April 2011

Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: The Health Buddy® Consortium (HBC)

Revised Final Report

Prepared for

Lorraine Johnson, Sc.D.
Centers for Medicare & Medicaid Services
Office of Research, Development, and Information
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Nancy McCall, Sc.D.
Jerry Cromwell, Ph.D.
Kevin Smith, M.S.
Carol Urato, M.A.
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: THE HEALTH BUDDY® CONSORTIUM (HBC)

by Nancy McCall, Sc.D.
Jerry Cromwell, Ph.D.
Kevin Smith, M.S.
Carol Urato, M.A.

Federal Project Officer: Lorraine Johnson, Sc.D.

RTI International*

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

April 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 the Health Buddy® Consortium's (HBC) Care Management for High Cost Beneficiaries (CMHCB) demonstration program. The HBC CMHCB demonstration program was run by a consortium of four organizations collaborating to deliver care management services to high-cost Medicare beneficiaries with diabetes, heart failure (HF), and/or chronic obstructive pulmonary disease (COPD). The consortium was coordinated by Health Hero Network (HHN) [acquired by Robert Bosch Healthcare (RBHC) in December 2007], the American Medical Group Association (AMGA), Bend Memorial Clinic (BMC) in Central Oregon, and Wenatchee Valley Medical Center (WVMC) in North Central Washington.

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 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 the HBC program 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 (CMOs), the HBC program 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 the HBC program 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 HBC's ***effectiveness*** in engaging the ***full*** group of pre-identified FFS beneficiaries. Consequently, the HBC's relatively low participation rates (40-45%) 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:

- Implementation. To what extent was the HBC able to implement its program?
- Reach. How well did the HBC engage its intended audiences?
- Effectiveness. To what degree did the HBC improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings?

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

E.1 Scope of Implementation

The HBC launched its program on February 1, 2006. The HHN and AMGA worked closely with its CMS project officer and analysts from Actuarial Research Corporation (ARC) to develop a methodology for selecting the starting population for the HBC CMHCB program. Beneficiaries had to meet all of the following inclusion criteria:

- Medicare FFS beneficiaries with a primary residence in designated counties of Oregon and Washington, with high costs in 2004 (i.e., Medicare costs greater than or equal to $6,000 in 2004) **or** high disease severity as indicated by Hierarchical Condition Categories (HCC) risk scores greater than or equal to 1.7,
- A plurality[1] of visits or two or more visits to the HBC medical practices as evidenced by one or more claims in 2004 associated with a provider at either BMC or WVMC, and
- One or more claims associated with diagnosis codes for at least one of the following conditions: HF, diabetes, or COPD.

The population was further restricted using the following exclusion criteria: receiving hospice care or the end-stage renal disease (ESRD) benefit, enrolled in a Medicare Advantage plan or have Medicare as a secondary payer or lack Medicare Part A or Part B coverage as of January 2, 2006, and at least one of the exclusionary diagnoses designated by the HBC (e.g., dementia or spinal cord disease). The remaining beneficiaries were randomly assigned to the intervention group. The randomization was done on a state-by-state basis by alphabetizing eligible beneficiaries by their last name. BMC was assigned 661 patients and WVMC was assigned 965 patients for a total of 1,626 beneficiaries.

[1] Plurality refers to receiving more care from one provider (e.g., BMC) than any other provider.

Following the development of the intervention group criteria, the HBC worked with CMS and RTI to develop specifications to select a comparison group. The methodology was an approach that identified Physician Group Practices (PGPs) identified by their Tax Identifications Numbers (TINs) that would be appropriate for comparison purposes. RTI then applied the exclusion, inclusion, and loyalty criteria to select beneficiaries associated with the designated TINs (See Section 1.2.3 for details). In Oregon, this method identified 17 TINs and 1,925 beneficiaries in 9 comparison counties. In Washington, the 18 TINs yielded 1,569 beneficiaries in 12 counties. Beneficiaries in these practices and in the Oregon and Washington intervention clinics were then stratified using a 9-cell diagnostic and cost matrix. The final step in the process was to draw the sample of comparison group beneficiaries, done by randomly choosing beneficiaries from the comparison pool to match the number in the corresponding cell in the intervention group. This produced a final group of 660 comparison beneficiaries in Oregon and 964 in Washington, the same as the size and distribution of the intervention groups in each state. The combined per beneficiary per month (PBPM) cost during the baseline period for the Washington and Oregon intervention groups was $1,289 PMPM. For the combined comparison groups, the cost was $1,280 PMPM.

After program implementation, the HBC requested from CMS reconsideration of its intervention population. The HBC requested removal of selected beneficiaries (carve-out) from its starting population and the addition of new beneficiaries using alternative inclusion/exclusion criteria at the time of their planned refreshment of their intervention and comparison populations. Of the starting comparison population, 788 beneficiaries (49%) were identified for the carve-out. The final refresh intervention population was 1,056 beneficiaries . The total number of comparison refresh beneficiaries drawn was 1,056 or the same size as the intervention refresh group with 322 comparison beneficiaries from Oregon and 734 comparison beneficiaries from Washington.

Of the HBC's original intervention group beneficiaries, 45% verbally consented to participate in the CMHCB demonstration at some point during the intervention period, 48% refused to participate, and 6% were not contacted or were unable to be located. Of the refresh intervention beneficiaries, 40% consented to participate at some point during the 26-month period. The percent that refused to participate was modestly lower (43%), the percent that were not contacted or were unable to be contacted increased to 18%. The HBC program ended March 31, 2009 or 38 months after the start date for the original population and 26 months after the start date for the refresh population.

The HBC negotiated a management fee of $120 for the original intervention group during the first year, $123.84 in year 2, and $127.80 in year 3. Fees for the refresh intervention group were $123.84 in year 2 of the overall study and $127.80 in year 3 of the overall study. Fees were paid on a monthly basis for all beneficiaries who did not opt out during the 6-month outreach period and remained eligible for the demonstration. The net savings requirements for the HBC program are 5% for the original cohort and 2.5% for the refresh cohort.

E.2 Overview of the HBC CMHCB Demonstration Program

The core of the intervention was a care management program augmented by the use of a Health Buddy® device. Once a beneficiary consented to participate in the HBC CMHCB

program, he or she was asked if they would be willing to receive a Health Buddy® device that would allow for routine communication with program staff through daily sessions that span 5 to 10 minutes. During each session, patients were expected to answer a series of questions related to symptoms, vital signs, knowledge, and health behaviors and receive educational information about their health conditions. Each device was set up with one of 13 disease-specific programs that most closely address the needs of each participant. For example, patients with diabetes participated in dialogs that focused primarily on issues related to diabetes. Four of the programs addressed comorbid conditions, such as COPD and diabetes, or COPD, HF, and diabetes. Nurse care managers at each of the program sites monitored patient responses to Health Buddy® questions on a daily basis using a web-based computer application, the Health Buddy® desktop. This system helped care managers determine the urgency of follow-up required for each patient, triage patients accordingly, and coordinate the provision of appropriate medical, psychological, or social services.

For individuals who did not have a chronic condition, there was a 14th program, a Senior Wellness program that addressed issues related to general health and safety, as well as psychosocial issues, such as depression. Patients who were unable or unwilling to use the Health Buddy® device had the opportunity to participate in the Health Buddy® program through routinely scheduled telephone calls with nurse care managers that occurred weekly, bi-weekly, or monthly depending on the patient's health status. This is referred to as the alternate program.

The Health Buddy® disease management program content is drawn from evidence-based practice guidelines, and each program is designed to collect standard outcome measures including utilization, patient satisfaction, quality of life, and compliance with treatment regimens. Advisory boards at each of the participating medical practices reviewed the Health Buddy® disease management programs and developed associated care protocols and care plans to guide care managers responses to alerts associated with each disease-specific program.

The Health Buddy® Program provided physicians with information about patient symptoms, vital signs, and behaviors during the time period between office visits. As a result, providers had the opportunity to intervene with patients when they had early symptoms of health problems, potentially avoiding hospitalizations or emergency room visits. Further, physicians could review trends in patients' Health Buddy® responses prior to scheduled office visits, which could help them to identify health issues that required attention.

During the first site visit, physicians at both sites reported that they were initially very enthusiastic about the Health Buddy® program, because it offered a promising way to effectively support patients with chronic disease. The Health Buddy® technology coupled with telephonic care management support was viewed as an effective way to maintain and improve patient health and identify symptoms of complications early, so that timely medical intervention could be used to prevent serious problems requiring hospitalization. Once the physicians received the list of patients who were eligible for the Health Buddy® program, they reported that they became frustrated with the project because they felt that many of the patients selected would not benefit from participating. Further, physicians reported disappointment that many of the patients they believed could be helped by the program were not eligible to participate in the program based upon the claims-based algorithm developed by HHN.

Using information gleaned from its early experience with the program, the HBC made a series of changes and enhancements to its operations as reported to us at our second site visit. Over the course of the demonstration, there were no major changes in the overall programmatic approach; however, there were enhancements focused on tailoring the Health Buddy® programs and dialogues to meet the needs of participants with multiple comorbidities and to maintain participant interest. The HBC program staff reported that they had made adjustments to effectively address the multiple comorbidities characteristic of the high-cost Medicare population. The content of most of the disease management programs deployed via the Health Buddy® device in the HBC program addressed one particular health condition. However, Medicare beneficiaries with high health care utilization typically have multiple comorbid conditions. Feedback obtained during the site visit indicated that care managers were providing telephonic support for conditions that were not addressed by the 14 programs then available to the HBC CMHCB population. As a result, the HBC program was in the process of introducing additional disease management programs in an effort to address the needs of as many members of the intervention population as possible. The care managers at each of the clinics also expanded their focus to pay greater attention to social issues that affected their participants' ability to manage their clinical conditions.

WVMC implemented Merck's "Journey for Control" program to facilitate education and peer-to-peer discussion among patients with diabetes. This program provides tools and resources for diabetes educators to help support their patients' knowledge of diabetes and provide supplemental tools for enhanced patient self-management. A part of the Journey for Control program is the AADE7 Self-Care Behaviors framework, developed by the American Diabetes Association to encourage patients to develop seven behaviors it believes are key to better self-management of diabetes. WVMC incorporated this program into its on-site Wellness Days, which emphasized flu shots, education, and greater use of community resources. WVMC also explored the use of the Johns Hopkins Guided Care model. BMC held educational programs at the local senior center in Bend.

HHN, continuing as RBHC, contracted with Noridian Administrative Services, a CMS contractor that processes Part A claims in the States of Washington and Oregon, to obtain hospitalization and emergency room utilization data on intervention beneficiaries. Care managers received sentinel event data from Noridian about 1 month after the acute care event occurs. They used these reports to gain participation among beneficiaries who were eligible but not participating and as an outcome to assess whether the sentinel events were preventable. The care managers had expected that the receipt of the sentinel event data would allow them to engage nonparticipating beneficiaries. However, the care managers felt that this was not a successful strategy. Nor were the sentinel event data useful for identifying any pattern of hospitalizations, emergency room visits, or deaths that they felt could have been prevented. Retrospective review of their medical records by care managers at WVMC showed that most of the beneficiaries were admitted for reasons—including pain control, surgical hip repair, motor vehicle accidents, laceration repairs, and so forth— other than exacerbation of their primary clinical conditions.

E.3 Key Findings

In this section, we present key findings based upon the 38 months of the HBC operations with its original population and 26 months with its refresh population. Our findings are based on

the experience of approximately 3,600 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 the HBC program matured. One drawback to separate analyses of each group is the smaller samples available for statistical testing. Only 763 and 1,028 intervention beneficiaries were available for analysis in the original and refresh groups and comparable numbers in the corresponding 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 5% confidence levels. To better understand the statistical power underlying RTI's analyses, in subsequent chapters we present detailed statistics including (1) effect sizes for beneficiary experience with care measures; (2) confidence intervals for quality of care and acute care utilization measures; and (3) a detectable threshold for cost savings, or the rate of savings that would allow us to reject the null hypothesis of no savings.

Eight key findings on participation, intensity of engagement in the HBC program, 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 the HBC financially responsible for financial savings but does not hold the HBC financially responsible for quality of care improvements.

Key Finding #1: The HBC program was able to engage beneficiaries who were at higher risk of acute clinical deterioration as measured by the concurrent HCC score.

Of the HBC original intervention beneficiaries, 45% verbally consented to participate in the CMHCB demonstration at some point during the intervention period; 40% of the refresh population agreed to participate. For the HBC program, we find that beneficiaries with medium and high concurrent HCC scores were more likely to be participants. Beneficiaries with higher prospective HCC scores and baseline Charlson comorbidity scores were less likely to be participants. This suggests that the HBC program was less able to engage the historically sicker Medicare beneficiaries but more able to engage those at higher risk of acute clinical deterioration as measured by the concurrent HCC score.

Key Finding #2: Thirty-six percent of the intervention population agreed to use the Health Buddy® device.

A cornerstone of the HBC's program was the Health Buddy® device and interactions with care managers to address gaps in knowledge or self-management of their chronic diseases. Of the roughly 1,800 intervention beneficiaries, 668 beneficiaries (36%) agreed to participate in the program and used the device to complete at least one survey. Among the beneficiaries that did agree to participate in the HBC program, use of the Health Buddy® device was high (88%). Under an intent-to-treat model, active engagement of less than one-half of the total number of intervention beneficiaries requires that the HBC program has a large intervention effect on the beneficiaries with whom the HBC program staff members are actively engaging to achieve the desired outcomes.

Key Finding #3: The HBC program did not substantially affect beneficiary reported experience with care, level of physical activity, and self-reported physical health. Among the 19 outcomes covered by the survey, the HBC intervention resulted in a higher frequency of medication compliance for beneficiaries in the intervention group relative to the comparison group.

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 HBC CMHCB intervention: helpfulness of discussions with their health care team and quality of communication with their health care team. In addition, the survey instrument collected information about beneficiary self-care frequency and self-efficacy related to medications, diet, and exercise and Clinician and Group Adult Primary Care Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

The HBC demonstration program employs strategies to improve quality of care for high cost Medicare beneficiaries while reducing costs by empowering Medicare beneficiaries to better manage their care and mitigate acute flare-ups in the chronic conditions. 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, the HBC intervention resulted in a higher frequency of medication compliance for beneficiaries in the intervention group relative to the comparison group.

Key Finding #4: Rates of compliance with 3-of-4 quality-of-care process measures were high at baseline providing limited opportunity for improvement. The general trends during the demonstration were stable or decreasing rates of compliance in both the intervention and comparison groups.

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. Within the original and refresh intervention and comparison populations, we generally observe stable or negative trends in the rates. The original intervention group's rates tended to fall more than its comparison group's rates in 6-of-8 measurements; while the refresh intervention group's rates tended to fall less than its comparison group's rates in 3-of-4 measurements. The difference-in-differences (D-in-D) rates per 100 beneficiaries ranged from 1 to -7 per 100 beneficiaries for the original population and 8 to -7 per 100 beneficiaries for the refresh population. Of these differences, there is one that is statistically significant. The rate of receipt of the influenza vaccine among the refresh intervention beneficiaries declined by 1 percentage point while the rate of receipt among the refresh comparison beneficiaries increased 6 percentage points. Thus, the D-in-D change is -7 per 100 beneficiaries.

Key Finding #5: Rates of acute care utilization increased during the demonstration in the original and refresh intervention and comparison groups with one exception; all-cause hospitalizations declined within the refresh intervention group while the rate of all-cause hospitalizations increased within the comparison group. Although we observe no other statistically significant differential rates of growth in acute care utilization, we do observe a trend toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures. We do not observe differential use of the Medicare hospice benefit.

During the course of the HBC demonstration, we observed increasing rates of all-cause and ambulatory care sensitive conditions (ACSC) hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations with one exception. All-cause hospitalizations declined within the refresh intervention group while the rate of all-cause hospitalizations increased within the comparison group for a rate of -154 hospitalizations per 1,000 beneficiaries D-in-D rate (p=0.02). This represents a 26% lower rate than what would have been expected[2]. Although we observed no other statistically significant differential rates of growth in all-cause or ACSC hospitalizations or ER visits or 90-day readmissions, we observed a trend toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures with a number of the D-in-D rates appearing to be of clinical significance although not statistically significant. Further, we do observe wide confidence intervals for several of the readmission estimates due likely to small sample sizes. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rate of the Medicare hospice benefit or in mean and median number of days in hospice.

Key Finding #6: We observe a lower rate of mortality among intervention beneficiaries that used the Health Buddy® device.

We do observe a statistically significant lower rate of mortality in the original population's intervention group. Over the 38-month demonstration period for the original population, 35% of the original intervention group beneficiaries died while 40% of the comparison group beneficiaries died; a 5 percentage point lower rate of mortality in the intervention group (p=0.04). Over the 26-month demonstration period for the refresh population, 21% of the refresh intervention group beneficiaries died and 23% of the comparison group beneficiaries died; a 2 percentage point lower rate of mortality in the intervention group (p=0.33).

We estimated multivariate models of survival, whereby we controlled for potential imbalances in baseline beneficiary characteristics that may be related to mortality and not adequately accounted for in the development of a comparison group. When doing so, the

---

2 The percentage change in the D-in-D intervention rate is calculated by estimating the percent change in the comparison group's utilization between baseline and the demonstration period and applying the percent change to the intervention group's baseline rate. This produces an expected rate based upon the observed change in the comparison group. The percent change for the intervention rate is calculated using the expected rate as the baseline rate.

observed survival benefit for the intervention group within the original population was no longer present. However, when we introduced into our model a variable that captures the impact of intervention beneficiaries using the Health Buddy® device, we observed an incremental increase in survival benefit among both the original and refresh populations' intervention beneficiaries who used the Health Buddy® device. Because we did not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the survival benefit. However, given this important finding, additional study is warranted.

Key Finding #7: Medicare cost growth was slower in the intervention group in both the original and refresh populations, but neither trend was statistically significant.

No statistically significant savings were found for the intervention group in the original population. Costs rose $117 slower in the original intervention group (8.1% of the comparison group's costs), but savings needed to exceed 12.1% to be considered statistically significant. The HBC's trend in gross savings averaged -$73 in the refresh intervention group (6.0% of the comparison group's monthly costs), but savings needed to exceed 14.3% to be statistically significant. Insignificance may have been due to small numbers of intervention beneficiaries: 763 (original population); 1,028 (refresh population). A few material imbalances were found in cost, severity, and other patient characteristics between the original and refresh intervention and comparison groups in the base period. Still, controlling for imbalances had little effect on our overall final conclusion of no detectable statistically significant savings.

