-80,189

NOTES: Observations unweighted. MMC's CGP = Montefiore Medical Center's Care Guidance 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 (5/19/10).

Table 7-2 shows the unweighted cost effects of short-term eligible beneficiaries in the *refresh* population. Again, short-eligibility beneficiaries were several times as costly per month as those with more than 1 year's eligibility. Maximum PBPM costs were $47,596. Note that mean costs in the refresh intervention population are roughly one-half of the original intervention population costs. This is primarily due to not using a minimum cost threshold in the refresh group.

Table 7-2
The CGP PBPM mean costs by eligible days, intervention group, demonstration period: Refresh population

Eligible days[1]	N (%)	PBPM	Range
< 10	2 (0.2)	$4,140	$1,826-6,455
11-30	1 (0.1)	—	—
31-60	11 (1.2)	9,249	0-47,596
61-90	5 (0.6)	13,214	1,016-42,693
91-365	60 (6.7)	6,555	0-35,219
366-730	73 (8.2)	5,403	6-26,658
730-1,095	101 (11.3)	3,671	29-19,535
>1,095	643 (71.8)	1,675	0-19,600
Mean	896	2,692	0-47-596

NOTES: Observations unweighted. CGP = Care Guidance 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 (4/20/10).

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

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, 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

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. For MMC's CGP, its negotiated management fee was $120 for the original intervention group during the first year and $123.84 in years 2 and 3. Fees for the refresh intervention group fees were $123.84 in year 2 and $127.80 in year 3. No monthly fees were paid in the last two months of the demonstration for both intervention and comparison group beneficiaries and fees for intervention beneficiaries who were in the self-directed program were stopped as of September 2008 at the request of the MMC CGP. 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. Lower participation rates produce lower average fees spread across all participating and nonparticipating intervention beneficiaries.

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 group 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, the CGP 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 the CGP 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 population approach. First, we used each beneficiary's own mean PBPM costs in the base year just prior to the CGP'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.[11] 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-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

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

zero to assure CMS that the measured savings rate was not due to chance.[12] 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 the CGP.

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

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

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

A two-sided test[13] 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 population 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 CGP'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 $35,123 in the original comparison group, and $0 to $23,051 in the refresh comparison group. The coefficient of variation (CV), or the standard deviation of beneficiary-level PBPM costs divided by the mean, is fairly large (1.58) in the base year (standard deviations roughly 58% greater than mean costs). CVs in the original and refresh

---

12 Chance savings can occur primarily because of random fluctuations in the utilization of health services 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.

13 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.

comparison populations were quite similar in the base and demonstration periods. Some of the variation is reduced after weighting observations when determining intervention savings later in this chapter.

Table 7-3
The CGP 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)	(1,785)	(2,891)	(1,785)	(2,891)
Minimum	$0	$0	$0	$0
<10%	98	115	223	243
<25%	232	258	568	536
Median	709	662	1,559	1,478
>75%	2,095	2,157	3,480	3,754
>90%	4,426	4,489	6,961	7,586
Maximum	35,123	24,441	54,400	80,189
Mean	1,685	1,689	3,003	3,177
CV	1.58	1.44	1.53	1.59

NOTES: Observations unweighted. CGP = Care Guidance 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 (5/19/10).

Table 7-4
The CGP 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)	(868)	(896)	(868)	(896)
Minimum	$0	$0	$0	$0
<10%	0	0	205	187
<25%	0	95	485	405
Median	542	619	1,190	1,102
>75%	1,994	2,104	3,418	3,318
>90%	4,848	5,069	7,068	6,680
Maximum	23,051	25,720	108,907	47,596
Mean	1,697	1,756	3,030	2,692
CV	1.68	1.64	1.97	1.64

NOTES: Observations unweighted. CGP = Care Guidance 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 (5/19/10).

The difference between median and mean PBPM costs indicates how skewed costs actually are. Mean costs are more than double median costs in the original population's base year, indicating a strong right tail of very high costs. Costs were similarly skewed in the refresh group (***Table 7-4***). Note the low PBPM costs of the lowest 25% of beneficiaries in the base year. Maximum values show how high PBPM costs can be before weighting, e.g., $108,907 per month for one demonstration period beneficiary. 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] = 1.58, ***Table 7-3***), coupled with adjustments for the repeated nature of the

experimental design, the power afforded by the original intervention and comparison population sizes was very low, i.e., about 16%.[14]

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 a 36-month demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. Thus, to avoid paying back any fees for 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 group's PBPM costs, then the 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

---

[14] Power for a comparison of two mean changes in PBPM costs is given by $\Phi[-1.96 + (vn\Delta/(\sigma_d v2)]$ (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, and ρ = correlation between observations between the base and intervention periods. The intervention and comparison standard deviations in the base period were $2,489 and $2,176, respectively. Assuming a .33 intra-patient correlation, σ_d = $3,024. If there were no increase in the comparison group's PBPM over time, then Δ = .05($1,492) = $74.60 (see Table 7-5). The treatment n = 2,891. Thus, power = Φ[-1.96 + ($74.60•53.76/3,024•1.41= .938) = 1 – Φ[1.02] = .16. With the CGP's intervention population, we had 16% likelihood of finding a significant difference if the true mean change in the intervention PBPM was $74.60 less than the change in the comparison PBPM cost.

intervention had more success with those beneficiaries it actively engaged, then savings should be greater for participants than nonparticipants.

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
The CGP 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	2,891	$1,443	46.3	$2,222	56.8	$779**	58.5
Participants	2,159	1,448	48.2	2,296	65.2	808**	67.4
Nonparticipants	732	1,280	77.5	1,953	116.1	673**	117.8
Comparison	1,785	1,492	51.5	2,263	70.7	792**	74.6
Differences							
I - C	—	-29	66.5	-41	91.4	-13	95.0
Participants - C	—	16	71.2	33	96.6	17	100.9
Nonparticipants - C	—	-191*	97.4	-310*	136.9	-119	142.9
Participants - Nonparticipants	—	208*	100.1	343**	137.8	135	142.0

NOTE: CGP = Care Guidance 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(5/19/10).

Overall. The weighted base year average PBPM cost was -$29 (2%) less (*p*=insig) in the intervention group versus the comparison group ($1,443 versus $1,492). The intervention-comparison difference in PBPM Medicare costs increased slightly to -$41 (*p*=insig) in the demonstration period ($2,222 versus $2,263). Intervention beneficiaries remained 1.8% 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 $792 (*p*<.01), while the intervention group's PBPM average Medicare costs rose more slowly by $779 (*p*<.01). Consequently, the intervention group's PBPM cost rose -$13 more slowly (*p*=insig) than the comparison group's PBPM cost.

Participation Status. The participation rate, based on beneficiaries used in this cost analysis, was 75% (2,159/2,891 - 1). Participant costs in the CGP's intervention group were 1% higher ($16; p=insig) than in the comparison group in the base period. Nonparticipants were $191 less costly (p = <.05). Participant costs rose slightly relative to comparison costs over the demonstration period. Nonparticipants became -$310 less costly (p<.05) during the demonstration period versus -$191 in the base period. Thus, the -$13 slower growth in intervention PBPM costs appears to be due entirely to slower growth in the nonparticipant group (-$119; p=insig).

7.4.2 Refresh Population

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 $42 more (*p*=insig) in the

intervention versus comparison group ($1,589 versus $1,547). The intervention-comparison gap in PBPM Medicare costs reversed (-$82; p=insig) in the demonstration period ($2,154 versus $2,236). The average comparison group PBPM increased $688 (p<.01) while the intervention group's PBPM average Medicare costs increased $565 (p<.01). As a result, the intervention group's PBPM cost increased -$124 slower (*p*=insig) compared with the comparison group's PBPM cost. Intervention beneficiaries, who were 2.7% more costly at baseline, were 3.7% less costly than the comparison group, on average, after 18 months between the mid-points of the baseline and demonstration periods.

Table 7-6
The CGP PBPM cost growth 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	896	$1,589	88.9	$2,154	106.3	$565**	122.0
Participants	725	1,557	94.5	2,215	117.2	657**	132.2
Nonparticipants	171	1,759	245.7	1,833	252.2	75	312.7
Comparison	868	1,547	92.4	2,236	107.0	688**	117.1
Differences							
I - C	—	42	128.2	-82	150.9	-124	169.5
Participants - C	—	10	132.4	-21	158.2	-31	175.7
Nonparticipants - C	—	211	246.0	-402	279.0	-614*	312.2
Participants - Nonparticipants	—	-201	243.2	381	290.6	583	333.4

NOTE: CGP = Care Guidance 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 (5/19/10).

The participation rate, based on beneficiaries used in the refresh cost analysis, was 81% (725/896 – 1). Participants in the base period in the CGP's intervention group were $10 more costly (p=insig) than comparison group beneficiaries and nonparticipants were $211 more costly (p=insig). Participants became -$21 less costly (p=insig) during the demonstration period. Non-participants became -$402 less costly (p=insig) during the demonstration period. Consequently, the participant group's PBPM cost rose -$31 slower (p=insig) than the comparison group's cost while the nonparticipant group's PBPM cost rose -$614 more slowly (p<.05) than the comparison group's PBPM cost. Thus, the -$124 in gross savings in the refresh population appears to be due to unexplained changes in costs among intervention beneficiaries not engaged in the intervention.

7.5 Savings and Budget Neutrality

7.5.1 Original Population

Table 7-7 presents summary statistics on savings from the CGP'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 CGP'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 $779 in the intervention group and $792 in the comparison group. The result was a -$13 relative decrease in PBPM cost growth in the intervention group. This negative difference implies *gross savings* at a rate of -0.6% of the comparison group's demonstration period PBPM cost. These savings were statistically insignificant.

The minimal detectable savings threshold was -$186 at the 95% confidence level. This threshold rate was 8.2% 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.[15]

MMC's average monthly fee for the CGP was $90 when averaged over all intervention beneficiaries, which amounted to 4% of the comparison group's PBPM cost during the demonstration period. It was much lower than the stated fee because it was paid on less than all intervention beneficiaries. Thus, the CGP would have had to achieve 9% (4% + 5%) savings in order to retain all of its fees—at least according to RTI's calculations, which are not official under financial reconciliation.