Actuarial Research Corporation (ARC), under separate contract to CMS, conducted an actuarial reconciliation of financial performance of the HBC program and also found gross savings for the intervention. ARC-determined savings differed from savings reported by RTI in three ways. First, ARC capped high-cost beneficiaries at the top 1% threshold. RTI did not cap outliers because we did not want to inadvertently bias results against the intervention if it was particularly successful in reducing costs of the very high-cost beneficiaries. Second, ARC adjusted for base period differences in intervention-comparison group costs without taking beneficiary eligibility during the demonstration period into account. RTI down-weighted base period costs for beneficiaries with shorter demonstration period exposure. Third, ARC made no independent assessment of the statistical reliability of their cost estimates. RTI conducted all analyses at the individual beneficiary level to be able to test the reliability of savings.

Simulation analyses showed that ARC's level of savings was sensitive to its outlier trimming and its estimates of base year average costs. Without trimming and using RTI's method for calculating base year costs, ARC's gross savings would have been $3.5 million in the original sample instead of $4.6 million and $1.7 million in the refresh sample instead of $2.0 million. Using ARC's gross savings based on RTI methods would have resulted in the HBC retaining $967,000 in fees instead of $2.8 million. That savings are still positive using a modified ARC approach and RTI's statistical approach suggest than the HBC's intervention is an approach worthy of continued study.

Key Finding #8: Beneficiaries in the refresh population using the Health Buddy® device exhibited a slower rate of cost growth.

Although the HBC program performance summarized in other findings is based on the entire intervention population, we were interested in whether beneficiaries using the Health Buddy® device had a slower rate of cost growth. Controlling for age, gender, minority status, and other beneficiary characteristics, those using the Health Buddy® device in the refresh population exhibited slower cost growth of over $200, significant at the 5% level of confidence. No difference was found in the original population. Because we could not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the cost savings and not the Health Buddy® device itself. Nevertheless, the lower rate of growth in Medicare costs and the lower observed rate of mortality supports continued study of the cost effectiveness of using monitoring devices in the home.

E.4 Conclusion

Based on extensive quantitative analysis of performance using statistical tests at standard 5% confidence levels, we did not detect improvement in key processes of care, beneficiary self-reported experience with care, self-management, and functional status, or use of the Medicare hospice benefit. The HBC program was successful in reducing the rate of all-cause hospitalizations within its refresh intervention group with a trend (not statistically significant) toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures. We also observed an incremental increase in survival benefit among the original and refresh populations' intervention beneficiaries who used the Health Buddy® device relative to the comparison group (and the rest of the intervention group). Although PBPM costs rose slower in the original and refresh intervention groups relative to the comparison groups, statistically significant savings were not achieved in the *overall* intervention groups. Nevertheless, we observed significantly lower cost increases among refresh intervention beneficiaries who used the Health Buddy® device.

What might explain the lack of *overall* program effectiveness? One factor may be relatively small sample sizes and lack of statistical power. Only 763 and 1,028 intervention beneficiaries were available for analysis in the original and refresh groups and comparable numbers in the corresponding comparison groups. In addition, wide variation in beneficiary costs over time made precise estimates of program success difficult with such small samples. CMS and Responding to the HBC's request, CMS selected a very costly, complex set of Medicare beneficiaries for their intervention and comparison groups. Mean per beneficiary per month base year claims costs (weighted by fraction of time eligible for the intervention) were approximately $1,000 in both groups, a figure considerably higher than in the general Medicare population. Further, we observed extreme regression-to-the-mean (RtoM) behavior among the HBC's selected beneficiaries. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by over $1,000. Over the same time period, beneficiaries with monthly costs over $3,000 saw their average costs decline by $1,500-$2,500. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings.

A second factor may be the HBC's beneficiary recruitment strategy. Given the HBC program's monthly management fee (roughly $120 per month) and the population-based design of this demonstration, engagement of less than 50% of the intervention population required the HBC program to have been extremely successful with the participating beneficiaries.

And, a third factor 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 HBC program's fee. A cornerstone of the HBC's program was health coaching interactions with care manager nurses in response to alerts generated by the Health Buddy® device. Nearly all participating beneficiaries using the Health Buddy® device received at least one call from a care manager and nearly 60% received more than 20 calls. This is a relatively high contact rate compared to other care management programs that we have evaluated. However, the Health Buddy® nurse care managers often were not in direct proximity to their beneficiaries' primary care physicians, thereby potentially affecting their interactions with the beneficiaries' primary providers, changing medical care plans, or mitigating deterioration in health status. The care manager served primarily as an adjunct to the patients' primary physicians. Interviewed physicians felt that care management would be more effective and efficient if care managers were colocated with primary care physicians. Further, not all intervention beneficiaries had primary care physicians in the two study sites, therefore the care managers had to interact with community-based providers with whom they had little or no prior relationship. During our site visits, the care managers cited several challenges working with these physicians, in particular, because of communication barriers. Lastly, by complementing, not substituting, for the primary care physician, the nurse care managers were not directly determining whether a patient was admitted to a hospital or what service intensity the beneficiaries would receive during the demonstration period.

Yet, we do observe an incremental increase in survival benefit and lower cost increases among intervention beneficiaries who used the Health Buddy® device. As noted before, because we could not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the survival benefit and cost savings and not the Health Buddy® device itself. These two substantive findings require further evaluation by analysis of the HBC Phase II demonstration experience. It will be important to explore with the HBC what beneficiary characteristics they believe lead them to agree to use the Health Buddy® device. With this information, we may be able to develop an alternative comparison group that more closely aligns with the subset of beneficiaries that use the Health Buddy® device.

This page intentionally left blank.

CHAPTER 1
INTRODUCTION TO THE MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION AND THE HEALTH BUDDY® CONSORTIUM'S (HBC) PROGRAM

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 Health Buddy® Consortium's (HBC) 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 Healthcare (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 and Massachusetts General Physicians Organization (MGH)

4. Montefiore Medical Center (MMC)

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 17 of the intervention period. The HBC's survey was conducted between June 11, 2007 and October 10, 2007. 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 38 months of the HBC CMHCB program operations with its original population and 26 months with its refresh population. We start by reporting on the degree to which the HBC was able to engage its intervention populations. We measure degree of engagement in two ways: (1) participation rates and characteristics of participants; and (2) number and nature of contacts between the HBC and participating beneficiaries from encounter data provided to RTI from the HBC. We then report findings related to the effectiveness of the HBC program to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings.

1.2 The HBC's CMHCB Demonstration Program Design Features

1.2.1 The HBC Organizational Characteristics

The Health Buddy® Consortium (HBC) was one of six organizations providing care management support as part of the Medicare CMHCB Demonstration coordinated by CMS. The regions covered by the HBC's CMHCB program during Phase I included central Oregon and central Washington. The HBC launched its program on February 1, 2006.

The HBC CMHCB program is a systems-based approach to chronic care based on the Chronic Care Model and supported by the Health Buddy®, a health monitoring device that collects qualitative and quantitative information from patients on a daily basis. Care managers monitor patient responses to surveys conducted via the device and follow up with patients to help them address clinical issues and initiate interventions as needed to maintain their health. The device engages and educates patients so they may better understand their health conditions and proactively manage their disease by modifying high-risk behaviors. Routine monitoring of patient health status and symptoms alerts providers to health issues that require early intervention in an effort to avert serious complications requiring hospitalization. The overarching goal of the project is to demonstrate that multi-specialty medical groups, applying a consistent model of care management augmented by an integrated technology solution, are uniquely positioned to improve the lives and reduce the costs associated with high-cost beneficiaries insured by traditional Medicare FFS.

The HBC CMHCB demonstration program was run by a consortium of four organizations collaborating to deliver care management services to high-cost Medicare beneficiaries with diabetes, heart failure (HF), and/or chronic obstructive pulmonary disease (COPD). The consortium was coordinated by a medical technology company, Health Hero Network (referred to throughout this report as HHN) until HHN was acquired by and became known as Robert Bosch Healthcare (RBHC) in December 2007. HHN received support from the AMGA to implement the Health Buddy Program's consistent chronic care management process at two multi-specialty practices, Bend Memorial Clinic in Central Oregon and Wenatchee Valley Medical Center in North Central Washington.

Health Hero Network

Founded in 1998, HHN was a health technology company that developed and marketed computerized decision support tools, health monitoring technologies, and clinical information

databases to support the delivery of care management services.[3] The Health Buddy® System, one of HHN's primary technologies, is a device that allows patients to communicate with health care providers on a daily basis by answering a series of questions about symptoms, behaviors, and knowledge related to their health conditions. HHN enlisted the services of Cobbwebb Associates to assist in non-clinical aspects of the project, such as contributing to the proposal and program design development. HHN also contracted with Milliman, a consulting firm with expertise in Medicare claims analysis, to review claims data provided by CMS.

Robert Bosch Healthcare

Robert Bosch Healthcare Inc. is a fully-owned subsidiary of the Bosch Group. The product spectrum ranges from patient terminals through fast and secure transmission of health data (vital parameters), to evaluation software for healthcare professionals. The Bosch Group is a global supplier of technology and services. It comprises Robert Bosch GmbH and its roughly 300 subsidiary and regional companies in over 60 countries.

American Medical Group Association

AMGA is a professional organization that advocates for the multi-specialty group practice model of health care delivery and for the patients served by medical groups through innovation and information sharing, benchmarking, and efforts to continuously improve patient care. At the outset of the program, AMGA managed relationships with the two partner medical groups, supported HHN in its ongoing relationship with CMS, convened weekly conference calls and quarterly on-site meetings among consortium members to discuss ongoing operations issues, and conducts annual audits at each of the participating sites to identify opportunities to improve program quality and performance. Over the course of the demonstration, AMGA's role evolved more into a consultative role as quality of care did not become a performance metric within the demonstration contract.

Physician Practices: Bend Memorial Clinic and Wenatchee Valley Medical Center

Located in Bend, Oregon, Bend Memorial Clinic (BMC) is the largest multi-specialty group practice in Central Oregon. Prior to this project, BMC had been interested in implementing a care management program, but had not participated in a formal disease management program and had limited experience with performance monitoring. Therefore, BMC viewed the CMHCB program as an opportunity to benefit from support provided by the Health Buddy® Consortium to gain experience implementing care management systems and protocols, establish associated workflows, and offer the Health Buddy® to its most ill patients, a tool that BMC believed would help these individuals.

The Wenatchee Valley Medical Center (WVMC) is the second largest multi-specialty group practice in the Pacific Northwest, employing over 170 physicians who staff over 50 different clinical departments. WVMC was interested in the CMHCB program as a way to decrease multiple hospitalizations among patients with chronic illnesses, expand its case

[3] Health Hero Network was originally founded as Raya Systems in 1989.

management programs, offer the Health Buddy® technology to its patients, and replicate the care model in its various clinics.

Both of these physician practices were responsible for enrolling patients in the Health Buddy® program, reviewing information collected from patients using the Health Buddy® device on a daily basis, and contacting patients by telephone and coordinating with physicians, as needed, to assist patients with clinical issues identified by the monitoring technology.

1.2.2 Market Characteristics

Central Oregon and central Washington are primarily rural areas, where the population is widely dispersed over a large geographic area. Approximately 40% of the population qualifies for Medicare, and the elderly populations in these areas are growing. Healthy retired persons are attracted to this region to enjoy the extensive recreational opportunities. A significant proportion of the elderly are "snowbirds"—spending 6 months each year in the Northwest and 6 months in warmer states, such as Arizona, California, Florida, or Hawaii. There is also a large population of Medicare beneficiaries with chronic illnesses, who are attracted to the area for its existing health care infrastructure and/or the presence of family members to provide care. The populations of both states are approximately 80% White and 8% Hispanic.

BMC has a close relationship to a nearby hospital, St. Charles Medical Center, and many of BMC's physicians have admitting privileges at this facility. WVMC operates its own 21-bed hospital and has a strong relationship with Central Washington Hospital. WVMC reported that its service area and surrounding areas are characterized by significant access to care problems due to a shortage of physicians and poor public transportation access. In addition, some specialists have begun turning away Medicaid patients, so WVMC often serves as a provider of last resort seeing patients who live more than 3 hours from its facility. Overall, the region has a low penetration of managed care.

1.2.3 The HBC Intervention and Comparison Populations

Intervention population. HHN and the AMGA worked with their CMS project officer and analysts from ARC to develop a methodology for selecting the starting population for the HBC program. HHN and AMGA chose to focus on patients with diabetes, HF, and COPD because initial analyses indicated that 60% of those with high health care utilization and costs had at least one of these three conditions and would therefore be captured in the demonstration population. Inclusion criteria for eligibility in the HBC CMHCB demonstration program included the following:

- Medicare fee-for-service beneficiaries with a primary residence in designated counties of Oregon and Washington, with high costs in 2004 (i.e., Medicare costs greater than or equal to $6,000 in 2004) **or** high disease severity as indicated by Hierarchical Condition Categories (HCC) risk scores greater than or equal to 1.7,

- A plurality[4] of visits or two or more visits to the HBC medical practices as evidenced by one or more claims in 2004 associated with a provider at either Bend Memorial Clinic (BMC) or The Wenatchee Valley Medical Center (WVMC), and

- One or more claims associated with diagnosis codes for at least one of the following conditions: HF, diabetes, or COPD.

The population was further restricted using the following exclusion criteria: receiving hospice care or care for end-stage renal disease (ESRD) benefit, enrolled in a Medicare Advantage plan or have Medicare as a secondary payer or lack Medicare Part A or Part B coverage as of January 2, 2006, and at least one of the exclusionary diagnoses designated by HHN and AMGA (e.g., dementia or spinal cord disease). The remaining beneficiaries were randomly assigned to the intervention group. The randomization was done on a state-by-state basis by alphabetizing eligible beneficiaries by their last name. BMC was assigned 661 patients and WVMC had 965 patients for a total of 1,626 beneficiaries.

The CMHCB demonstration program was designed using an ITT model, which means that the CMOs are held accountable for outcomes across the full intervention population, not just those who agree to participate. This model provides CMOs with flexibility in their operations and strong incentives to keep evolving toward outreach and intervention strategies that are most effective in improving population outcomes. Once individuals were assigned to either the intervention or comparison group, they remained in their assigned group for all days in which they were eligible. Eligibility for the HBC program and hence membership in either the intervention or comparison group was lost for any period(s) during which the beneficiary:

- enrolled in an MA plan,

- lost eligibility for Medicare Part A or B,

- got a new primary payer (i.e., Medicare becomes the secondary payer),

- moved out of the HBC program service area,

- developed ESRD,

- elected the hospice benefit, or

- died.

Comparison population. Following the development of the intervention group criteria, HBC 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 program. The methodology was an approach that identified Physician Group Practices (PGPs) identified by their Tax Identifications Numbers (TINs) that would be appropriate for comparison purposes. Claims data were selected for all Medicare beneficiaries residing in each of the 18

[4] Plurality refers to receiving more care from one provider (e.g., BMC) than any other provider.

nominated counties and were processed through the ARC programs to determine initial eligibility based upon the general exclusion, cost and HCC score criteria. The total number of unique beneficiaries associated with each TIN was calculated. For the top 10 TINs within each county, defined by number of Medicare beneficiaries, RTI calculated the proportion of total physician payments by 24 major types of services (Berenson-Eggers Type of Service categories, or BETOS). Examples of BETOS categories are office visits, hospital visits, anesthesia, specialty consults, durable medical equipment, radiology, diagnostic testing, etc. RTI also calculated this distribution for the intervention population separately for BMC and WVMC. Office visits and diagnostic laboratory testing were the two most common types of services provided by the two intervention clinics. Thus, RTI selected a threshold of 20% of payments from office visits and/or diagnostic laboratory testing, and selected all TINs in each of the counties that met this criterion. This step effectively removed TINs representing single specialty practices, i.e., anesthesiology practices, hospital-based practices providing emergency room coverage, etc. RTI then applied the exclusion, inclusion, and loyalty criteria to the designated TINs.

Beneficiaries in these practices and in the Oregon and Washington intervention clinics were then stratified using a 9-cell diagnostic (HF only; HF + 1-2 other diseases; diabetes or COPD only) and cost (<$5,520; $5,520-$15,029; >$15,029) matrix. In Washington, this method initially produced a pool of only 962 eligible comparison beneficiaries with insufficient numbers for matching purposes in five of the stratification cells. RTI therefore added three more counties (Stevens, Ferry, and Lincoln) and additional TINs to increase the size of the pool. In Oregon, this method identified 17 TINs and 1,925 beneficiaries in 9 comparison counties. In Washington, the 18 TINs yielded 1,569 beneficiaries in 12 counties.

The final step in the process was to draw the sample of comparison group beneficiaries from the pool of those identified as eligible by the TIN volume approach. For each of the nine cells in the diagnosis by cost matrix, this was done by randomly choosing beneficiaries from the comparison pool to match the number in the corresponding cell in the intervention group. This produced a final group of 660 comparison beneficiaries in Oregon and 964 in Washington, the same as the size and distribution of the intervention groups in each state.

As a final check, RTI compared the claims-based measures for the intervention and comparison samples. The values were very similar in the two groups. None of the differences exceeded 0.15 standard deviations or appeared to be of clinical significance. Total Medicare payments did not differ significantly between the two groups. RTI's ultimate objective was to select a group of comparison beneficiaries whose baseline costs were equivalent to those in the intervention group. The combined per member per month (PMPM) cost during the baseline period for the Washington and Oregon intervention groups was $1,289 PMPM. For the combined comparison groups, the cost was $1,280 PMPM. Thus, RTI achieved both cost equivalence and of operationalizing a process to select a comparison group loyal to a PGP that closely replicated the process used for the intervention group.

After the payment comparisons had been completed, two additional modifications were made to the groups. The first modification was to eliminate comparison group beneficiaries who were enrolled in other CMS demonstrations, such as the BIPA demonstration or the Medicare Health Support pilot. RTI submitted a finder file containing all 1,624 comparison group beneficiary identification numbers to Mathematica Policy Research to be run against their list of

beneficiaries participating in any of these demonstration programs. One beneficiary was identified and removed from the comparison group, leaving the comparison group at 1,623.

The second modification involved a subsequent eligibility "true-up" of beneficiaries to a common date. In August 2006, it was discovered that eligibility for the intervention and comparison groups for the HBC program was not determined at the same time, eligibility for the intervention group was determined on January 2nd, 2006 while eligibility for the comparison group in Washington was determined on March 11th, 2006 and eligibility for the comparison group in Oregon was determined on March 18th, 2006. As a result, there were a number of beneficiaries in the intervention group who became ineligible (e.g., due to death, moving out of service area, etc.) between the time the intervention group was established and finalization of the comparison groups. CMS was concerned that this would bias the comparison of the intervention and comparison groups and decided to conduct an eligibility true-up process. To correct for the difference in eligibility between the two groups, a historical eligibility pull was completed on August 10th, 2006 for both the intervention and comparison groups. Using this view and a new eligibility date of March 18th, 2006, 29 comparison beneficiaries were removed as the result of the true-up. Following these modifications, the final sample size was 1,594 beneficiaries in the comparison group.

Carve-out and refresh population. After program implementation, the HBC requested from CMS reconsideration of its intervention population. The HBC requested removal of selected beneficiaries (carve-out) from its starting population and the addition of new beneficiaries using alternative inclusion/exclusion criteria at the time of their planned refreshment of their intervention and comparison populations. The HBC requested that the analyses to identify beneficiaries for retention be conducted using claims dated August 1, 2005 through July 31, 2006. HBC submitted diagnostic criteria that would retain beneficiaries with diabetes, heart failure, COPD, and hypertension/coronary artery disease (CAD), a set of diagnoses that would lead to exclusion (malignant neoplasms, dementias, substance abuse disorders, mental health disorders, and cerebral or neurological disorders), and a complex algorithm that evaluated principal and secondary diagnoses on different claim types and identified thresholds of frequency. In addition, a somewhat more restrictive set of evaluation and management codes were used. Of the starting comparison population of 1,595, 788 beneficiaries (49%) were identified for the carve-out. Final eligibility was determined as of February 3, 2007 with 1,056 beneficiaries in the final refresh intervention population.

In selecting the refresh comparison population, RTI employed the same loyalty definition of "plurality or two or more visits" for the selection of the refresh intervention group with the revised claims-based inclusion and exclusion criteria. The total number of comparison refresh beneficiaries drawn was 1,056, the same size as the intervention refresh group with 322 comparison beneficiaries from Oregon and 734 comparison beneficiaries from Washington.

1.2.4 Overview of the HBC CMHCB Demonstration Program

HBC launched its program February 1, 2006. RTI conducted two site visits to the HBC program offices in Bend, Oregon and Wenatchee, Washington, the two locations where the HBC CMHCB program was implemented. The first site visit was conducted 6 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 HBC program and reporting early experiences with program implementation and engagement of eligible beneficiaries, providers, and CMS. The second site visit, 26 months into the demonstration, focused on HBC staff's impressions and interpretation of its 26-month experience in working on the demonstration program. 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 HBC's CMHCB demonstration program and its activities in this report reflects HBC'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 HBC program participants and physicians, as well as the clinical protocols/analytic tools to support the HBC nurse care managers and other health professionals who delivered these services. Second, we discuss program changes and enhancement activities that occurred as the program evolved.

Participant Support Services. The core of the intervention was a care management program augmented by the use of a Health Buddy® device. Once a beneficiary consented to participate in the HBC CMHCB program, he or she was asked if they would be willing to receive a Health Buddy® device that would allow for routine communication with program staff through daily sessions that span 5 to 10 minutes. During each session, patients were expected to answer a series of questions related to symptoms, vital signs, knowledge, and health behaviors and receive educational information about their health conditions. Each device was set up with one of 13 disease-specific programs that most closely address the needs of each participant. For example, patients with diabetes participated in dialogs that focused primarily on issues related to diabetes. Four of the programs addressed comorbid conditions, such as COPD and diabetes, or COPD, HF, and diabetes. Nurse care managers at each of the program sites monitored patient responses to Health Buddy® questions on a daily basis using a web-based computer application, the Health Buddy® desktop. This system helped care managers determine the urgency of follow-up required for each patient, triage patients accordingly, and coordinate the provision of appropriate medical, psychological, or social services.

For individuals who did not have a chronic condition, there was a 14th program, a Senior Wellness program that addressed issues related to general health and safety, as well as psychosocial issues, such as depression. Patients who are unable or unwilling to use the Health Buddy® device had the opportunity to participate in the Health Buddy® program through routinely scheduled telephone calls with nurse care managers that occurred weekly, bi-weekly, or monthly depending on the patient's health status. This is referred to as the alternate program.

The Health Buddy® disease management program content is drawn from evidence-based practice guidelines, and each program is designed to collect standard outcome measures including utilization, patient satisfaction, quality of life, and compliance with treatment regimens. Advisory boards at each of the participating medical practices reviewed the Health Buddy® disease management programs and developed associated care protocols and care plans to guide care managers' responses to alerts associated with each disease-specific program.

Physician Support Services. The Health Buddy® Program provided physicians with information about patient symptoms, vital signs, and behaviors during the time period between office visits. As a result, providers had the opportunity to intervene with patients when they have early symptoms of health problems, potentially avoiding hospitalizations or emergency room visits. Further, physicians could review trends in patients' Health Buddy® responses prior to scheduled office visits, which could help them to identify health issues that required attention.

During the first site visit, physicians at both sites reported that they were initially very enthusiastic about the Health Buddy® program, because it offered a promising way to effectively support patients with chronic disease. The Health Buddy® technology coupled with telephonic care management support was viewed as an effective way to maintain and improve patient health and identify symptoms of complications early, so that timely medical intervention could be used to prevent serious problems requiring hospitalization. Once the physicians received the list of patients who were eligible for the Health Buddy® program, they reported that they became frustrated with the project because they felt that many of the patients selected would not benefit from participating. Further, physicians reported disappointment that many of the patients they believed could be helped by the program were not eligible to participate in the program because they had not been identified through the claims based algorithm developed by HHN.

Using information gleaned from its early experience with the program, the HBC made a series of changes and enhancements to its operations and as reported to us at our second site visit.

Refresh Population. The HBC negotiated with CMS to revise the beneficiary selection criteria with the goal to select beneficiaries whom the clinics thought were more clinically appropriate for the demonstration. The HBC applied more stringent diagnostic inclusion and exclusion criteria specifying that only inpatient, outpatient hospital, and physician claims should be analyzed when determining the presence of diabetes, heart failure, or COPD. The HBC also developed more stringent utilization thresholds than those used in the initial population selection, and included coronary artery disease and hypertension when determining utilization. The HBC also created a fourth diagnostic category, comorbid (consisting of any of the three diseases), to identify any target beneficiaries who may have been missed using a single disease diagnosis and the more stringent utilization thresholds. In selecting their refresh population, the HBC employed the same loyalty definition of "plurality or two or more visits" for the selection of their refresh intervention group. However, they used a somewhat more restrictive set of evaluation and management codes.

Engagement of the refresh population. BMC conducted its outreach campaign in stages, starting with those with trimorbid conditions, then those with bimorbid conditions, then groups of beneficiaries with a single diagnosis. This strategy enabled BMC to first reach beneficiaries whom it believed would be sicker and would thus benefit most from the Health Buddy® program. WVMC stratified its refresh list geographically, rather than by morbidity, because its care managers were located geographically. Care managers at WVMC, with their case management responsibilities, did not have sufficient time to make multiple rounds of calls and send the mailings needed to reach the people. RBHC added its own staff and then contracted a third-party vendor, Health Contact Partners, to assist with making calls.

Enhancement of the Health Buddy® program, tailoring of the Health Buddy® programs to meet the need of participants with multiple comorbidities, and maintenance of participant interest. Over the course of the demonstration, there were no major changes in the overall programmatic approach; however, there were enhancements focused on tailoring the Health Buddy® programs and dialogues to meet the needs of participants with multiple comorbidities and to maintain participant interest. The HBC program reported that they had made adjustments to effectively address the multiple comorbidities characteristic of the high-cost Medicare population. The content of most of the disease management programs deployed via the Health Buddy® device addressed one particular health condition. However, Medicare beneficiaries with high health care utilization typically have multiple comorbid conditions. Feedback obtained during the site visit indicated that care managers were providing telephonic support for conditions that were not addressed by the 14 programs then available to the HBC CMHCB population. As a result, the HBC program was in the process of introducing additional disease management programs in an effort to address the needs of as many members of the intervention population as possible. The care managers at each of the clinics also expanded their focus to pay greater attention to social issues that affected their participants' ability to manage their clinical conditions.

Implementation of Merck's "Journey for Control" program, the "Guided Care" model, and Wellness/Education Days. WVMC implemented Merck's "Journey for Control" program to facilitate education and peer-to-peer discussion among patients with diabetes. This program provides tools and resources for diabetes educators to help support their patients' knowledge of diabetes and provide supplemental tools for enhanced patient self-management. The program offers Conversation Map® training that was developed in collaboration with the American Diabetes Association. The training helps educators to learn methods of engaging groups of patients in a nonthreatening environment to learn about their diabetes and self-management needs and to develop action plans through practical knowledge gained from the educator as well other members of the group. A part of the Journey for Control program is the AADE7 Self-Care Behaviors framework, developed by the American Diabetes Association (ADA) to encourage patients to develop seven behaviors the ADA believes are key to better self-management of diabetes. WVMC incorporated this program into its on-site Wellness Days, which emphasized flu shots, education, and greater use of community resources. WVMC also explored the use of the Johns Hopkins Guided Care model. BMC held educational programs at the local senior center in Bend.

Receipt of sentinel event data. HHN, continuing as RBHC, contracted with Noridian Administrative Services, a CMS contractor that processes Part A claims for the States of Washington and Oregon, to obtain hospitalization and emergency room utilization data on intervention beneficiaries. Care managers received sentinel event data from Noridian about 1 month after the acute care event occurs. They used these reports to solicit participation among beneficiaries who were eligible but not participating and to assess as an outcome whether the sentinel events were preventable. The care managers had expected that the receipt of the sentinel event data would allow them to engage nonparticipating beneficiaries. However, this was not a successful strategy. Nor were the sentinel event data useful for identifying any pattern of hospitalizations, emergency room visits, or deaths that they felt could have been prevented. Retrospective review of their medical records showed that most of the beneficiaries were admitted for reasons other than exacerbation of their primary clinical conditions.

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

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 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.

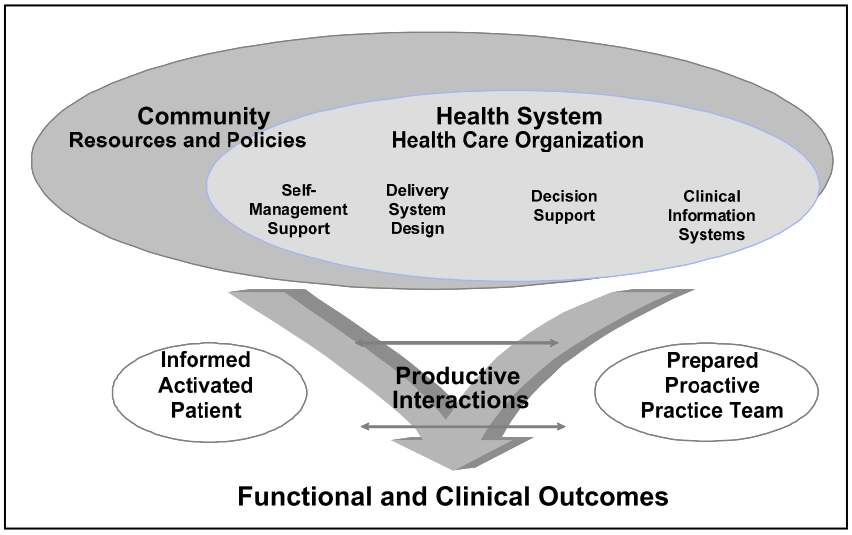
2.1.2 Emerging Approaches to Chronic Care

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).[5] 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

[5] 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.

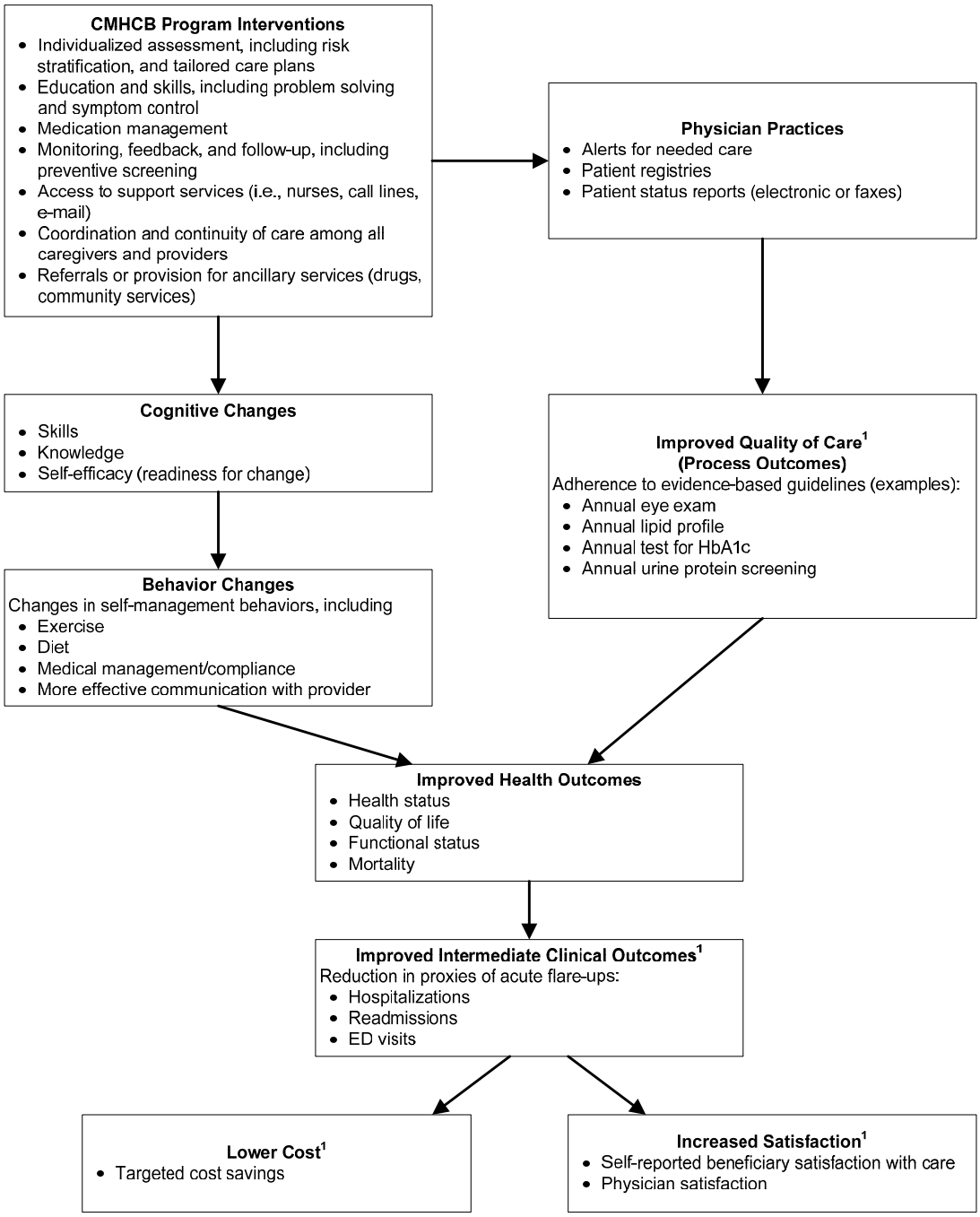
- *Monitoring, feedback, and follow-up*. Activities in this domain include ongoing biomonitoring of beneficiaries by placing scales or other equipment in their homes or by having the beneficiaries self-report their weights, blood sugars, or other measures. When data on preventive services, screenings, 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 HBC program was able to engage its intervention population and achieve four outcomes. ***Table 2-1*** presents a summary of research questions and data sources, organized by three evaluation domains: Reach, Implementation, and Effectiveness. The HBC program implementation experience was is reported in ***Chapter 1***.

Table 2-1
Evaluation research questions and data sources

Research questions	Site visits	CMO data	Claims	Survey
IMPLEMENTATION: To what extent was the HBC able to implement its program?				
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 HBC program?	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 HBC program engage physicians and physician practices in their programs?	Yes	No	No	No
REACH: How well did the HBC program 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 HBC program 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 HBC program?	No	Yes	Yes	No
4. To what extent were the intended audiences exposed to the HBC programmatic 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 HBC demonstration intervention versus a low level of intervention?	No	Yes	Yes	No
EFFECTIVENESS: To what degree was the HBC program 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 HBC program 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 HBC program 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 HBC program result in greater engagement in health behaviors?	No	No	No	Yes
3. Did the HBC program 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 HBC demonstration program improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?	No	No	Yes	No
2. Did the HBC program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?	No	No	Yes	No
3. Did the HBC program 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 38 or 26 months of the demonstration for the intervention and the comparison groups?	No	No	Yes	No
2. What were the levels and trends in PBPM costs for intervention group participants and nonparticipants? Did nonparticipation, alone, materially reduce the intervention's overall cost savings?	No	No	Yes	No
3. How variable were PBPM costs in this high cost, high risk, population? What was the minimal detectable savings rate given the variability in beneficiary PBPM costs?	No	No	Yes	No
4. How did Medicare savings for the 38- or 26-month period compare with the fees that were paid out? How close was the HBC program 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 differences 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: CMO = care management organization; HBC = Health Buddy® Consortium; CMS = Centers for Medicare & Medicaid Services; CMHCB = Care Management for High Cost Beneficiaries; ER = emergency room; PBPM = per beneficiary per month.