[15] If minimal savings were based just on differences in PBPM costs during the demonstration period, the intervention would have to achieve a 6.1% savings rate (70.7(1.96)/$2,263) based on RTI's weighting method.

Table 7-7
The CGP average PBPM gross savings, fees, and budget neutrality status: Original population

Description	PBPM cost change
Intervention group	$779
Comparison group	792
Difference	-$13
Gross (dis)saving %[1]	-0.6%
Minimal Detectable Savings[2]	
Absolute	-$186
% of comparison PBPM[3]	-8.2%
Monthly Fee	
Absolute[4]	$90
% of comparison PBPM[3]	4.0%
Net Fee	
Absolute[5]	$77
% of comparison PBPM[3]	3.4%

NOTES: CGP = Care Guidance Program; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison demonstration PBPM (= $2,263). 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 ($2,263) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of $120, $123.84, $0.00 (last 2 months for all intervention beneficiaries and for self-directed beneficiaries as of September 2008) fees paid in outreach period and thereafter through month 36. Weights = fee-eligible member months.

[5] Absolute Net Fee = Absolute Monthly Fee + Difference in PBPM outlay change.

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 MMC Phase I, May 14, 2010, Tables 3, 5 and 6.

7.5.2 Refresh Population

Table 7-8 presents summary statistics on savings from the CGP intervention with the refresh population. Over the course of the 24-month intervention, average monthly costs increased $565 in the intervention group and $688 in the comparison group. The result was a -$124 lower relative increase in PBPM costs in the intervention group. This negative difference implies *gross savings* at a rate of 5.5% of the comparison group's PBPM cost.

Table 7-8
The CGP average PBPM gross savings, fees, and budget neutrality status: Refresh population

Description	PBPM cost change
Intervention group	$565
Comparison group	$688
Difference	-$124
Gross (dis)saving %[1]	-5.5%
Minimal Detectable Savings[2]	
Absolute	-$332
% of comparison PBPM[3]	-14.9%
Monthly Fee	
Absolute[4]	$99
% of comparison PBPM[3]	4.4%
Net Fee	
Absolute[5]	-$25
% of comparison PBPM[3]	-1.1%

NOTES: CGP = Care Guidance Program; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison PBPM (= $2,236). 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 ($2,236) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of $123.84, $0.00 (last 2 months for all intervention beneficiaries and for self-directed beneficiaries as of September 2008) fees paid in outreach period and thereafter through months 13-36. Weights = fee-eligible member months.

[5] Absolute Net Fee = Absolute Monthly Fee + Difference in PBPM outlay change.

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 MMC Phase I, May 14, 2010, Tables 3, 5 and 6.

With less than 1,000 beneficiaries in each of the intervention and comparison refresh groups, the minimal detectable refresh savings threshold was -$332 at the 95% confidence level. This rate is -14.9% 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. Ignoring the fact that the -$124 in intervention savings was not statistically different from zero, the net fee to Medicare was reduced from $99 per beneficiary per month to -$25, resulting in a net Medicare cost of -1.1% of the comparison group's average monthly outlay on claims.

7.6 Imbalances between Intervention and Comparison Populations

Because the CGP comparison group was not based on a randomized design, it is possible that material imbalances remained between study and comparison groups simply by chance. If the distribution of high-cost and high-risk beneficiaries differs between the CGP 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 $25,000/month, or the top 25% of cases in either population based on their costs the year prior to selection. The HCC high-risk threshold was set at 2.0.

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. Intervention beneficiaries, compared with comparison beneficiaries, were less likely to be non-White or eligible for Medicaid. They were more likely to have been in a SNF during the base year. They appear equally balanced in cost and slightly less likely to be just high risk.

Table 7-9
The CGP frequency distribution of beneficiary characteristics, intervention and comparison groups, base year: Original population

Characteristics	Intervention (%)	Comparison (%)
Cost-Risk Group		
High-cost > =$ 25,000	8.6%	8.2%
Both	12.5	13.6
High-risk: HCC > 2.0	8.1	11.6
Neither	70.8	66.7
Age Group		
<65	6.1	9.1
65-69	8.8	10.0
70-74	15.8	17.7
75-79	22.0	19.9
80-84	22.2	20.0
85+	24.1	23.2
Gender		
Female	60.8	66.4
Male	39.2	33.6
Race		
Non-White	32.5	59.6
White	67.5	40.4
Medicaid Eligible		
No	92.1	80.8
Yes	7.9	19.2
Disabled		
No	94.0	91.1
Yes	6.0	8.9
Urban residence		
No	0.0	0.0
Yes	100.0	100.0
Long-term care		
No	99.9	99.9
Yes	0.1	0.1
SNF		
No	92.6	95.6
Yes	7.4	4.4
Institutionalized		
No	96.7	97.8
Yes	3.3	2.2

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. CGP = Care Guidance Program; HCC = Hierarchical Condition Category; SNF = skilled nursing facility; Institutionalized = in nursing home 1+ months during demonstration's first 6 months.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/7/10).