2.1.4 General Analytic Approach

The CMHCB initiative is what is commonly called a "community intervention trial" (Piantadosi, 1997). It is a "community" in the sense of being population based for a 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-intervention assignment" 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

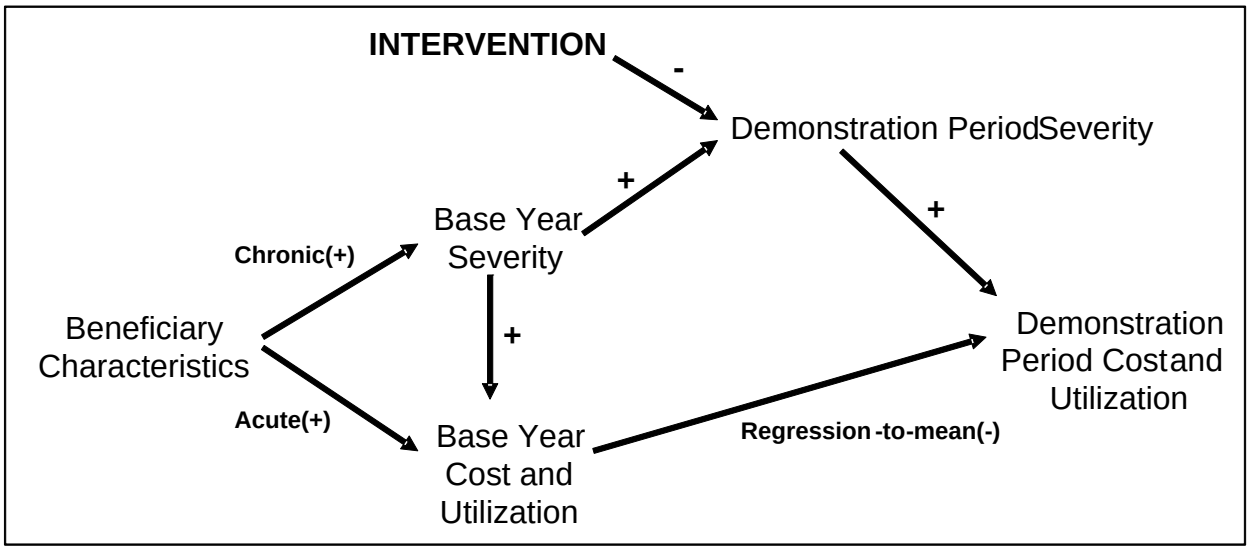
ε = a regression error term.

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 HBC CMHCB demonstration program. As noted in ***Chapter 1***, we also conducted a survey of the HBC CMHCB demonstration beneficiaries to assess their satisfaction with the CMHCB program and semi-structured interviews with a small number of physicians to assess their awareness of and satisfaction with the CMHCB program. The data used to make those assessments are described in ***Chapter 3***.

2.2.1 Data

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

- *Participant status files*. We received participant status files from ARC. The participant status information originates from the HBC program and was submitted to ARC. This file was updated quarterly and logged status changes among the intervention groups by the HBC program. 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 of the HBC program beneficiaries was assigned—intervention or comparison—for both the original and refresh populations.

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

  - ARC provided RTI with an EDB file for the HBC program comprised of all original and refresh beneficiaries, excluding the carve-out beneficiaries from the original population. RTI used this file to determine daily eligibility based on the HBC program 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 HBC program launch date. We used the files to identify days of eligibility during the 12-month baseline period and the intervention periods of the demonstration and to select claims data during periods of eligibility in both the baseline and intervention periods. *Only beneficiaries who had at least 1 day of eligibility in the baseline and the demonstration periods are included in our evaluation.*

  - RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (January 2, 2006) for the HBC's original population.

  - RTI conducted an EDB extraction to obtain demographic characteristics at the time of assignment (January 2, 2007) for the HBC's 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 HBC program 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*. The HBC uses a health monitoring device that collects qualitative and quantitative information from patients on a daily basis. The intervention data files provided to us only collect information from patients that use the device. Quarterly, the HBC program sent RTI beneficiary-level intervention files that contained summary counts of intervention activities, such as the number of surveys completed, counts of the number of inbound calls to a care manager from a patient and outbound calls to a patient from a care manager, as well as counts of calls between care managers and doctors regarding the patient. Information about high risk responses was also collected. In June 2010, the HBC program provided new quarterly files that contained updated and more complete information for the entire demonstration period. 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 HBC program

Ineligibility reasons	Description
Death	Ineligible beginning on day following date of death.
Hospice	Ineligible on hospice coverage start date. Eligible on day following hospice coverage end date.
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 HBC program. 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: HBC = Health Buddy® Consortium; ESRD = end-stage renal disease; MA = Medicare Advantage; GHO = Group Health Organization.

Table 2-3 contains the HBC program's evaluation start and end dates, both baseline and intervention period, for the original and refresh populations.

Table 2-3
Analysis periods used in the HBC CMHCB demonstration 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				
2/1/06	3/31/09	38	2/1/05	1/31/06
Refresh Population				
2/1/07	3/31/09	26	2/1/06	1/31/07

NOTES: CMHCB = Care Management for High Cost Beneficiaries; HBC = Health Buddy® Consortium.

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, January 2, 2006 for the original population and the refresh assignment date (January 2, 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 HBC program was active in the demonstration. The numerator is the number of days the beneficiary is eligible during that time period. The HBC program participated in the demonstration for the full 38 months, so the number of days in the denominator for each original population beneficiary in the HBC program is 1,155 (the HBC end date minus the HBC 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 1,155, or 0.364.

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 assignment.

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. 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, COPD 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:

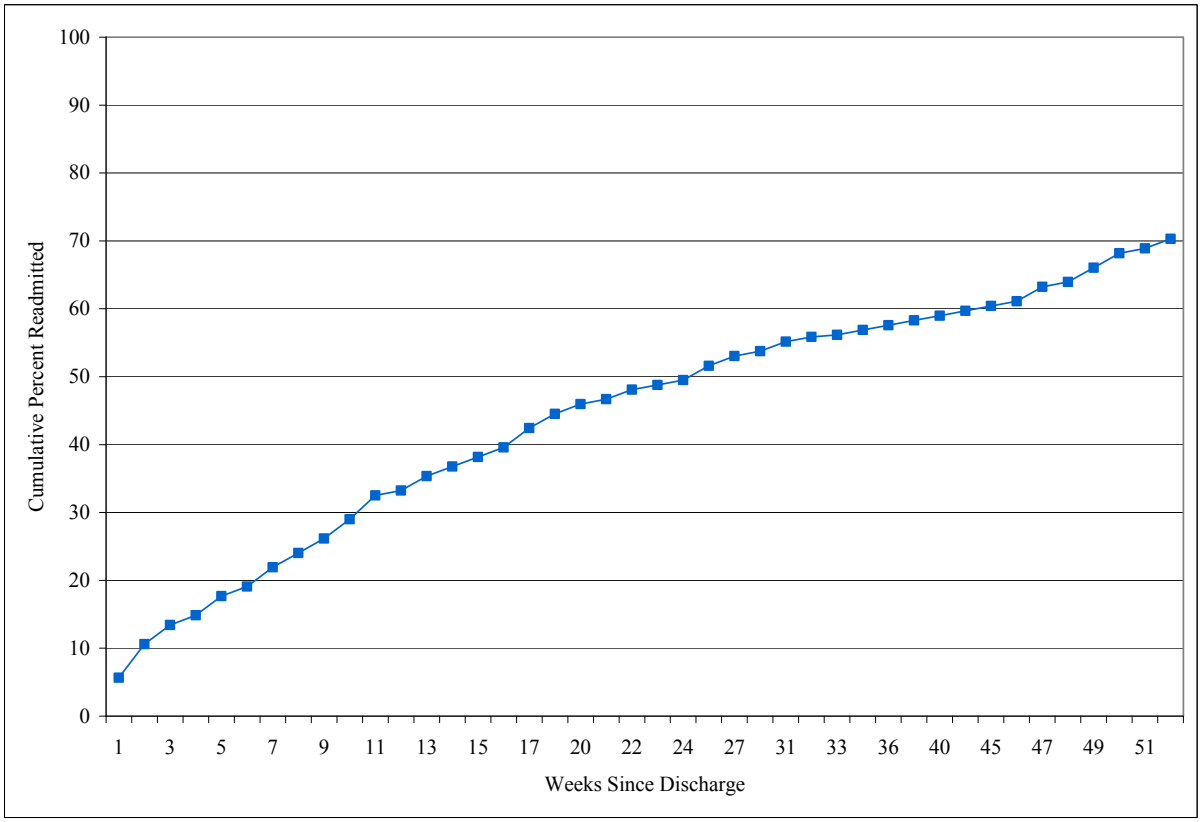
- the number of acute hospitalizations,
- 90-day readmissions, and
- emergency room visits, including observation bed stays.

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 HBC's CMHCB program 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 one-third of subsequent admissions in our analyses[6].

Figure 2-4
Percent with readmission for any diagnosis: The HBC'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 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

6 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.

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 24 through 35 and included readmissions through months 21 and 38, respectively. The intervention period for the refresh population examined admissions during months 12 through 23 and readmissions through month 26. 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) 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 38-month original intervention period and the 26-month refresh intervention period.

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

   First, we selected a measure that is broadly applicable to the Medicare fee-for-service population, influenza vaccination. Second, we selected several measures that are specific to beneficiaries with diabetes and heart failure as these populations are prevalent in the HBC demonstration population. The study populations were subset to the appropriate clinical cohorts to construct these measures.

   The selected measures and relevant disease population are as follows:

   - Rate of influenza shots for adults > 50 years – 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 methodology used to create these measures can be found in ***Supplement 2A***. CMS requested that we use existing, widely adopted specifications for evidence-based measures of care. Based on that request, RTI selected the National Quality Forum (NQF)–endorsed National Voluntary Consensus Standards for Physician-Focused Ambulatory Care. While the NQF-endorsed specifications restrict the diabetes quality-of-care measures to beneficiaries ages 18 to 75, we did not use this age restriction because no such restriction is used by the HBC program. 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 HBC program in case the services occurred after exposure to the CMHCB demonstration intervention and during the intervention period. One could envision that the HBC program 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 the HBC program, we constructed counts of the number of telephonic contacts with the participants (both inbound and outbound) and between caregivers—as well as total contacts (both), and number of surveys completed.

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[7].

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 methodology 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).

3.1.1 Survey Instrument Design

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

---

7 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 HBC demonstration program team. We do so for the primary reason that we are asking the comparison population the same question and we desire to isolate the effect of the HBC intervention on the beneficiaries' assessment of satisfaction that their full health care team is helping them to cope with their chronic conditions.

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

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

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

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

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

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

converted into composite scores ranging from 0 (never to all items) to a maximum of 100 (always to all items).

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

3.1.1.2 Self-Management Measures

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

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

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

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

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

3.1.1.3 Physical and Mental Health Function

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

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

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

Activities of daily living—A related measure of beneficiary functioning is the ability to perform basic activities of daily living (ADLs). The questionnaire collected information about six standard activities—bathing, dressing, eating, getting in and out of chairs, walking, and using the toilet. Respondents were first asked if they had any difficulty performing each activity. Possible responses were that they were unable to perform, had difficulty, or did not have difficulty doing the activity. They were then asked, with responses of yes or no, if they needed help from another person to perform the activity. An ADL difficulty score was created by counting the number of activities that the beneficiary had difficulty with or was unable to do. The ADL help score was the number of activities for which the beneficiary needed help. Each score ranges from 0 to 6.

3.1.1.4 Background Characteristics

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

3.1.2 Analytic Methods

We conducted a series of statistical analyses to explore intervention-comparison differences and CMHCB 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 Health Buddy® Consortium'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 HBC 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 a Medicare Advantage plan, etc.) were ineligible for the survey frame. To maximize the number of eligible respondents in the frame, we performed a Medicare EDB run prior to sampling to identify decedents and other beneficiaries who had recently become ineligible.

3.1.2.3 Data Collection Procedures

We surveyed beneficiaries by mail with a telephone follow-up of nonrespondents. We used a multiple-mode, multiple-contact approach that has proved very successful on surveys conducted with the Medicare population and incorporates suggestions from Jenkins and Dillman's best mail survey practices guidelines (Jenkins and Dillman, 1997). Beneficiaries were surveyed once during the intervention period. The HBC program's survey was conducted between June 11, 2007 and October 10, 2007 among the original population beneficiaries only (excluding those that were carved out). Thus, the intervention beneficiaries had an opportunity to have had up to 21 months of programmatic exposure at the time of the survey given the original population's program start date of February 1, 2006.

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. The sample frame included only beneficiaries from the original population after the carve-out or roughly 800 beneficiaries in each of the intervention and comparison groups. From the sample

frame for each group, we randomly selected 300/.7 = 429 beneficiaries. The response rate for the HBC program was 82%. 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 HBC Program

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

Table 3-1
Medicare Health Services Survey: Estimated intervention effects for experience and satisfaction with care,
HBC
(N = 673)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.	Effect size
Helping to cope with a chronic condition (1 to 5)	3.76	3.70	0.08	N/S	0.08
Number of helpful discussion topics (0 to 5)	2.02	1.92	0.11	N/S	0.06
Discussing treatment choices (1 to 4)	3.13	3.15	-0.01	N/S	0.01
Communicating with providers (0 to 100)	75.7	75.4	2.2	N/S	0.10
Getting answers to questions quickly (0 to 100)	62.9	59.7	4.3	N/S	0.16
Multimorbidity Hassles score (0 to 24)	3.06	3.41	-0.44	N/S	0.12

NOTES: HBC = Health Buddy® Consortium; 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 3.8 for the intervention group and 3.7 for the comparison group, or about midway between "very good" and "good" ratings. Over sixty percent of the HBC beneficiaries rated their experience as "excellent" or "very good" and approximately a third selected "good." It is not uncommon among the elderly to see high satisfaction ratings. For that reason, the mean scale score was used in the analyses so that transitions between all response categories would be captured. For this overall satisfaction measure, we observe no statistically significant intervention effect for the HBC, indicating that CMHCB demonstration did not produce greater beneficiary satisfaction with care.

Across the six measures of experience and satisfaction with care, we observe no statistically significant positive intervention effects for the HBC, suggesting that the HBC intervention failed to produce a difference in any of the experience and satisfaction with care domains measured by the Medicare Health Services Survey. Effect sizes, which express the group difference as a fraction of the pooled standard deviation for the outcome, ranged from 0.01 to 0.16 for these measures.

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 HBC program.

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. In the intervention group about 56% of the HBC beneficiaries report receiving help with setting goals and 51% report receiving help making a care plan. Similarly, in the comparison group 55% report receiving help on each of these respective outcomes. The ANCOVA results reveal the HBC was not effective at increasing the proportion of intervention beneficiaries who had received help to set goals for self-care management. The HBC was also not effective at increasing the proportion of intervention beneficiaries reporting that they had help from their health care team in making health care plans. For the HBC, there were a few other covariates that predicted receiving help on these two measures: females were significantly more likely to get help setting goals, and mail survey respondents were more likely to receive help with making a care plan. The HBC intervention beneficiaries with additional insurance coverage were less likely to receive help with making a care plan.

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, the HBC 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, the HBC beneficiaries in the intervention group rated their confidence in taking prescription medications

Table 3-2
Medicare Health Services Survey: Estimated intervention effects, self-management,
HBC
(N = 673)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.	Effect size
Percent receiving help setting goals	55.5	54.9	2.3	N/S	0.05
Percent receiving help making a care plan	50.6	54.9	-2.8	N/S	-0.06
Self-efficacy ratings	4.47	4.41	0.09	N/S	0.08
Take all medications (1 to 5)					
Plan meals and snacks (1 to 5)	4.00	3.94	0.10	N/S	0.08
Exercise 2 or 3 times weekly (1 to 5)	3.47	3.31	0.20	N/S	0.14
Self-care activities					
Prescribed medications taken (mean # of days)	6.79	6.64	0.18	*	0.17
Followed healthy eating plan (mean # of days)	5.12	5.05	0.15	N/S	0.07
30 minutes of continuous physical activity (mean # of days)	2.76	2.79	-0.03	N/S	-0.01

NOTES: HBC = Health Buddy® Consortium; 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

4.5, compared to 4.4 in the comparison group. Confidence in planning meals and snacks was rated 4.0 and 3.9, respectively, and confidence in exercising was rated as 3.5 and 3.3 respectively. The confidence levels mirrored somewhat the frequency with which beneficiaries reported performing particular self-care activities, as reported later in this section. For the HBC, we found no significant intervention effects in beneficiary confidence in taking medications, planning healthy meals and snacks, and engaging in physical exercise.

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 among other activities (exercise). For example, the mean number of days that the HBC beneficiaries said they take their medications as prescribed ranged from 6.8 to 6.6; the mean number of days that the HBC beneficiaries reported following a healthy eating plan was about 5.1, and the mean number of days the HBC beneficiaries reported exercising was 2.8 out of 7 days. For self-care activities, we found one

significant positive intervention effect for the frequency of prescribed medications behavior with an effect size of 0.17. The HBC intervention increased their beneficiaries' compliance with prescription drugs. The intervention effect for healthy eating was not significant, and the effect on physical activity was slightly negative.

Physical and mental function—***Table 3-3*** displays the mental and physical functioning outcomes for the Health Buddy® Consortium. The mean PHC score for the intervention group was 27.3, compared to 26.7 for the comparison group. The mean MHC score for the intervention group was 37.3 and the PHQ-2 score of 2.0, compared to 36.4 and 2.2 for the comparison group. The PHC and MHC scores are far below the normative scores of 50 for the general adult population. The ANCOVA estimation revealed no statistically significant intervention effects for physical and mental function outcomes. The effect sizes for these three functioning outcomes were quite small, ranging from 0.11 to 0.14. Several characteristics in the ANCOVA model proved to be significant predictors for physical and mental health outcomes. PHC scores for the HBC beneficiaries increase significantly with age and additional years of education and significantly decrease for those with higher HCC scores and for those who completed the survey by mail. For MHC scores the pattern is similar: the scores increase significantly with age and additional years of education and decrease with higher HCC scores and for mail respondents.

Table 3-3
Medicare Health Services Survey: Estimated intervention effects, physical and mental health function, HBC
(N = 673)

Outcome	Intervention mean	Comparison mean	ANCOVA-adjusted intervention effect	Stat. sig.	Effect size
PHC score (physical health, mean =50, std=10)	27.3	26.7	0.8	N/S	0.11
MHC score (mental health, mean =50, std=10)	37.3	36.4	1.2	N/S	0.11
PHQ-2 score (depression, 0 to 6)	2.01	2.19	-0.28	N/S	0.14
Number of ADLs difficult to do (0 to 6)	2.36	2.56	-0.17	N/S	0.09
Number ADLs receiving help (0 to 6)	0.90	1.07	-0.17	N/S	0.11

NOTES: HBC = Health Buddy® Consortium; 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

In addition, for the HBC beneficiaries the MHC scores are significantly lower for those who used a proxy to respond to the survey. As higher PHQ scores indicate greater depressive symptoms, the results for PHQ mirror those for MHC: PHQ scores decrease with higher age and more years of education. PHQ scores are lower for females, proxy respondents, and those with additional health coverage.