Table 7-10
The CGP frequency distribution of beneficiary characteristics, intervention and comparison groups, base year: Refresh population

Characteristics	Intervention (%)	Comparison (%)
Cost-Risk Group		
High-cost > =$ 25,000	10.1%	8.1%
Both	12.5	13.3
High-risk: HCC > 2.0	8.2	9.6
Neither	69.2	69.0
Age Group		
<65	9.3	9.3
65-69	13.9	12.0
70-74	17.6	16.4
75-79	20.4	20.1
80-84	18.1	21.0
85+	19.7	21.2
Gender		
Female	63.5	67.0
Male	36.5	33.0
Race		
Non-White	39.7	59.2
White	60.3	40.8
Medicaid Eligible		
No	92.6	83.5
Yes	7.4	16.5
Disabled		
No	90.6	90.4
Yes	9.4	9.6
Urban residence		
No	0.0	0.0
Yes	100.0	100.0
Long-term care		
No	99.9	100.0
Yes	0.1	0.0
SNF		
No	92.8	96.5
Yes	7.2	3.5
Institutionalized		
No	98.8	98.5
Yes	1.1	1.5

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. CGP = Care Guidance Program; HCC = Hierarchical Condition Category; SNF = skilled nursing facility; Institutionalized = in nursing home 1+ months during demonstration's first 6 months.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/7/10).

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 $1,022 in the base year compared with $3,500 for high-cost only beneficiaries (3.5 times greater) and both high- cost and high-risk beneficiaries ($5,530; 5.5 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 or 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 find no statistically significant differences between the original intervention and comparison groups' growth rates.

7.6.2.2 Refresh Population

Table 7-12 presents results on PBPM cost trends by the four cost-risk groups for the refresh population. The high-cost & high-risk refresh group and the high-cost only refresh group showed costs declining faster among the intervention beneficiaries than the comparison beneficiaries. PBPM costs in the high-risk only and neither group increased over time. However, none of the four cost-risk groups showed statistically significant differences in intervention versus comparison group rates. The large standard errors for the refresh population are noteworthy. We had little power to detect savings rates even as large as several hundred dollars per month given the small group sizes and high cost variance from year to year.

7.7 Regression-to-the-Mean (RtoM)

Tables 7-13 and ***7-14*** demonstrate RtoM effects 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 $1,685 in the comparison group's base period in the original population (***Table 7-13***), with an overall increase of $1,318. Cost increases are inversely correlated with a beneficiary's base period PBPM costs. At the extremes, beneficiaries with less than $250 in base period PBPM costs saw their average costs increase by $1,825 while those with initial costs greater than $4,000 experienced average decreases of $1,662. Mean costs in both periods are well above median costs and indicate a strong skewness in PBPM costs. Also note that any RtoM effects are dominated by the strong upward trend in costs in the New York area with few high cost groups exhibiting any cost reductions.

Regression-to-the-mean effects are more obvious in the refresh population (***Table 7-14***). Unweighted mean costs increased $1,332 due mostly by much larger cost increases for beneficiaries with base year costs under $250 per month. This suggests that for the intervention to be successful, it would need to identify initially low cost beneficiaries most likely to experience major cost increases.

Table 7-11
The CGP 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)	(486; 17%)	—	(257: 9%)	—	(252: 12%)	—	(1,896; 66%)	—
Base Year	$5,530	154.0	$3,500	101.7	$1,022	36.1	$519	10.7
Demonstration	4,586	224.9	3,007	219.3	2,692	202.3	1,654	50.1
Difference	-944**	251.0	-493*	242.7	1,670**	200.2	1,135**	49.5
% Change	-17%	—	-14%	—	163%	—	219%	—
Comparison (N)	(301; 17%)	—	(142; 8%)	—	(223; 12%)	—	(1,319; 63%)	—
Base Year	5,170	185.9	3,759	135.6	983	40.8	523	14.2
Demonstration	4,025	245.3	2,709	273.1	2,939	264.9	1,723	64.4
Difference	-1,094**	274.3	-1051**	303.0	1,956**	261.5	1,200**	63.7
% Change	-21%	—	-28%	—	199%	—	229%	—
Difference-in-Differences	151	382.1	558	513.9	-286	325.1	-64	81.3

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. CGP = Care Guidance 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 $25,000 in base period (top 25%).

High-Risk: HCC > 2.0 in base period.

% Change: Difference/Base Year.

*p<.05; **p<.01

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/7/10).

Table 7-12
The CGP 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)	(1,373; 15%)	—	(90; 10%)	—	(70: 8%)	—	(599; 69%)	—
Base Year	$6,348	341.8	$3,997	241.0	$965	64.3	$453	21.9
Demonstration	3,969	356.6	2,380	344.0	2,226	379.0	1,785	115.9
Difference	-2,378**	467.2	-1,617**	352.1	1,261**	389.1	1,332**	116.7
% Change	-37%	—	-40%	—	131%	—	294%	—
Comparison (N)	(136; 16%)	—	(170; 8%)	—	(83; 10%)	—	(599; 66%)	—
Base Year	6,213	374.4	4,086	227.4	1,040	68.2	417	23.1
Demonstration	4,105	364.8	2,758	445.8	2,521	379.9	1,773	107.8
Difference	-2,108**	396.3	-1,328**	508.2	1,481**	368.4	1,356**	109.9
% Change	-34%	—	-33%	—	142%	—	325%	—
Difference-in-Differences	-270	612.9	-289	600.3	-221	535.4	-24	160.8

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. CGP = Care Guidance 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 $25,000 in base period (top 25%).

High-Risk: HCC > 2.0 in base period.

% Change: Difference/Base Year.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/7/10).

Table 7-13
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	478	$125	$1,951	$1,825
251-500	281	358	1,690	1,333
501-750	158	621	2,869	2,248
751-1,000	121	867	2,847	1,980
1,001–1,250	99	1,113	2,573	1,440
1,251-1,500	81	1,370	2,737	1,367
1,501-1,750	69	1,631	4,378	2,747
1,751-2,000	34	1,886	4,685	2,799
2,001-2,250	48	2,120	4,297	2,177
2,251-2,500	37	2,382	2,786	404
2,501-2,750	41	2,626	5,382	2,756
2,751-3,000	29	2,869	4,270	1,401
3,001-3,250	26	3,117	2,902	-215
3,251-3,500	26	3,371	5,397	2,026
3,501-3,750	17	3,613	3,602	-11
3,751-4,000	28	3,884	4,084	200
> 4,000	212	7,192	5,530	-1,662
Total/Mean	1,785	1,685	3,003	1,318
Median	—	909	1,559	850

NOTES: Observations unweighted. CGP = Care Guidance Program; PBPM = per beneficiary per month; N = number of beneficiaries in base year.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/3/10).

Table 7-14
Regression-to-the-Mean in comparison group PBPM costs:
Refresh population

Base year PBPM cost level	N	Base year PBPM costs	Demonstration period PBPM costs	Change
< $250	545	$39	$2,899	$2,859
251-500	84	366	1,414	1,048
501-750	52	630	1,365	735
751-1,000	29	873	1,607	734
1,001–1,250	46	1,139	2,894	1,755
1,251-1,500	43	1,362	2,519	1,157
1,501-1,750	32	1,627	2,072	445
1,751-2,000	21	1,883	3,795	1,913
2,001-2,250	20	2,107	2,467	360
2,251-2,500	15	2,375	4,213	1,888
2,501-2,750	14	2,609	3,588	980
2,751-3,000	14	2,896	2,907	11
3,001-3,250	6	3,120	4,859	1,737
3,251-3,500	10	3,388	2,223	-1,164
3,501-3,750	11	3,577	4,938	1,362
3,751-4,000	9	3,884	3,189	-695
> 4,000	117	7,575	5,711	-1,864
Total/Mean	868	1,697	3,029	1,332
Median	—	542	1,190	648

NOTES: Observations unweighted. CGP = Care Guidance Program; PBPM = per beneficiary per month; N = number of beneficiaries in base year.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (6/3/10).

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 Intervention coefficient indicates the average difference in cost growth of the intervention versus the comparison group.

In the first column of results controlling only for each beneficiary's base period PBPM cost, the Intervention coefficient of -29 is statistically insignificant implying no reliable success on the part of the intervention in slowing beneficiary cost increases. The base period PBPM cost coefficient (0.424; $p < .01$), when combined with the intercept coefficient, implies modest RtoM effects on costs (0.424 - 1 = -0.576, 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 264% during the intervention period, while the "high cost" beneficiary's PBPM cost would decline by almost 12%.[16] Whereas cost differences were 8:1 in the base period, they would now be compressed to 1.9:1.

RtoM effects are quite substantial but not in one direction. Including only high cost beneficiaries in the original population 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 only marginally affected when applying the controls and still insignificant. This is true even though all the three cost-risk groups experienced much larger cost increases relative to the neither group. Minor changes in the two intervention coefficients are due to the few imbalances between the intervention and comparison groups. The PBPM base coefficient is smaller, implying more RtoM within each of the cost-risk groups.

---

[16] The calculation is as follows based on Table 7-15, column 1:

PBPM[base]	PBPM[demo]	PBPM Change	%Change
$500	$1,820	$1,320	+264%
$4,000	$3,516	-$484	-12%

Table 7-15
The CGP Regression results: Intervention gross savings controlling for base period PBPM cost and beneficiary characteristics: Original population

Independent Variable	Model 1 PBPM_ Demo Coefficient	Model 1 t-stat	Model 2 PBPM_ Demo Coefficient	Model 2 t-stat	Model 3 PBPM_ Demo Coefficient[a]	Model 3 t-stat
Intercept	1,608	26.8	676	0.6	660	0.6
Intervention	-29	0.3	39	0.4	31	0.4
PBPM_Base	0.424	22.2	0.311	9.7	0.302	9.6
High-cost–high risk	—	—	1,367	6.9	1,266	6.6
High-cost	—	—	446	2.5	403	2.3
High-risk	—	—	974	6.7	940	6.6
Male	—	—	108	1.2	94	1.1
Non-White	—	—	30	0.3	18	0.2
Age 65-69	—	—	796	0.7	832	0.8
70-74	—	—	642	0.6	663	0.6
75-79	—	—	672	0.6	661	0.6
80-84	—	—	887	0.8	901	0.8
85+	—	—	948	0.8	890	0.8
Medicaid	—	—	87	0.7	119	0.9
Disabled	—	—	764	0.7	786	0.7
SNFB	—	—	-681	3.4	-744	3.7
Institutionalized	—	—	—	—	510	2.1
R^2	.095	—	.115	—	0.12	—
N	4,675	—	4,675	—	4.441	—

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. PBPM = per beneficiary per month; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries. Institutionalized = in nursing home 1+ months during demonstration's first 6 months.