Activities of daily living—On average, the HBC respondents reported limitations on about two and a half activities of daily living (ADLs) and received help with an average of one activity of daily living. We found no statistically significant intervention effects in ADL outcomes for the HBC. Among the HBC members, when other characteristics are held constant, females report significantly more ADL limitations than males, and African Americans report more than members of other races, and proxy respondents more than self-respondents. As expected, those with higher baseline HCC score also report higher levels of functional impairment. The HBC members who have additional health insurance coverage report fewer ADL limitations than those who only have Medicare. In terms of needing help with ADLs, the patterns are similar: females, proxy respondents, and members with higher baseline HCC score report needing help on a significantly higher number of ADLs. Somewhat counter-intuitively, those who live alone report needing help on fewer ADLs.

3.1.4 Conclusions

The HCB 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 disease management demonstration involves the use of a proprietary device, the Health Buddy,® in participants' homes to collect information on vital signs, symptoms, behaviors and knowledge of individuals' health conditions and transmit to multi-specialty medical groups. The HBC demonstration staff hypothesized that better communication with providers will mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions.

In summarizing the HBC intervention effect on changes in beneficiary experience and satisfaction with care, self-management behaviors, and self-reported physical and mental health functioning, survey results indicate that among 19 CMHCB demonstration survey outcome measures, the HBC achieved a positive intervention effect on only one measure within the self-management survey domain. The HBC intervention resulted in a higher frequency of medication compliance for beneficiaries in the intervention group relative to the comparison group. This was also the outcome with the largest effect size (0.17).

The absence of substantively important intervention effects is also reflected by the effect sizes computed for each survey outcome. An effect size of 0.20 is frequently considered to represent a "small" effect, and larger effects are required for clinical relevance. The study had sufficient statistical power to detect effect sizes even smaller than 0.20. However, the largest intervention effect size (and the only statistically significant one) was 0.17. The mean effect size across all 19 outcomes was 0.08.

3.2 Physician Satisfaction

RTI made two site visits to meet with the HBC program staff during the demonstration period. The first site visit was conducted on August 2-4, 2006 in Bend, Oregon and Wenatchee, Washington, the two locations where the HBC CMHCB program is being implemented. During this visit, RTI evaluators consulted with the senior management of the HBC and key HBC CMHCB program staff. The second visit to the Health Buddy® Consortium that included in-person visits to the BMC and WVMC was conducted on April 8 and 10, 2008. During this visit, RTI staff met with representatives from RBHC, the HBC clinical and managerial staff from the two clinics, and physicians who have Medicare beneficiaries from their primary care panels participating in the demonstration. Following the site visits, representatives of the AMGA and RBHC were consulted via a conference call for additional information.

3.2.1 The HBC Provider Feedback

Care managers reported that the Health Buddy® program, in general, works very well as a long-term program; however, they perceived that the Health Buddy® device may lose some of its effectiveness after 6 to 9 months for some patients. It was felt that the Health Buddy® device could be a powerful adjunct to care management during times of transition in medical treatment or between care settings.

Most physicians supported the general concept and potential benefits of the program but also expressed frustration with several aspects of the current demonstration design: too few patients whom they believed would benefit were participating because there was no physician referral process, and care managers were not embedded in their physical practice locations. The physicians also felt that the Health Buddy® device might work best as a short-term adjunct for some of their patients. When they were asked to summarize their experiences with the demonstration program, they identified four types of benefits: patient education, patient compliance, more appropriate physician contact, and more focused visits.

3.2.2 The HBC Provider Recommendations

The HBC clinic staff identified features of the demonstration that they felt interfered with beneficiaries receiving the full benefit of the intervention and suggested enhancements to the current demonstration design and care management operations. In the future, the program should consider several changes. One improvement proposed was featuring a care management structure that pairs care managers and participants' primary care physicians in the same physical location. The proximity of the care managers to the primary care physicians varied between WVMC and Bend and within the clinics at WVMC. The second proposed improvement had to do with excluding beneficiaries from practices outside the care management organizations, if a systematic means of communicating with clinicians from these practices is not established. The staff also suggested implementing a physician referral model to gain physician buy-in and to identify sufficient numbers of patients to make a financially viable care management program. A physician referral model could increase enrollment by more than 10 times, according to one physician's estimate, with which others agreed. Interviewed physicians and care managers felt that a physician referral model would increase the appropriateness of patients referred for care management services. It was recommended that patient-specific clinical or educational goals

accompany an open physician referral model in order to ensure that participants have clearly identified goals against which to measure their progress.

This page intentionally left blank.

CHAPTER 4
PARTICIPATION RATES IN THE HBC CMHCB DEMONSTRATION PROGRAM AND LEVEL OF INTERVENTION

4.1 Introduction

Our participation analysis is designed to critically evaluate the level of engagement by the HBC program 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 the HBC programmatic 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 HBC program engage their intended audiences? Specific research questions include the following:

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

- How many individuals did the HBC program 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 the HBC program?

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

- What beneficiary characteristics predict a high level of HBC demonstration 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. In our August 2006 site visit, HBC staff reported that BMC had enrolled 162 patients (25% of its population) and WVMC had 333 patients (35% of its population). Approximately 25% of BMC patients and 17% of WVMC patients had opted out of the program and refused to be contacted further about the Health Buddy® program (Brody and McCall, 2006). By February 2008, the participation rates among the original and refresh populations were nearly identical within each clinic. BMC recruited 149 participants (30 percent of the original population plus refresh eligible beneficiaries), while WVMC engaged 445 participants (44 percent of the original population plus refresh eligible beneficiaries) (McCall 2010). In our first analysis of participation in the CMHCB demonstration, we examined participation during the initial 6-month outreach period of the demonstration (McCall et al., 2008). In this report, we

examine the level of participation for the full intervention period and the beneficiary characteristics that predict participation.

We also examine the level of intervention between the HBC program and its beneficiaries with the Health Buddy® device. The main intervention for the HBC program is the Health Buddy® health monitoring device, which collects qualitative and quantitative survey information from beneficiaries on a daily basis. The HBC program also offers an alternate program for beneficiaries who are unable or unwilling to use the Health Buddy® device. This program also involves care management support provided through routinely scheduled telephone calls with nurse care managers or telephone calls in response to data transmitted through the Health Buddy®. During the routine calls, nurses ask participants who do not use the Health Buddy® device similar questions to those programmed into the device. However, these responses are not entered into the Health Buddy® desktop—the data repository used to create the intervention data files. Thus, the intervention data files contain only information from beneficiaries who use the device. Therefore, we examine the number of telephonic contacts between HBC staff and their participants with the Health Buddy® device. For each participating beneficiary, the HBC program provided RTI with a count of the number of telephonic contacts by type: inbound and outbound. The HBC program also provided information on who was contacted (e.g., caregiver, patient, or physician) and number of completed surveys.

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 HBC 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 HBC program) and the percentage of beneficiaries participating for more than 75% of their eligible months.[8] These latter two sets of numbers provide an estimate of the number of beneficiaries with whom the HBC program 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

[8] 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), develops ESRD, elects the hospice benefit, or dies.

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 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 day versus at least 75% of eligible months). The participation model investigates whether group membership is influenced by beneficiary demographic attributes, clinical characteristics, and utilization and cost factors previously defined in ***Chapter 2***. The demographic variables included in the model are defined as follows from the Medicare EDB and determined as of the date of assignment for the original population (January 2, 2006) and the refresh population (January 2, 2007):

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

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

- baseline HCC score medium and high, two dichotomous variables set at 1 if the prospective HCC score was between 2.0 and 3.1 (medium) and greater than 3.1 (high); HCC score less than 2.0 is the reference group;

- baseline Charlson score medium and high, two dichotomous variables set at 1 if the Charlson index score was 2 or 3 (medium) and 4 or greater than (high); Charlson score of less than 2 is the reference group for the original and refresh populations.

- 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 HBC original demonstration program was greater than or equal to $366 and less than $1,055 (medium) and $1,055 or greater (high); PBPM cost less than $366 is the reference group for the original population. For the refresh population, baseline PBPM costs greater than or equal to $307.50 and less than $1,082 were assigned to the medium group and $1,082 or greater to the high category; PBPM cost less than $307.50 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 0.606 but less than 1.54 (medium) and greater than or equal to 1.54 (high); concurrent HCC score less than or equal to 0.606 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 between 0.447 and 1.2 and values greater than or equal to 1.2 were assigned to the high category; a concurrent HCC score less than or equal to 0.447 is the reference group.

4.2.2 Level of Intervention Analysis Methods

The HBC program provided RTI with the number and nature of contacts with participating beneficiaries at the beneficiary level for the first 13 quarters of the CMHCB demonstration (February 1, 2006 through April 30, 2009). Because data were submitted quarterly and we wanted to include all the data for the full 38 months of the original population and 26 months of the refresh, we have an extra month of data included in the last quarter. We use these data to develop estimates of the level of intervention provided to Health Buddy® device participants. The core of the intervention was one-on-one care manager support provided via telephone. Care managers monitor patient responses to surveys conducted via the device and follow up with patients to help them address clinical issues and initiate interventions as needed to maintain their health (Brody and McCall, 2006). The device engages and educates patients so they may better understand their health conditions and proactively manage their disease by modifying high-risk behaviors. Routine monitoring of patient health status and symptoms alerts providers to health issues that require early intervention in an effort to avert serious complications requiring hospitalization.

Using the encounter data submitted by the HBC program, we constructed counts of the number of telephonic contacts with Health Buddy® device participants (both inbound and outbound), in total, and by who was contacted or doing the contacting: patient, provider, or caregiver. We report the mean and median number of total contacts and the distribution of beneficiaries across six categories of contacts (0, 1, 2-4, 5-9, 10-19, and 20 or more). We also estimate a multivariate logistic regression model of the likelihood of being in the high total contact category relative to the low total contact category. A dichotomous dependent variable was created and set at 1 for beneficiaries who had a high level of contact with the HBC program and 0 for beneficiaries who had a low level of contact based upon the distributional properties of number of contacts. Beneficiaries who had a medium level of contact with the HBC program 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 15-26 for the refresh population; and

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

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

4.3 Findings

4.3.1 Participation Rates for the HBC Program 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 the HBC program is 38 months for the original population and 26 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 as the total number of beneficiaries weighted by the number of days eligible in the intervention period divided by the total number of days in the intervention period. For example, a beneficiary in the HBC program had a

total of 38 months (or 1,155 days) of possible enrollment. If he/she died after 90 days, their FTE value would be 90/1,155 or 0.078 FTEs. If someone were eligible for all 38 months, then his or her value is 1. The sum of this value across all beneficiaries gives the total FTE value reported.

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

The ratio of FTEs to the total number of eligible beneficiaries in the original intervention population is 0.77 for the entire intervention period (months 1-38) compared with higher ratios (greater than 0.90) for each individual year of the demonstration. These differences in ratios illustrate the effect of subsetting to beneficiaries in the different time periods and attrition over time of the original beneficiaries due primarily to death. Beneficiaries also became ineligible for participation in the HBC demonstration program if they joined a Medicare Advantage (MA) plan, lost Medicare Part A or B eligibility or Medicare became a secondary payer, developed ESRD, elected the hospice benefit, or moved out of the service area.

Forty-five percent of the original intervention and 51% 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 HBC program's nonparticipant group was eligible only 71% 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 49% for the nonparticipant group.

Table 4-2 displays eligibility data for the refresh population, which is nearly 40% larger than the size of the original population which excludes the carve-out beneficiaries. The ratio of total number of beneficiaries to FTEs was lower for the full 26 months (0.85) 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 (78%) than nonparticipants (66%) or the comparison group (67%), but the difference narrows by the last 12 months of the demonstration (87%, 84%, and 83%, respectively).

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

Characteristics	Months 1-38	Months 1-12	Months 13-24	Months 27-38
Intervention group				
Number eligible[1]	763	763	640	517
Full time equivalent[2]	584	708	585	468
Number fully eligible	416	636	526	423
Participants				
Number eligible	346	317	301	196
Full time equivalent	288	303	290	182
Number fully eligible	212	281	266	164
Participants > 75%				
Number eligible	212	208	166	168
Full time equivalent	183	201	157	158
Number fully eligible	141	187	144	142
Non-participants				
Number eligible	417	446	339	321
Full time equivalent	295	405	296	286
Number fully eligible	204	355	260	259
Comparison group				
Number eligible	805	804	673	528
Full time equivalent	602	752	601	466
Number fully eligible	392	671	528	399

NOTES:

FFS = fee-for-service; HBC = Health Buddy® Consortium; 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/HealthBuddy/final/tables/tabHB-1.sas 23APR2010.

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

Characteristics	Months 1-26	Months 1-12	Months 15-26
Intervention group			
Number eligible[1]	1,038	1,038	870
Full time equivalent[2]	885	967	805
Number fully eligible	736	886	738
Participants			
Number eligible	413	404	341
Full time equivalent	378	395	320
Number fully eligible	324	375	296
Participants > 75%			
Number eligible	260	162	306
Full time equivalent	249	160	287
Number fully eligible	222	153	264
Non-participants			
Number eligible	625	634	529
Full time equivalent	507	572	486
Number fully eligible	412	511	442
Comparison group			
Number eligible	1,041	1,040	852
Full time equivalent	865	945	786
Number fully eligible	699	859	710

NOTES:

FFS = fee-for-service; HBC = Health Buddy® Consortium; 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/HealthBuddy/final/tables/tabHB-1.sas 23APR2010.

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

Participation rates for the HBC program original population. Of all HBC demonstration program original intervention group beneficiaries, 45% verbally consented to participate in its program at some point during the intervention period. We previously reported (McCall et al., 2008) that, among the carve-in population, 40% consented in the initial 6-month engagement period and we observe a slight increase in the HBC program's enrollment over the entire intervention period. Only 15% of beneficiaries were continuous participants (***Table 4-3***), which equates to one-third of participants. Among the HBC program beneficiaries, 48% refused to participate. The percent not contacted or unable to be located was 6%.

Participation rates were heavily influenced by length of eligibility during the intervention period. An alternative measure of participation is the percentage of beneficiaries who participated more than 75% of months they were eligible for the CMHCB demonstration. Of the HBC's original intervention beneficiaries, 28% participated for more than 75% of their eligible months, which is much higher than the continuous participant percentage. ***Table 4-3*** also reports the number of participants over time (for months 6, 12, 24 and 38, 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 HBC program 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. For the original population, beneficiaries had to have annual costs of $6,000 or more and an HCC score greater than or equal to 1.7. For the refresh population, those criteria were not used. Instead, the HBC specified tiers of qualification thresholds (based on beneficiary utilization of services) for each of the four diagnostic inclusion categories they specified (HF, DM, COPD, and comorbidity) After a number of iterations reviewing tabulations produced by ARC, HBC selected tier 1 (2 or more noninpatient encounters)for HF, tier 3 (4 or more noninpatient claims plus 1 or more inpatient claim) for DM, tier 2 (8 or more noninpatient claims) for COPD and tier 2 (4 or more noninpatient claims) for comorbid conditions. With the selection criterion change, there was no improvement in their participation rate, in fact it decreased (***Table 4-4***). Overall, 40% of the refresh intervention beneficiaries consented to participate at some point during the 26-month period. Of those, 22% were continuous participants, which equates to 55% of participants. The percent that refused to participate was modestly lower (43%), and the percent not contacted or unable to be contacted was higher at 18%.

Table 4-3
Participation in the HBC CMHCB demonstration program: Original population

Characteristics	Statistic
Number of intervention months	38
Participation rate (entire demonstration period)	45%
Length of participation	
Continuous participation after engagement	15%
After initial participation, became a continuous nonparticipant	20%
Intermittent participation	10%
Nonparticipation (never agreed)	55%
Refused to participate when contacted	48%
Not contacted/unable to be contacted	6%
Beneficiaries participating more than 75% of months	28%
Number of participants in selected months[1]	
Month 6	283
Month 12	262
Month 24	208
Month 38 (last month)	131

NOTES: HBC = Health Buddy® Consortium; 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/HealthBuddy/final/tables/tableHB-2.sas 23APR2010.

Table 4-4
Participation in the HBC CMHCB demonstration program: Refresh population

Characteristics	Statistic
Number of intervention months	26
Participation rate (entire demonstration period)	40%
Length of participation	
Continuous participation after engagement	22%
After initial participation, became a continuous nonparticipant	15%
Intermittent participation	2%
Nonparticipation (never agreed)	60%
Refused to participate when contacted	43%
Not contacted/unable to be contacted	18%
Beneficiaries participating more than 75% of months	25%
Number of participants in selected months[1]	
Month 6	294
Month 12	355
Month 26 (last month)	238

NOTES: HBC = Health Buddy® Consortium; 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/HealthBuddy/final/tables/tableHB-2.sas 23APR2010.

4.3.2 Characteristics of the HBC Program Intervention and Comparison Populations

In addition to evaluating the level of initial engagement by the HBC program, our participation analysis is designed to confirm that the selection procedures produced similar demographic, disease, and economic burden profiles between the intervention and comparison groups for both the original and refresh populations. Identifying any systematic baseline differences in demographic characteristics, health status, or baseline chronic condition patterns between the intervention and comparison group beneficiaries is important because the

contractual and financial benchmarks established as part of the CMHCB demonstration program are based on an ITT framework and an assumption that the intervention and comparison groups are equivalent or essentially equivalent at the start of the demonstration.

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

Characteristics of the HBC program original population—In ***Supplement*** 4A we report the beneficiary characteristics for the intervention and comparison groups for the original and refresh populations. ***Supplement 4A,*** also reports the beneficiary characteristics that predict participation in the HBC CMHCB 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 lower percentages of the disabled, men, beneficiaries under the age of 65 and ages 85 and older and higher rates of diabetes with complications, essential hypertension, valve disorders, and lipid metabolism disorders. The refresh population has fewer statistically significant differences in the beneficiary characteristics – primarily in higher percentages of heart-related comorbidities among the intervention beneficiaries. ***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 HBC 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 HBC demonstration program more than 75% of their eligible months. Model 1 reflects the initial recruitment emphasis by the HBC program, or characteristics of beneficiaries with whom the HBC program 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 HBC demonstration. For each model, we estimated two equations; an equation with just demographic characteristics and a full model equation that includes baseline and demonstration utilization and health status variables. Because there is correlation between beneficiary characteristics and the other variables, such as health status and baseline characteristics, we were most interested in examining which beneficiary characteristics had the greatest effect on willingness to participate before controlling for these other factors. The results for all three models were very similar in direction and magnitude of effect of beneficiary characteristics on the likelihood of participation so we do not display results of Models 1 and 2 in the body of the text (see ***Supplement 4A***).

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

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

Characteristics	Model 3A OR	p^3	Model 3B OR	p^3
Intercept	0.58	**	0.43	**
Beneficiary characteristics				
Male	0.79	N/S	0.79	N/S
African American/other/unknown	0.30	N/S	0.31	N/S
Age < 65 years	1.02	N/S	1.02	N/S
Age 75-84	0.96	N/S	0.99	N/S
Age 85 + years	0.51	*	0.61	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	0.89	N/S
Baseline HCC score high	N/I	N/I	0.71	N/S
Medium baseline PBPM	N/I	N/I	1.05	N/S
High baseline PBPM	N/I	N/I	1.21	N/S
Baseline Charlson score medium	N/I	N/I	1.38	N/S
Baseline Charlson score high	N/I	N/I	1.63	N/S
Demonstration period health status				
Died	N/I	N/I	0.59	N/S
Concurrent HCC score medium	N/I	N/I	1.35	*
Concurrent HCC score high	N/I	N/I	1.19	N/S
Number of cases	763	N/A	763	N/A
Chi-square (p<)	9.41	N/S	19.57	N/S
Pseudo R-square	0.01	N/A	0.03	N/A

NOTES: HBC = Health Buddy® Consortium; 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 < $366. 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, partab3b, and partab4b 27APR2010.

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

Characteristics	Model 3A OR	p^3	Model 3B OR	p^3
Intercept	0.42	**	0.29	**
Beneficiary characteristics				
Male	0.98	N/S	0.97	N/S
African American/other/unknown	0.39	N/S	0.39	N/S
Age < 65 years	0.92	N/S	0.87	N/S
Age 75-84	1.05	N/S	1.13	N/S
Age 85 + years	0.72	N/S	0.83	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.05	N/S
Baseline HCC score high	N/I	N/I	0.66	N/S
Medium baseline PBPM	N/I	N/I	0.87	N/S
High baseline PBPM	N/I	N/I	1,41	N/S
Baseline Charlson score medium	N/I	N/I	1.20	N/S
Baseline Charlson score high	N/I	N/I	1.30	N/S
Demonstration period health status				
Died	N/I	N/I	0.40	**
Concurrent HCC score medium	N/I	N/I	1.54	*
Concurrent HCC score high	N/I	N/I	1.57	*
Number of cases	1,038	N/A	1,038	N/A
Chi-square (p<)	6.62	N/S	30.78	**
Pseudo R-square	0.01	N/A	0.03	N/A

NOTES: HBC = Health Buddy® Consortium; 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 < $307.50. The baseline Charlson score reference group is < 2. The concurrent HCC score reference group is 0.447 or less.

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

Program: bene02, partab3b, and partab4b 27APR2010.

Model 3a shows that beneficiaries who were 85 years of age and older were less likely to be a participant, a proxy for poorer health status (***Table 4-5***). Examining Model 3b for the original population (***Table 4-5***), the introduction of baseline and demonstration period health status measures negates the influence of age participation status. Beneficiaries with medium concurrent HCC scores were more likely to participate than those with a low concurrent HCC score, holding other factors constant. Given that there were only 212 beneficiaries in the participant group, it is difficult to capture any statistically significant differences between the participants and nonparticipants.

There are a few noted differences in the results for the refresh population (***Table 4-6***) such as medium and high concurrent HCC scores were positive predictors of participation, indicating more success in engaging the sicker reference beneficiaries into their program. Also, beneficiaries that died were less likely to participate. During the second site visit, HBC physicians felt that too few patients whom they believed would benefit were participating because there was no physician referral process, and care managers were not embedded in their physical practice locations. HBC staff also noted that beneficiaries who were more difficult to convince to participate initially were not as likely to fully engage in the Health Buddy® program. Further, the care managers reported that they had limited success gaining participation agreement from beneficiaries who had initially declined and then subsequently experienced a sentinel event.

4.3.4 Level of Intervention

In this section, we report the frequency of interaction between the HBC program intervention beneficiaries for a subset of intervention population beneficiaries who had the Health Buddy® device at any point during the HBC program. HBC only provided encounter data for beneficiaries with the Health Buddy® device. The Health Buddy® is a health monitoring device that collects qualitative and quantitative information from patients on a daily basis. Care managers monitor patient responses to surveys conducted via the device and follow up with patients to help them address clinical issues and initiate interventions as needed to maintain their health. We also examine whether there is evidence of selective targeting of beneficiaries for intervention contacts based upon level of perceived need as determined by beneficiary demographic, health status, baseline costliness, and acute care utilization during the demonstration period. The HBC program's target population had a high prevalence of comorbid conditions, such as diabetes, HF, and COPD.

Descriptive statistics were performed using beneficiaries participating in the HBC demonstration program to determine the breadth and depth of contacts related to care management. RTI received quarterly data from HBC, thus, the reported 13 quarters of data represent information on beneficiaries with the Health Buddy® device at any point during 39 months for the original population and 27 months for the refresh population. ***Table 4-7*** provides counts of beneficiaries that had the Health Buddy® device by quarter and the percent of eligible beneficiaries with the device. Roughly 40% of the HBC program original and refresh eligible intervention beneficiaries used the Health Buddy® device during the demonstration period.

Tables 4-8 and 4-9 provide the number of beneficiaries that had the Health Buddy® device at any point during the demonstration, the length of time they had the device, and their utilization of the device (as measured by the number of surveys completed on the device). ***Table 4-8*** provides this

information for the original population. There were 346 beneficiaries that agreed to participate in the HBC program. Of these, 296 (86%) agreed to use the device for at least 1 quarter during the full 39 month period. On average, beneficiaries had the device for 8 of the 13 quarters and completed 463 surveys, which equates to about 56 surveys per quarter. Of the 463 surveys, 60 (13%) included high risk responses (knowledge, behavior, symptoms or general high risk), which were intended to be triggers for care managers responses. The majority of high risk responses were categorized as high risk symptoms responses.

Table 4-7
Frequency and percent of HBC eligible beneficiaries with the Health Buddy® device by quarter

Quarter	Number of beneficiaries – original population	Percent of eligibles	Number of beneficiaries – refresh population	Percent of eligibles
Never had a device	467	61.2	666	64.2
1	195	25.6	n/a	n/a
2	256	33.6	n/a	n/a
3	242	31.7	n/a	n/a
4	222	29.1	n/a	n/a
5	196	25.7	171	16.5
6	176	23.1	274	26.4
7	179	23.5	286	27.6
8	162	21.2	291	28.0
9	147	19.3	274	26.4
10	140	18.3	245	23.6
11	125	16.4	217	20.9
12	111	14.5	194	18.7
13	98	12.8	171	16.5

[1] Beneficiaries had to be eligible at any point during the HBC demonstration and have the Health Buddy® device

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab1.sas 04AUG2010

Table 4-8
Mean and median number of surveys and high risk responses completed by those beneficiaries with the Health Buddy® device: original population

Statistic	Number	
Number of beneficiaries with the Health Buddy® device[1]	296	—
FTE beneficiaries with the Health Buddy® device[2]	247	—
Measures of Health Buddy® device utilization	Mean	Median
Number of quarters with the Health Buddy® device	8	9
Number of completed surveys	463	389
Number of high risk knowledge responses	2	1
Number of high risk behavior responses	6	2
Number of high risk symptoms responses	48	22
Number of high risk general responses	4	3
Number of total high risk responses	60	33

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration and have the Health Buddy® device.

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

[3] Beneficiaries had to have completed at least one survey during the demonstration

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab2.sas 04AUG2010

Among the refresh population (***Table 4-9***), there were 413 beneficiaries that agreed to participate in the HBC program. Of these, 372 (90%) agreed to use the device for at least 1 quarter during the full 27 month period. On average, beneficiaries had the device for 6 of the 9 quarters and completed 317 surveys, which equates to about 54 surveys per quarter. Of those 463 surveys, 41 (13%) included high risk responses.

The HBC program provided data on the number of telephonic contacts per beneficiary with the Health Buddy® device by quarter. ***Table 4-10*** provides a summary of these contacts by type of contact (outbound and inbound) and by who was contacted (patient, physician, or care manager). The majority of contacts were made by the care managers to the patient (about 60%) followed by physician calls to the care manager (about 20%) for both the original and refresh populations. Outbound telephonic contact was the dominant form of contact (70%).

Table 4-9
Mean and median number of surveys and high risk responses completed by those beneficiaries with the Health Buddy® device: Refresh population

Statistic	Number	
Number of beneficiaries with the Health Buddy® device[1]	372	—
FTE beneficiaries with the Health Buddy® device[2]	341	—
Measures of Health Buddy® device utilization	Mean	Median
Number of quarters with the Health Buddy® device	6	6
Number of completed surveys	317	277
Number of high risk knowledge responses	1	0
Number of high risk behavior responses	6	2
Number of high risk symptoms responses	31	13
Number of high risk general responses	3	2
Number of total high risk responses	41	21

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration and have the Health Buddy® device.

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

[3] Beneficiaries had to have completed at least one survey during the demonstration

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

H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab2.sas 04AUG2010

Table 4-10
Frequency distribution of HBC Care Manager interactions: Total contacts[1,2]

Contacted	Original Frequency	Percent	Refresh Frequency	Percent
Outbound total	7,814	70.4	7,312	68.9
Patient	6,756	60.8	6,331	59.6
Physician	1,058	9.5	981	9.2
Inbound total	3,290	29.6	3,301	31.1
Physician to Care Manager	2,232	20.1	2,320	21.9
Patient to Care Manager	1,058	9.5	981	9.2
Total contacts	11,105	100.0	10,613	100.0

NOTES: HBC = Health Buddy® Consortium.

[1] Beneficiaries had to be eligible at some point during the HBC demonstration and have the Health Buddy® device.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab2.sas 04AUG2010

Table 4-11 displays the mean number of telephonic contacts and quarters of contact for the original population beneficiaries with the Health Buddy® device (n = 296). It also provides the overall distribution of telephonic contacts for the original population. Observations were weighted by the fraction of eligible days, accounting for fewer contacts due to attrition because of death, which resulted in 247 full-time equivalent beneficiaries. The mean number of contacts for each beneficiary was 45 and the median was 32. On average, there was at least one telephonic correspondence with or regarding the beneficiary in 7 of the 13 quarters. One-quarter of beneficiaries had less than 14 contacts and nearly 50% of beneficiaries had 36 or more contacts over the 13 quarter period.

Table 4-12 displays this same information for the refresh population. A total of 372 unique refresh population beneficiaries met the inclusion criteria for this analysis (341 full-time equivalents). The refresh population had a higher percentage of beneficiaries with less than 14 contacts (37%) and a lower percentage of beneficiaries with 36 or more contacts (27%). Given the shorter intervention period, it is not surprising to find these lower numbers.

Table 4-11
Distribution of number of contacts with participants[1.2.3] in the HBC program: Original intervention population

Statistic	Number	Percent
Mean number of contacts	45	—
Median number of contacts	32	—
Mean number of quarters of contact	7	—
Median number of quarters of contact	8	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-13 contacts	59	24.0%
14-35 contacts	73	29.6%
36+ contacts	115	46.4%
Total	247	100.0%

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Participants are defined as beneficiaries with the Health Buddy® device.

[2] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab3.sas 04AUG2010.

Table 4-12
Distribution of number of contacts with participants[1,2,3] in the HBC program: Refresh intervention population

Statistic	Number	Percent
Mean number of contacts	31	—
Median number of contacts	22	—
Mean number of months of contact	5	—
Median number of months of contact	5	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-13 contacts	127	37.4%
14-35 contacts	123	36.0%
36+ contacts	91	26.6%
Total	341	100.0%

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Participants are defined as beneficiaries with the Health Buddy® device.

[2] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab3.sas 04AUG2010.

Table 4-13 displays the percent of Health Buddy® device participants with care manager interactions – telephone contacts inbound and outbound, and total contacts (all telephonic) by frequency of contact over the full 39 months for the original population. Outbound calls are care manager calls to a patient or a physician. Inbound calls are defined as calls to the care manager from the beneficiary or a physician. Given that outbound telephonic contact is most frequent, we find that more beneficiaries have at least 1 outbound call (97% compared to 89% for inbound contact) and 57% have 20 or more outbound calls compared to inbound contacts (22%). Less than 3% of beneficiaries had no telephonic contact, with approximately six percent of beneficiaries having 1 to 4 contacts during the 39-month period. Nearly 70% had 20 or more telephonic contacts of some form. This indicates that beneficiaries with the Health Buddy® device were in frequent contact with their care manager and their care manager and physician were also in frequent contact. Similar results can be found for the refresh population (***Table 4-14***), except that there are lower percentages of beneficiaries receiving 20 or more calls, which is a function of the shorter time period (27 months).

Table 4-13
Percent distribution of participants[1] with HBC care manager interactions: Original intervention population

Type and frequency of contact	Number of FTE beneficiaries[2,3]	Percent
Telephonic inbound		
0	28	11.2
1	19	7.7
2-4	37	15.1
5-9	52	20.9
10-19	58	23.3
20+	54	21.7
Telephonic outbound		
0	6	2.6
1	6	2.4
2-4	14	5.8
5-9	31	12.7
10-19	48	19.3
20+	142	57.3
Total telephonic		
0	6	2.6
1	3	1.4
2-4	10	4.1
5-9	23	9.3
10-19	35	14.1
20+	170	68.6

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Participants are defined as beneficiaries with the Health Buddy® device.

[2] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

Data Sources: RTI analysis of 2006-2009 Medicare enrollment, eligibility, and HBC encounter data.
Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab4.sas 04AUG2010.

Table 4-14
Percent distribution of participants[1] with HBC care manager interactions: Refresh intervention population

Type and frequency of contact	Number of FTE beneficiaries[2,3]	Percent
Telephonic inbound		
0	52	15.2
1	31	9.2
2-4	82	24.0
5-9	77	22.7
10-19	53	15.6
20+	45	13.3
Telephonic outbound		
0	12	3.4
1	16	4.8
2-4	45	13.2
5-9	44	12.9
10-19	83	24.2
20+	141	41.5
Total telephonic		
0	7	2.1
1	14	4.1
2-4	28	8.3
5-9	42	12.3
10-19	71	20.7
20+	179	52.5

NOTES: HBC = Health Buddy® Consortium; FTE = full time equivalent.

[1] Participants are defined as beneficiaries with the Health Buddy® device.

[2] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab4.sas 04AUG2010

Table 4-15 displays the frequency of care manager contacts by baseline HCC score and type of telephonic contact. Contact by mode was not mutually exclusive in that a beneficiary could have a combination of inbound and outbound telephone contacts any time during the demonstration period. Beneficiaries were stratified into three HCC categories ranging from an HCC score greater than 3.1 to less than 2.0.

Table 4-15
Frequency of HBC contacts by HCC score:
Original intervention population

Contact mode	HCC Score Low (<2) N = 73 Frequency	%	HCC Score Medium (2-<3.1) N = 94 Frequency	%	HCC Score High (>3.1) N = 80 Frequency	%
Telephonic inbound						
0	8	11.2	12	12.5	8	9.8
1	7	9.3	7	7.1	6	7.0
2-4	12	16.6	13	14.0	12	15.0
5-9	10	13.8	20	20.9	22	27.3
10-19	20	27.2	22	23.2	16	19.9
20+	16	21.8	21	22.2	17	21.0
Telephonic outbound						
0	2	3.0	4	3.9	1	0.6
1	1	1.6	2	1.7	3	4.1
2-4	5	7.0	7	7.1	2	3.1
5-9	11	15.1	9	10.0	11	13.5
10-19	15	20.5	15	15.8	18	22.2
20+	38	52.8	58	61.5	45	56.5
Total telephonic						
0	2	3.0	4	3.9	1	0.6
1	0	0.2	2	1.7	2	2.1
2-4	1	1.6	5	5.5	4	4.9
5-9	13	18.0	7	7.4	3	3.6
10-19	7	10.2	11	11.6	16	20.4
20+	49	67.0	66	69.9	55	68.4

NOTES: HBC = Health Buddy® Consortium; HCC =Hierarchical Condition Category; N = number of beneficiaries.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab4.sas 04AUG2010

There is a high level of outbound telephonic contact across the three risk categories. When examining the two highest categories of outbound calls, there are no meaningful differences across the risk categories – more than 70% of participants received 10 or more calls during the demonstration period. The high risk group had nearly 80% of participants receiving 10 or more calls. There is no difference in the percent of beneficiaries that received one or more contacts when all both modes of telephone contact are combined – basically nearly every beneficiary received at least one contact. Over one-third of participants across all the risk categories had twenty or more contacts. Similar results are found for the refresh population (***Table 4-16***).

Table 4-16
Frequency of HBC contacts by HCC score:
Refresh intervention population

Contact mode	HCC Score Low (<2) N = 102 Frequency	%	HCC Score Medium (2-<3.1) N = 153 Frequency	%	HCC Score High (>3.1) N = 86 Frequency	%
Telephonic inbound						
0	15	14.7	25	16.0	12	14.4
1	6	6.3	16	10.3	9	10.7
2-4	27	26.8	38	25.1	16	18.5
5-9	24	23.2	33	21.5	21	24.2
10-19	17	16.5	24	16.0	12	13.8
20+	13	12.5	17	11.0	16	18.3
Telephonic outbound						
0	1	0.9	8	5.3	3	3.0
1	5	5.1	9	5.9	2	2.4
2-4	13	12.3	19	12.3	14	16.0
5-9	17	16.9	19	12.5	7	8.7
10-19	27	26.1	36	23.3	20	23.7
20+	40	38.8	62	40.7	40	46.2
Total telephonic						
0	0	0.0	5	3.3	2	2.3
1	2	2.0	9	6.2	3	3.1
2-4	11	10.8	11	7.4	6	7.0
5-9	12	11.9	19	12.7	10	12.0
10-19	23	23.0	29	19.2	18	20.7
20+	53	52.3	78	51.3	47	54.9

NOTES: HBC = Health Buddy® Consortium; HCC =Hierarchical Condition Category; N = number of beneficiaries.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/HealthBuddy/final/enctab4.sas 04AUG2010

[1] Beneficiaries had to be fully eligible and full participants in the last 18 months of the KTBH demonstration.