Observations weighted by beneficiary's fraction of eligible days during demonstration.

[a] Excludes beneficiaries who died in demonstration's first 6 months.

PBPM_Demo: Dependent variable: Beneficiary's average PBPM cost during demonstration.

PBPM_Base: Beneficiary's average PBPM cost in base period just prior to start date.

High-Cost-High-Risk: PBPM cost > $25,000 and HCC > 2.0 in base year.

High-Cost: PBPM cost > $25,000 and HCC< 2.0.

High-Risk: PBPM cost < $25,000 and HCC > 2.0.

SNFB = 1 if beneficiary had SNF payments in base year.

Institutionalized = 1 if beneficiary had 1+ months in custodial nursing home in demonstration's first 6 months.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (6/7/10); Costrun3a(8/18/10).

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 -95 is insignificant, implying no statistical

Table 7-16
The CGP Regression results: Intervention gross savings controlling for base period PBPM and beneficiary characteristics: Refresh population

Independent variable	Model 1 PBPM_ Demo Coefficient	Model 1 t-stat	Model 2 PBPM_ Demo Coefficient	Model 2 t-stat	Model 3 PBPM_ Demo Coefficient[a]	Model 3 t-stat
Intercept	1,647	15.0	3,015	1.8	2,869	1.9
Intervention	-95	0.7	-65	0.4	-52	0.4
PBPM_Base	0.319	11.8	0.264	6.0	0.259	6.2
High-cost–high risk	—	—	653	2.0	522	1.7
High-cost	—	—	-209	0.7	-165	0.6
High-risk	—	—	399	1.5	392	1.6
Male	—	—	223	1.4	243	1.7
Non-White	—	—	38	0.3	45	0.3
Age 65-69	—	—	-1,516	0.9	-1,438	0.9
70-74	—	—	-1,461	0.9	-1,333	0.9
75-79	—	—	-1,718	1.0	-1613	1.0
80-84	—	—	-1,521	0.9	-1,382	0.9
85+	—	—	1,383	0.8	-1,287	0.8
Medicaid	—	—	154	0.7	124	0.6
Disabled	—	—	-1,130	0.7	-1,053	0.7
Urban	—	—	—	—	—	—
LTCB	—	—	—	—	—	—
SNFB	—	—	101	0.3	-136	0.4
Institutionalized	—	—	—	—	912	1.4
R^2	.073	—	.083	—	.089	—
N	1,763	—	1,763	—	1,684	—

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. PBPM = per beneficiary per month; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries. Institutionalized = in nursing home 1+ months during demonstration's first 6 months.
Observations weighted by beneficiary's fraction of eligible days during demonstration.
[a] Excludes beneficiaries who died in demonstration's first 6 months.
PBPM_Demo: Dependent variable: Beneficiary's average PBPM cost during demonstration.
PBPM_Base: Beneficiary's average PBPM cost in base period just prior to start date.
High-Cost-High-Risk: PBPM cost > $25,000 and HCC > 2.0 in base year.
High-Cost: PBPM cost > $25,000 and HCC< 2.0.
High-Risk: PBPM cost < $25,000 and HCC > 2.0.
SNFB = 1 if beneficiary had SNF payments in base year.
Institutionalized = 1 if beneficiary had 1+ months in custodial nursing home in demonstration's first 6 months.
SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1(6/7/10); Costrun3a(8/19/10).

cost trend differences between intervention beneficiaries. The base period PBPM cost coefficient (0.319; $p < .01$), when combined with the intercept coefficient, again implies more substantial RtoM of costs in the refresh population (= 0.319 - 1 = -0.681, the RtoM effect).

The second regression model controls for cost-risk group and other patient characteristics determined during the base period. The Intervention coefficient remains insignificant. Most all control variables are statistically insignificant, partly because of the smaller populations in the refresh group. Age and disabled coefficients are unreliable because of the high correlation between the disabled and under-age 65 reference group in the intercept.

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. With only roughly 2,900 original and 900 refresh intervention beneficiaries and 1,800 original and 900 refresh comparison beneficiaries, we had limited our power to detect significant savings. Gross savings had to be 8.2% in the original intervention population and 14.9% in the refresh intervention population to be considered significant at the 95% confidence level.

No statistically significant savings were detected for the intervention in the original population. Costs rose -$13 slower in the original intervention group (0.6 % of comparison costs), but savings needed to exceed 8.2% to be considered statistically significant. The CGP, overall, performed better with its refresh population as gross savings averaged -$124 (5.5% of comparison monthly costs). This savings level, however, did not achieve statistical significance.

Intervention and comparison groups were somewhat unbalanced. Intervention beneficiaries were less likely to be non-White or Medicaid-eligible. Beneficiaries in the original sample were more likely to be in a nursing home than comparison beneficiaries, but the one percentage point difference was immaterial. Controlling for imbalances had little effect on our overall final conclusion of no significant savings.