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

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

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/ktbh/enctab1 16MAR2010

To more directly examine the targeting strategy of the HBC program, a multivariate logistic regression model was estimated with the number of total contacts (inbound and outbound telephone calls) as the dependent variable. The model estimates the likelihood of a participant receiving a high number of contacts. The medium contact group was omitted, thus comparing the high contact group to the low contact group. ***Tables 4-17*** (original population) and ***4-18*** (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 demonstration, and their number of contacts categorized either as low (0-13) or high (36+). Odds ratios are partial in the sense that all other variables are held constant. For example, the odds of a beneficiary younger than 65 years of age experiencing a high contact rate are 2.55 times greater than those for a beneficiary age 65 and older, adjusting for any baseline difference in HCC score and other characteristics.

For the original population, beneficiaries ages 75-84 were less likely to be in the high contact category. There were no other beneficiary characteristics or baseline characteristics found to be a statistically significant indicator of the likelihood of being in the high contact category (***Table 4-17***). Demonstration period acute care utilization was not a strong predictor of a high level of contact and likely reflects the challenges that the HBC staff expressed in knowing when one of their participants had been to an emergency room or hospitalized. None of the other demonstration period health status measures were found to be statistically significant. The explanatory power of the studied beneficiary characteristics was extremely low, suggesting that there is not a strong set of variables that predict likelihood of a beneficiary being in the high contact group. The pseudo R-square for this model was 0.11. Another challenge to finding statistically significant results is the very low number of observations: there are 84 beneficiaries in the low contact category and 125 in the high contact group. These numbers become even smaller once they are weighted by eligibility (59 and 115, respectively) which also indicates that a high percentage of the low contact category lost eligibility.

For the refresh population, none of the beneficiary, baseline, or demonstration period health status characteristics were found to be statistically significant indicators of the likelihood of being in the high contact category. Again, these models face the challenge of very small numbers of observations and the pseudo R-square for this model is even lower (0.04).

4.4 Summary

For the HBC program, we find that participants from the original population were healthier and younger than beneficiaries who never participated. The very old (85 years of age and older), the disabled, and beneficiaries with chronic comorbid conditions such as diabetes with complications were less likely to be participants. In the multivariate regression analysis, the same baseline health status characteristics had no impact on the likelihood of participation after controlling for baseline demographics and demonstration period health status. Beneficiaries with medium concurrent HCC scores were more likely to be participants. This suggests that the HBC program was unable to engage the historically sicker Medicare beneficiaries but did make some inroads with engaging those with acute clinical deterioration as measured by the concurrent HCC score. The results for the refresh population were similar to the original population, with two

Table 4-17
Logistic regression modeling results comparing the likelihood of being in the HBC program high contact category relative to the low contact category: Original intervention population

Characteristics	Odds ratio[1,2]	p[3]
Intercept	2.82	N/S
Beneficiary characteristics		
Male	0.55	N/S
Age <65	2.55	N/S
Age 75-84	0.38	*
Age 85+ years	0.45	N/S
Baseline characteristics		
Baseline HCC score medium	1.19	N/S
Baseline HCC score high	0.81	N/S
Medium base PBPM	1.68	N/S
High base PBPM	1.73	N/S
Baseline Charlson score medium	1.10	N/S
Baseline Charlson score high	2.06	N/S
Demonstration period health status		
Died	0.54	N/S
Concurrent HCC score medium	0.68	N/S
Concurrent HCC score high	1.53	N/S
One hospitalization	1.21	N/S
Multiple hospitalizations	0.98	N/S
Number of cases	209	N/A
Chi-square (p<)	23.36	N/S
Pseudo R2	0.11	N/A

NOTES: HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration and have the Health Buddy® device.

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

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

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

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $366. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .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/HBC/ enctab3, enctab5 04AUG2010.

Table 4-18
Logistic regression modeling results comparing the likelihood of being in the HBC program high contact category relative to the low contact category: Refresh intervention population

Characteristics	Odds ratio[1,2]	*P*[3]
Intercept	0.35	*
Beneficiary characteristics		
Male	0.84	N/S
Age <65	1.44	N/S
Age 75-84	1.71	N/S
Age 85+ years	1.32	N/S
Baseline characteristics		
Baseline HCC score medium	1.11	N/S
Baseline HCC score high	1.37	N/S
Medium base PBPM	1.51	N/S
High base PBPM	1.18	N/S
Baseline Charlson score medium	0.92	N/S
Baseline Charlson score high	1.10	N/S
Demonstration period health status		
Died	0.71	N/S
Concurrent HCC score medium	1.33	N/S
Concurrent HCC score high	1.81	N/S
One hospitalization	0.77	N/S
Multiple hospitalizations	0.66	N/S
Number of cases	239	N/A
Chi-square (p<)	9.67	N/S
Pseudo R2	0.04	N/A

NOTES: HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1] Beneficiaries had to be eligible at any point during the HBC demonstration.

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

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

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

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $307.50. The baseline Charlson score reference group is LT 2. The concurrent HCC score reference group is .447 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/HBC/ enctab3, enctab5 04AUG2010.

noted differences: both medium and high concurrent HCC scores were positive predictors of participation and beneficiaries that died were less likely to participate. These differences suggest that the HBC program was more successful engaging the sicker and more costly beneficiaries in their program as it matured.

A cornerstone of the HBC's program was the Health Buddy® device and interactions with care managers; however, nearly 60% of eligible beneficiaries never used the device. Of the beneficiaries participating in the program and using the Health Buddy® device, nearly all beneficiaries received at least one call from a care manager during the demonstration and nearly 60% received more than 20 contacts during this same time period. Other than routine contact with the Health Buddy® device , outbound telephone contact with the care managers was the most dominant form of contact. In our multivariate regression modeling of likelihood of being in a high contact versus low contact group for the original population, we found that beneficiary characteristics, baseline characteristics, and demonstration period acute care utilization were not indicators of being in the high contact category. The small sample sizes made it difficult to determine statistically significant differences.

CHAPTER 5
CLINICAL QUALITY PERFORMANCE

5.1 Introduction

RTI's analysis of quality of care focuses on measuring effectiveness of the Health Buddy® Consortium's demonstration program by answering the following evaluation question:

- *Clinical Quality of Care:* Did the HBC's demonstration program 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 HBC demonstration program 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 the HBC financially responsible for quality of care improvements, RBHC does not track clinical outcomes. However, WVMC did monitor a small number of quality of care measures and conducted analyses against their general population performance. At our most recent site visit, they did report seeing a higher rate of compliance among their Health Buddy® patients than their general population. This analysis will provide CMS with additional information on performance against the comparison populations.

5.2 Methodology

We created the process-of-care measures for the 12-month period immediately prior to the go-live date for the HBC program for its original and refresh populations and for two intervention periods (months 7-18 and months 27-38) for its original population and for one intervention period (months 15-26, 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 HBC

Statistics	All	Diabetes	Ischemic vascular disease
Original beneficiaries			
Months 7-18			
Intervention			
Total number of beneficiaries	710	363	206
Full time equivalents[1]	709	362	206
Comparison			
Total number of beneficiaries	755	330	238
Full time equivalents[1]	753	329	238
Months 27-38			
Intervention			
Total number of beneficiaries	517	265	137
Full time equivalents[1]	517	265	137
Comparison			
Total number of beneficiaries	528	250	163
Full time equivalents[1]	527	249	163
Refresh beneficiaries			
Months 15-26			
Intervention			
Total number of beneficiaries	870	415	308
Full time equivalents[1]	867	414	307
Comparison			
Total number of beneficiaries	852	411	318
Full time equivalents[1]	851	411	318

NOTES: HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries.

[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: gcc01, gcc02, gcctab, gcc_rob, gcctabx, gcctab1 23APR2010.

Medicare claims for the full baseline and intervention period were included regardless of beneficiary eligibility for the HBC demonstration (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 HBC program 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 HBC demonstration program—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 HBC demonstration program 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 the HBC demonstration program 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 the HBC demonstration 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 period; 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 either period. We report on the natural trends observed in the comparison and

intervention populations over the 3-year period.[9] 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 HBC 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). Rates of the four measures calculated for the pre-demonstration period in the original comparison group are relatively high ranging from 60% for influenza vaccine to 90% for HbA1c testing for beneficiaries with diabetes. Yet, we observe even higher rates of baseline HbA1c and LDL-C testing among the original intervention beneficiaries. The baseline rate of HbA1c testing among beneficiaries with diabetes is 95%. We do not observe this pattern for influenza vaccination; the rate of vaccination among intervention beneficiaries is 12 percentage points lower than the comparison group at baseline. Over the course of the two demonstration periods for the original comparison population, we observe stable or declining rates of receipt. Most notable is a 10 percentage point decline in rate of LDL-C testing among beneficiaries with IVD in the last 12 months of the demonstration period. We observe declines in the rates of all measures for the original intervention population. Not surprisingly, we observe only modest separation in the difference-in-differences rates; none are statistically significant.

For the refresh comparison population, the rate of influenza vaccination is 9 percentage points lower than observed for the original comparison group; however, we do observe a 6 percentage point increase during the demonstration period. The other three rates are either the same or higher than observed for the original comparison population; and, the comparison group rates are more closely aligned with the refresh intervention rates at baseline. Over the last 12 months of the demonstration, there is one statistically significant difference. The rate of receipt of the influenza vaccine among the refresh intervention beneficiaries declined by 1 percentage point while the rate of receipt among the comparison beneficiaries increased 6 percentage points. Thus, the D-in-D change is -7 percentage points, which is a statistically significant negative intervention effect. There are no other statistically significant differences.

[9] 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 ill fee-for-service (FFS) beneficiaries.

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

Process-of-care measures	Rate per 100 Baseline I[1]	Rate per 100 Baseline C[1]	Rate per 100 Demo period I[1]	Rate per 100 Demo period C[1]	D-in-D Rate per 100	D-in-D OR	D-in-D P	D-in-D CI Low	D-in-D CI High
ORIGINAL POPULATION									
Months 7-18									
All beneficiaries									
Influenza vaccine	48	60	44	60	-4.60	0.83	0.23	0.61	1.12
Beneficiaries with diabetes									
HbA1c test	95	88	92	84	1.12	0.89	0.76	0.42	1.88
LDL-C test	79	63	76	67	-6.81	0.71	0.17	0.44	1.16
Beneficiaries with IVD[2]									
LDL-C test	78	68	72	65	-2.48	0.85	0.60	0.46	1.56
Months 27-38									
All beneficiaries									
Influenza vaccine	48	62	43	57	0.26	1.02	0.93	0.71	1.45
Beneficiaries with diabetes									
HbA1c test	95	90	92	88	-0.76	0.78	0.59	0.31	1.94
LDL-C test	81	66	78	66	-2.61	0.85	0.59	0.48	1.52
Beneficiaries with IVD[2]									
LDL-C test	85	74	74	64	-1.52	0.77	0.52	0.35	1.71
REFRESH POPULATION									
Months 15-26									
All beneficiaries									
Influenza vaccine	38	51	37	57	-7.05	0.75	0.04	0.57	0.99
Beneficiaries with diabetes									
HbA1c test	93	92	93	88	3.97	1.55	0.22	0.77	3.14
LDL-C test	80	77	81	71	7.81	1.52	0.08	0.95	2.45
Beneficiaries with IVD[2]									
LDL-C test	81	74	78	64	5.65	1.24	0.43	0.73	2.10

NOTES: HBC = Health Buddy® Consortium; 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; CMO = care management organization.

[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 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.

SOURCE: RTI analysis of 2005-2009 Medicare enrollment, eligibility, claims and encounter data; Computer runs: gcc01, gcc02, gcctab, gcc_rob, gcctabx, gcctab1 23APR2010.

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

- always compliant, meaning compliant in both baseline and intervention periods;
- became noncompliant, meaning compliant in the baseline period but noncompliant in the intervention period;
- never compliant, meaning noncompliant in both the baseline and intervention 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 HBC CMHCB demonstration: 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	79%	88%	52%	67%	55%	66%	46%	32%
Became noncompliant	10	6	12	13	17	16	15	15
Never compliant	5	2	21	10	18	11	26	41
Became compliant	6	4	14	10	10	7	13	12
Refresh population	C	I	C	I	C	I	C	I
Always compliant	82	87	59	69	50	67	40	25
Became noncompliant	10	5	19	12	24	15	12	13
Never compliant	3	2	10	6	12	7	31	49
Became compliant	6	5	12	13	14	11	17	12

NOTES: HBC = Health Buddy® Consortium; CMHCB = Medicare Care Management for High Cost Beneficiaries; LDL-C = low-density lipoprotein cholesterol; IVD = ischemic vascular disease; C = comparison population; I= intervention population; CMO = care management organization.

[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 10MAY2010.

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 HbA1c and LDL-C testing than the comparison beneficiaries with the exception of influenza vaccination; the intervention beneficiaries were 14 percentage points less likely to always be compliant than the comparison beneficiaries. It is not surprising that we see lower rates of never compliant for HbA1c and LDL-C testing and a higher rate of never compliant for influenza vaccination among the intervention beneficiaries. There are not many differences in the rates of beneficiaries that became compliant or noncompliant. Given the higher baseline rates of compliance for the intervention group, we observe somewhat smaller percentages of beneficiaries becoming compliant during the last 12 months of the demonstration. A similar 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 HBC demonstration program on quality of care. Specifically, we report findings for the key research question: did HBC 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 baseline rates are because of differences in beneficiary or provider behavior. In general, we observe higher rates of baseline compliance among the intervention beneficiaries. The noted exception is influenza vaccination.

Within the original and refresh intervention and comparison populations, we generally observe stable or negative trends in the rates. The original intervention group's rates tended to fall more than its comparison group's rates in 6-of-8 measurements; while the refresh intervention group's rates tended to fall less than its comparison group's rates in 3-of-4 measurements. The difference-in-differences (D-in-D) rates per 100 beneficiaries ranged from 1 to -7 per 100 beneficiaries for the original population and 8 to -7 per 100 beneficiaries for the refresh population. Of these differences, there is one that is statistically significant. The rate of receipt of the influenza vaccine among the refresh intervention beneficiaries declined by 1 percentage point while the rate of receipt among the refresh comparison beneficiaries increased 6 percentage points. Thus, the D-in-D change is -7 per 100 beneficiaries.

This page intentionally left blank.

CHAPTER 6
HEALTH OUTCOMES

6.1 Introduction

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

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

- Did the HBC program 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 HBC demonstration 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 Methodology

6.2.1 Rates of Hospitalizations and Emergency Room Visits

Rates of hospitalization and ER visits were constructed for the 12-month period immediately prior to the launch of the HBC demonstration program 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*** provides 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 the HBC CMHCB demonstration 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 incidence rate ratio's associated *p* value and 95% confidence interval.

6.2.2 Rates of 90-Day Readmissions

We estimated the percent of beneficiaries with at least one readmission within 90 days of discharge and the readmission rate per 1,000 beneficiaries. Readmissions are estimated for index admissions that occurred during 12-month spans in the baseline and demonstration periods. For the baseline, we included index admissions in the 12-month period immediately prior to the go-live date of the HBC's program. 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 24 through 35 and included readmissions through months 21 and 38, respectively. The intervention period for the refresh population examined admissions during months 12 through 23 and readmissions through month 26.

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 HBC program 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 (OR) from the logistic regressions and the incidence rate ratio (IRR) 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 HBC program 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 enrollment data base (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 a multivariate logistic regression analysis to determine the predictors of mortality controlling for baseline differences in beneficiary demographic and health status characteristics between the intervention and comparison groups with inclusion of an indicator for being in the intervention group and completing at least one Health Buddy® survey. Both a logistic model of the likelihood of death and a Cox proportional hazard model of survival were estimated testing the relationship of a large number of independent variables with likelihood of death or time to death. The independent variables used in the final multivariate Cox proportional hazard model of survival reported are defined as follows:

- intervention status, set at 1 for beneficiaries in the intervention group;
- intervention status interacted with completion of a Health Buddy® survey, set at 1 if the intervention beneficiary agreed to use a Health Buddy® device and completed at least one survey;
- male, a dichotomous variable, set at 1 for males;

- minority race, a dichotomous variable, set at 1 for beneficiaries whose race code is African American, other, or unknown;

- age, three dichotomous variables set at 1 for age less than 65 years, age 75-84 years, and age greater than or equal to 85 years, age 65-74 years is the reference group;

- urban, a dichotomous variable, set at 1 for beneficiaries with ZIP codes within metropolitan statistical areas;

- four dichotomous variables for presence of each of the following chronic conditions: diabetes with complications, essential hypertension, valve disorders, lipid metabolism disorders;

- baseline Hierarchical Condition Category (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 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 HBC original demonstration program was greater than or equal to $366 and less than $1,055 (medium) and $1,055 or greater (high), PBPM cost less than $366 is the reference group for the original population. For the refresh population, baseline PBPM costs greater than or equal to $307.50 and less than $1,082 were assigned to the medium group and $1,082 or greater to the high category, PBPM cost less than $307.50 is the reference group; 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 0.606 but less than 1.54 (medium) and greater than or equal to 1.54 (high), concurrent HCC score less than or equal to 0.606 is the reference. These scores were re-calculated for the first 6-months of the refresh intervention period with the medium category assigned to values between 0.447 and 1.2 and values greater than or equal to 1.2 were assigned to the high category, a concurrent HCC score less than or equal to 0.447 is the reference group.

Prior to estimating the final Cox proportional hazard model, we conduct statistical tests of the underlying assumption of proportionality by creating time-dependent covariates for each of our independent variables through interaction of each variable with the logarithm of time. Thus, we are testing whether the proportional effect of an independent variable varies with time. Violation of this assumption requires alternative modeling, either through the estimation of a parametric regression model, inclusion of a time-varying variable for the nonproportional predictors, or stratification. The final models presented in this report do not violate the proportionality without the need for use of a parametic model, inclusion of time-varying parameters, or stratification.

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 HBC 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.

Not unexpectedly, the baseline rates of hospitalization and ER visits were high in the HBC intervention and comparison populations. The baseline rate of all-cause hospitalization was 506 per 1,000 original intervention group beneficiaries (***Table 6-1***). And, the baseline rate of all-cause ER visits was 1,080 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 HBC demonstration program 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 and ER visits increased similarly in the original intervention and the comparison groups between the baseline and both demonstration periods. The trend D-in-D rate is negative for all the hospitalization and ER rates but the 10 ACSC hospitalization rate during months 7-18 of the demonstration, indicating that the intervention rates increased less than the comparison group's rates, with none of the differences being statistically significant. In months 7-18, the D-in-D rate for all-cause hospitalizations is 2 per 1,000 beneficiaries lower in the intervention group than the comparison group (p= 0.99). In months 27-38, the D-in-D rate for all-cause hospitalizations is 79 per 1,000 beneficiaries lower in the intervention group than the comparison group (p=0.41). Although there was a 4% lower rate of growth in the intervention group's all-cause ER utilization rate during months 27-38 of the demonstration, this was not a statistically significant difference[10].