The CGP involved a select group of high cost, more severely ill beneficiaries. As a result, the comparison group exhibited both rapidly rising costs during the intervention period ($792 in the original and $688 in the refresh group) as well as sizable RtoM effects. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by nearly $2,000. Over the same time period, beneficiaries with costs over $4,000 saw their average costs decline by over $3,500. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period 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. The “low cost” beneficiary problem was exacerbated by the more than one-year lag between selection and start date. Many originally high cost beneficiaries two years prior to start date became much lower cost one year prior to start date.

CHAPTER 8
KEY FINDINGS FROM THE MONTEFIORE MEDICAL CENTER'S (MMC) CARE MANAGEMENT FOR HIGH COST BENEFICIARIES DEMONSTRATION EVALUATION

The purpose of this report is to present the findings from RTI International's evaluation of the MMC Care Guidance Program (CGP). Our evaluation focuses upon three broad domains of inquiry:

1. **Implementation.** To what extent was MMC able to implement its program?
2. **Reach.** How well did the CGP engage its intended audience?
3. **Effectiveness.** To what degree was the CGP 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 the CGP's operations with its original population and 24 months with its refresh population. Our findings are based on the experience of approximately 6,000 ill Medicare beneficiaries split across 4 groups for analysis purposes (original and refresh intervention and comparison groups) limiting statistical power somewhat to detect differences. CMS required RTI to analyze the original and refresh populations separately to be consistent with the financial reconciliation. Doing so allowed us to quantify intervention effects over time as MMC's CGP matured. One drawback to separate analyses of each group is the smaller samples available for statistical testing. Only 2,891 and 896 intervention beneficiaries were available for analysis in the original and refresh intervention groups and 1,785 and 868 comparison beneficiaries in the corresponding original and refresh comparison groups. Wide variation in beneficiary costs over time make precise estimates of program success difficult with such small samples. Key findings presented below are based on the resulting statistical tests at standard 95% confidence levels. To better understand the statistical power underlying RTI's analyses, in subsequent chapters we present detailed statistics including confidence intervals for quality of care and acute care utilization measures and a detectable threshold for cost savings, or the rate of savings that would allow us to reject the null hypothesis of no savings.

Six key findings on participation, intensity of engagement in the CGP, beneficiary satisfaction and experience with care, clinical quality, health outcomes, and financial outcomes have important policy implications for CMS and future disease management or care coordination efforts among Medicare FFS beneficiaries. The CMHCB demonstration program holds MMC financially responsible for financial savings but does not hold MMC financially responsible for quality of care improvements.

Key Finding #1: The CGP did not preferentially engage beneficiaries who were at highest risk of acute clinical deterioration as measured by the concurrent HCC score.

Of the CGP's original intervention beneficiaries, 75% verbally consented to participate in the CMHCB demonstration at some point during the intervention period; 81% of the refresh population agreed to participate. In spite of this fairly high level of participation, we find that beneficiaries from the original population with high baseline HCC scores or medium or high baseline PBPM costs more likely to be participants. Demonstration period health status as measured by the concurrent HCC score had no impact after controlling for baseline health status characteristics and demographics. This suggests that the CGP was able to engage the historically sicker Medicare beneficiaries but did not preferentially engage those with acute clinical deterioration as measured by the concurrent HCC score. None of these measures were statistically significant for the refresh population. However, only 32% of the refresh population, or 386 beneficiaries, were in the reference group making it difficult to determine statistically significant differences.

Key Finding #2: The CGP did not substantially affect beneficiary reported experience with care, level of physical activity, and self-reported physical health.

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 the CGP's intervention: helpfulness of discussions with their health care team and quality of communication with their health care team. Because we used the same survey questions for both the intervention and comparison groups, we did not ask specifically about the helpfulness of discussions with staff of the CGP. Thus, demonstration programs with a telephonic care management approach might be at a disadvantage because beneficiaries might not consider the telephonic care managers as being their health care providers. 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, only one statistically significant positive group difference was detected—members of the CGP's intervention group reported fewer limitations in their activities of daily living than those in the comparison group. This difference, however, was not reflected in another measure of physical health-PHC scores. We did not detect any statistically significant intervention effects on any measures of beneficiary's satisfaction and experience with care, nor on any of the self-management outcomes for the CGP.

Key Finding #3: We did not detect improvement in 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 Medicare 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 did not detect systematic improvement in quality of care among the original or refresh intervention beneficiaries. We generally observed negative trends in rates, not positive, due to either improvement of the rate or less of a decline in the rate of receipt among the comparison beneficiaries. However, we would like to note that claims data are likely to produce underestimates of the rate of influenza vaccination as they do not capture flu vaccines that people receive in pharmacies, supermarkets, senior centers, or city-funded health care centers because services received in those settings do not result in Medicare claims. Although we do not have self-reported rates of influenza vaccines for the comparison group beneficiaries, MMC's CGP reported to RTI that their survey of participants revealed over three-quarters reported having received a flu vaccine during each of the three years of the demonstration.

Key Finding #4: We did not detect systematic reductions in acute care utilization as measured by rate of hospitalization, ER visits, or 90-day readmissions nor did we detect any difference in the use of the Medicare hospice benefit between the intervention and comparison groups.

During the course of CGP, we observed 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. Out of 30 acute care utilization comparisons, we observe one statistically significant positive intervention effect; the rate of all-cause ER visits grows at a slower rate during the last 12 months of the demonstration within the original intervention group than within the comparison group, or -275 per 1,000 beneficiaries ($p=0.02$).

We do not detect any other statistically significant positive intervention effects; however, we do detect one statistically significant negative intervention effect. During months 7 to 18 of the CGP demonstration period, the percent of original intervention beneficiaries with an all-cause hospitalization increased from 38% to 41% while the percent of comparison beneficiaries with an all-cause hospitalization declined from 41% to 40% ($p=0.04$).

We did not detect any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rates of the Medicare hospice benefit or in mean or median number of days of hospice.

Key Finding #5: We did not detect a difference in the rate of or time to death between the original and refresh intervention and comparison beneficiaries.

We did not detect a statistically significant differential rate of mortality between the intervention and comparison groups of the original or refresh population. Over the 36-month demonstration period for the original population, 29% of beneficiaries in the intervention and comparison groups died. During the 24-month demonstration period for the refresh population, 17% of beneficiaries in the intervention and comparison groups died.

Key Finding #6: Medicare cost growth in the intervention group was not statistically different from the rate of growth in the comparison group.

No statistically significant savings were detected for the intervention in the original population. Costs rose -$13 slower in the original intervention group (0.6 % of comparison costs), but savings needed to exceed 8.2% to be considered statistically significant. The CGP, overall, performed better with its refresh population as gross savings averaged -$124 (5.5% of comparison monthly costs). This savings level, however, did not achieve statistical significance. Intervention and comparison groups were somewhat unbalanced in baseline characteristics. Intervention beneficiaries were less likely to be non-White or enrolled in Medicaid. However, controlling for these baseline imbalances in characteristics had no effect on our overall final conclusion of no statistically significant savings. For differences in 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.

8.2 Conclusion

Based on extensive quantitative analysis of performance using statistical tests at standard 95% confidence levels, we did not detect improvement in key processes of care, beneficiary self-reported experience with care, self-management, and functional status, increase in use of the Medicare hospice benefit, or decrease in mortality. We did detect one statistically significant positive intervention effect; the rate of all-cause ER visits grew at a slower rate during the last 12 months of the demonstration within the original intervention group than within the comparison group. However, we also detected one statistically significant negative intervention effect. During months 7 to 18 of the CGP demonstration period, the percent of original intervention beneficiaries with an all-cause hospitalization increased from 38% to 41% while the percent of comparison beneficiaries with an all-cause hospitalization declined from 41% to 40%.

Although PBPM costs rose slower in the original and refresh intervention groups relative to the comparison groups, statistically significant savings were not detected in the *overall* samples. This may have been due to relatively small sample sizes and lack of statistical power. 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 2,900 original and 900 refresh intervention beneficiaries and 1,800 original and 900 refresh comparison beneficiaries, we had limited power to detect significant savings. Gross savings had to be 8.2% in the original intervention population and 14.9% in the refresh intervention population to be considered significant at the 95% confidence level.

What might explain the lack of *overall* program effectiveness? One explanation may be the targeting of beneficiaries at greatest risk of intensive, costly, service use (as distinct from the need for general care management). MMC and CMS agreed upon a predicted costly set of Medicare beneficiaries for their intervention and comparison groups using 2004 claims data. Mean per beneficiary per month base year claims costs (weighted by fraction of time eligible for the intervention) were approximately $1,600 in both groups, a figure considerably higher than in the general Medicare population. During the intervention period, the comparison group

exhibited both rapidly rising costs ($792 in the original and $688 in the refresh group) as well as sizable RtoM effects. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by nearly $2,000. Over the same time period, beneficiaries with costs over $4,000 saw their average costs decline by over $3,500. The large churning of beneficiaries in both the intervention and comparison groups from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period 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. The "low cost" beneficiary problem was exacerbated by the more than one-year lag between selection and start date. Many originally high cost beneficiaries two years prior to start date became much lower cost one year prior to start date.

A second explanation may be the CGP's beneficiary recruitment strategy. Given the program's monthly management fee (roughly $120 per month) and the population-based financial risk feature of this demonstration, less than full engagement of the intervention population required the CGP to have been extremely successful in reducing costs associated with the participating beneficiaries. The CGP was not successful in broadly reducing hospitalizations during the demonstration period. The lack of substantive improvements in acute care utilization broadly across their intervention population translated into limited financial savings.

And, a third explanation may be the model of intervention itself. Prior evaluations of Medicare care management programs that were primarily telephonic have not demonstrated savings sufficient to cover fees similar to the CGP's fee. A cornerstone of the CGP was health coaching interactions with care manager nurses. However, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments, and the most dominant form of contact was telephonic.

Furthermore, CGP nurse care managers were not part of the beneficiaries' primary health care teams, thereby hindering their interactions with the beneficiaries' primary providers, changing medical care plans, or mitigating deterioration in health status. And, not all intervention beneficiaries had primary care physicians at MMC, therefore the care managers had to interact with nonMMC providers with whom they had little or no prior relationship.

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.