[10] The percentage change in the D-in-D intervention rate is calculated by estimating the percent change in the comparison group's utilization between baseline and the demonstration period and applying the percent change to the intervention group's baseline rate. This produces an expected rate based upon the observed change in the comparison group. The percent change for the intervention rate is calculated using the expected rate as the baseline rate.

Table 6-1
Comparison of rates of utilization for months 7-18 and the last 12 months of the HBC CMHCB demonstration with rates of utilization for a 1-year period prior to the start of the HBC CMHCB demonstration: 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	506	519	628	643	-2	1.00	0.99	0.78	1.29
10 ACSCs[5]	183	191	235	236	7	1.04	0.86	0.69	1.55
ED/Obs visits									
All-cause	1,080	1,274	1,322	1,603	-88	0.97	0.81	0.77	1.23
10 ACSCs	323	344	366	454	-67	0.86	0.39	0.61	1.21
Months 27-38									
Hospitalizations									
All-cause	486	457	686	737	-79	0.88	0.41	0.64	1.20
10 ACSCs	176	156	271	258	-7	0.93	0.78	0.56	1.54
ED/Obs visits									
All-cause	997	1,127	1,445	1,708	-132	0.96	0.75	0.73	1.25
10 ACSCs	288	281	410	440	-37	0.91	0.66	0.59	1.40

NOTES: HBC = Health Buddy® Consortium; 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

[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 23APR2010.

Rates of hospitalization and ER visits per 1,000 refresh population beneficiaries for the year prior to go-live and months 15-26 of the HBC refresh demonstration period are presented in ***Table 6-2.*** We observe roughly similar refresh baseline rates of use among the intervention and comparison groups as we do for the original intervention and comparison groups. And, we observe an increase in the hospitalization and ER visit rates for both the intervention and comparison groups during the demonstration period with a noted exception. The rate of all-cause hospitalizations declines within the intervention group during the demonstration while the rate of all-cause hospitalizations increases within the comparison group. The overall D-in-D net effect in the all-cause hospitalization rate is -154 per 1,000 beneficiaries (p=0.02). This represents a 26% lower rate than what would have been expected. We observe no other statistically significant D-in-D rates of ACSC 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 HBC CMHCB demonstration with rates of utilization for a 1-year period prior to the start of the HBC CMHCB demonstration: 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 15-26									
Hospitalizations									
All-cause	528	474	485	585	-154	0.74	0.02	0.58	0.95
10 ACSCs[5]	129	134	205	212	-3	1.00	1.00	0.67	1.50
ED/Obs visits									
All-cause	904	1,042	1,171	1,367	-58	0.99	0.90	0.80	1.22
10 ACSCs	231	276	329	426	-52	0.92	0.65	0.66	1.30

NOTES: HBC = Health Buddy® Consortium; 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

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 with at least one day of eligibility in the baseline and demonstration period are included.

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 23APR2010.

6.3.2 Rates of 90-Day Readmissions

Table 6-3 displays the number of beneficiaries included in the readmission analyses. ***Table 6-4*** displays the percent of original population beneficiaries with an admission and 90-day readmission and rate of 90-day readmission per 1,000 beneficiaries. Data are displayed for all-cause and ACSC admissions and readmissions. In general, we observe a pattern of increasing percentage of both intervention and comparison beneficiaries being hospitalized or having a readmission over the course of the demonstration. We observe 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. 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 HBC program was more successful at reducing readmissions for beneficiaries with frequent readmissions than for beneficiaries with less frequent readmissions relative to the comparison group.

Table 6-3
Number of beneficiaries included in analyses of readmissions for the HBC CMHCB demonstration

Counts of beneficiaries	Intervention	Comparison
Original beneficiaries		
Months 7-18		
Total number of beneficiaries	710	755
Full time equivalents[1]	709	753
Months 24-35		
Total number of beneficiaries	542	547
Full time equivalents[1]	542	546
Refresh beneficiaries		
Months 12-23		
Total number of beneficiaries	903	883
Full time equivalents[1]	900	882

NOTES: HBC = Health Buddy® Consortium.

[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 23APR2010.

Table 6-4
Change in 90-day readmission[1] rates between the year prior to the HBC CMHCB demonstration and months 7-18 and months 24-35 of the demonstration: Original population

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo period rate per 1,000[1,2,3] I	Demo period rate per 1,000[1,2,3] C	D-in-D	OR/IRR[4]	*p*	Low CI	High CI
Months 7-18									
Hospitalizations									
Percent with an admission	35	32	34	33	-2	0.92	0.61	0.67	1.26
Percent with ACSC[5] admission	14	14	16	14	2	1.18	0.45	0.77	1.80
All-cause 90-day readmission									
Percent with readmission	24	27	30	32	1	1.05	0.87	0.59	1.86
Readmission rate / 1,000	324	484	541	636	64	1.27	0.35	0.77	2.10
ACSC same-cause 90-day readmission									
Percent with readmission	6	11	11	13	4	1.73	0.42	0.46	6.57
Readmission rate / 1,000	101	123	129	167	-16	0.94	0.93	0.26	3.35
Months 24-35									
Hospitalizations									
Percent with an admission	33	30	34	35	-4	0.82	0.28	0.57	1.18
Percent with ACSC admission	13	13	15	15	-0	0.97	0.90	0.59	1.59
All-cause 90-day readmission									
Percent with readmission	24	22	38	33	3	1.11	0.77	0.56	2.18
Readmission rate / 1,000	326	408	694	844	-69	1.03	0.93	0.55	1.92
ACSC same-cause 90-day readmission									
Percent with readmission	7	9	13	4	10	3.85	0.14	0.65	22.94
Readmission rate / 1,000	130	100	199	79	90	1.94	0.44	0.37	10.20

NOTES: HBC = Health Buddy® Consortium; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odds 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 26APR2010.

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 observe 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.

6.3.3 Mortality

In ***Table 6-6***, we display mortality rates during the HBC CMHCB demonstration for both the original and refresh intervention and comparison populations. Over the 38-month demonstration period for the original population, 35% of the intervention group beneficiaries died while 40% of the comparison group beneficiaries died. This five percentage point lower mortality rate within the intervention group is a statistically significant (p=0.04) difference from the observed mortality rate within the comparison group. During the 26-month demonstration period for the refresh population, just under one-quarter of both groups of beneficiaries died. The 2 percentage point difference between the intervention and comparison group of refresh beneficiaries is not statistically significant.

We further explored the lower rates of mortality in both the original and comparison populations by estimating a multivariate Cox proportional hazard model of survival. ***Figures 6-1 and 6-2*** displays unadjusted survival curves for the original and refresh populations, respectively. Each survival curve has three lines displaying the intervention and comparison groups' unadjusted survival differences with the intervention group split between those that agreed to a Health Buddy® device and completed a survey versus all other intervention group beneficiaries. The LifeTest procedure reveals that there is a statistically significant difference in survival among the three original groups and the three refresh groups of beneficiaries. Both original and refresh intervention beneficiaries, who agreed to the Health Buddy® device and completed at least one survey, appears to have superior survival performance[11]; however, these survival curves are not adjusted for covariates that could reflect willingness to use the Health Buddy® device (i.e., health status) and vary by cohort.

Table 6-7 displays three 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 large number of baseline covariates to control for any differences between the two groups at baseline. Model 3 includes a dichotomous interaction term set equal to 1, if the beneficiary is in the intervention group and completed one or more surveys with the Health Buddy® device; the variable is set to 0 otherwise. The hazard ratios and associated p values are displayed for all three sets of the 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.

---

[11] The LifeTest procedure for the unadjusted survival curves of the full intervention group versus the comparison group revealed that there is a statistically significant difference between the original full intervention group versus the comparison group but not between the intervention and comparison group of the refresh population mirroring the statistical findings from the analysis of unadjusted mortality rates.

Table 6-5
Change in 90-day readmission[1] rates between the year prior to the HBC CMHCB demonstration and months 12-23 of the demonstration: Refresh population

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo period rate per 1,000[1,2,3] I	Demo period rate per 1,000[1,2,3] C	D-in-D	OR/IRR4	p	Low CI	High CI
Months 12-23									
Hospitalizations									
Percent with an admission	37	33	28	29	-5	0.80	0.14	0.60	1.07
Percent with ACSC[5] admission	11	11	11	13	-1	0.88	0.55	0.58	1.34
All-cause 90-day readmission									
Percent with readmission	20	22	30	39	-8	0.73	0.26	0.42	1.26
Readmission rate / 1,000	269	297	542	675	-105	0.89	0.60	0.56	1.40
ACSC same-cause 90-day readmission									
Percent with readmission	10	9	12	15	-4	0.71	0.60	0.20	2.49
Readmission rate / 1,000	102	102	204	191	13	1.07	0.91	0.33	3.48

NOTES: HBC = Health Buddy® Consortium; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odds 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 26APR2010.

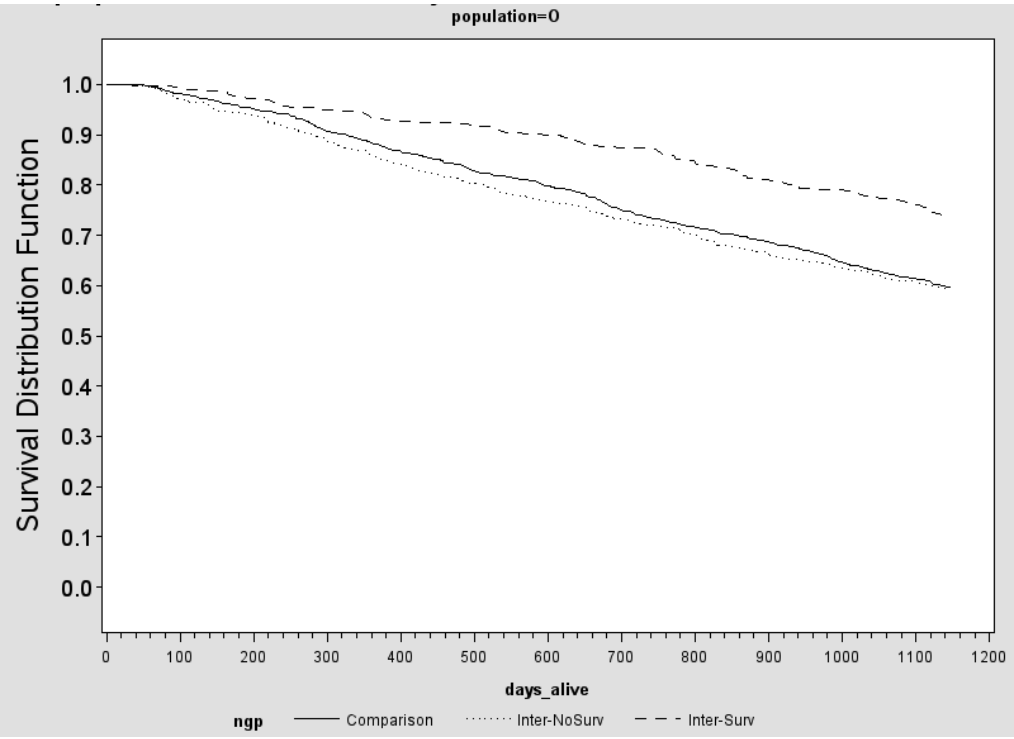
Table 6-6
Mortality rates during the HBC CMHCB demonstration: Original and refresh populations

Description	Intervention number of deaths	Percent	Comparison number of deaths	Percent	Difference	P value
Original population (38 months)	268	35.1%	324	40.3	-5.1	0.04
Refresh population (26 months)	218	21.0%	237	22.8	-1.8	0.33

NOTES: HBC = Health Buddy® Consortium; CMHCB = Medicare Care Management for High Cost Beneficiaries.

SOURCE: RTI analysis of Medicare enrollment and intervention data; Computer runs: mortality.sas 21APR2010.

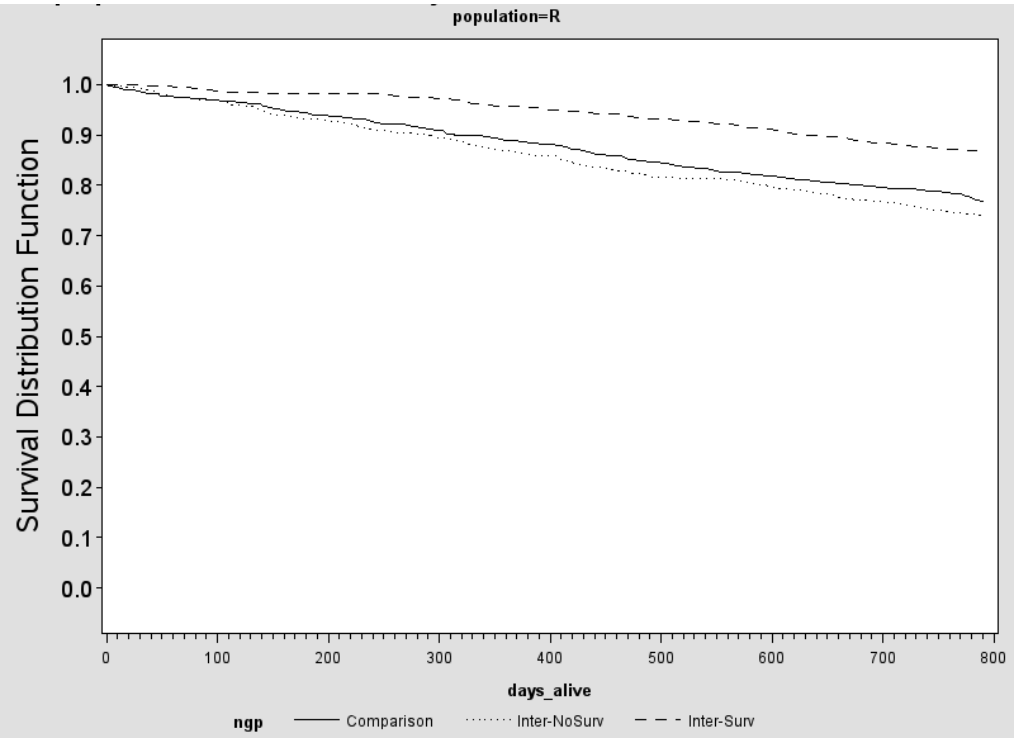
Figure 6-1
Cox proportional hazard model unadjusted survival curves for the HBC demonstration original population



NOTES: HBC = Health Buddy® Consortium; Inter-NoSurv = intervention beneficiaries who did not use the Health Buddy® device to complete a survey; Inter-Surv = intervention beneficiaries who did use the Health Buddy® device to complete at least one or more surveys; Comparison = comparison group.

SOURCE: RTI analysis of Medicare enrollment and intervention data; Computer runs: dietab3.sas 6May2010.

Figure 6-2
Cox proportional hazard model unadjusted survival curves for the HBC demonstration refresh population



NOTES: HBC = Health Buddy® Consortium; Inter-NoSurv = intervention beneficiaries who did not use the Health Buddy® device to complete a survey; Inter-Surv = intervention beneficiaries who did use the Health Buddy® device to complete at least one or more surveys; Comparison = comparison group.

SOURCE: RTI analysis of Medicare enrollment and intervention data; Computer runs: dietab3.sas 6May2010.

Table 6-7
Cox proportional hazard survival models for the HBC demonstration program: Original population

Original	**Model 1** Hazard Ratio	**Model 1** p value	**Model 2** Hazard Ratio	**Model 2** p value	**Model 3** Hazard Ratio	**Model 3** p value
Intervention	0.845	0.0417	0.880	0.1293	1.053	0.5778
Intervention_survey	N/I	N/I	N/I	N/I	0.590	<.0001
Male	N/I	N/I	1.213	0.0204	1.229	0.0131
Medicaid	N/I	N/I	0.712	0.4517	0.712	0.4514
Age <65	N/I	N/I	0.660	0.0680	0.648	0.0570
Age 75-84	N/I	N/I	1.669	<.0001	1.659	<.0001
Age ≥ 85	N/I	N/I	3.410	<.0001	3.310	<.0001
Baseline HCC score medium	N/I	N/I	0.955	0.6870	0.962	0.7300
Baseline HCC score high	N/I	N/I	1.280	0.0450	1.307	0.0293
Medium baseline PBPM	N/I	N/I	0.949	0.6759	0.934	0.5846
High baseline PBPM	N/I	N/I	1.340	0.0403	1.324	0.0495

NOTES: HBC = Health Buddy® Consortium

Program: Dietab3c; Dietab4c, Dietab4d, May 6, 2010; August 11, 2010.

Table 6-8
Cox proportional hazard survival models for the HBC demonstration program: Refresh population

Refresh	Model 1 Hazard ratio	Model 1 p value	Model 2 Hazard ratio	Model 2 P value	Model 3 Hazard ratio	Model 3 p value
Intervention	0.922	0.381	0.932	0.4459	1.114	0.2760
Intervention_survey	N/I	N/I	N/I	N/I	0.519	<.0001
Male	N/I	N/I	1.084	0.3917	1.111	0.2682
Medicaid	N/I	N/I	1.294	0.5011	1.233	0.5846
Age <65	N/I	N/I	0.680	0.1162	0.685	0.1230
Age 75-84	N/I	N/I	1.538	0.0006	1.533	0.0006
Age ≥ 85	N/I	N/I	3.045	<.0001	2.964	<.0001
Baseline HCC score medium	N/I	N/I	0.831	0.1616	0.835	0.1735
Baseline HCC score high	N/I	N/I	1.045	0.7574	1.092	0.5316
Medium baseline PBPM	N/I	N/I	1.520	0.0044	1.515	0.0046
High baseline PBPM	N/I	N/I	2.854	<.0001	2.743	<.0001

NOTES: HBC = Health Buddy® Consortium; N/I = not included

Program: Dietab3c; Dietab4c, Dietab4d, May 6, 2010; August 11, 2010

In Model 1 for the original population, we observe that the intervention variable has a hazard ratio of 0.845 implying a survival advantage to the intervention group (***Table 6-7***). It is statistically significant at the 0.04 level – similar to what we observed for the mortality rate calculation. In Model 2, we no longer observe a survival advantage among the original intervention beneficiaries when baseline covariates are added to adjust for any imbalances between the intervention and comparison groups at baseline. Beneficiaries who are male, age 75 and above, and with high baseline costs or HCC risk scores are far more likely to die that those without these characteristics. In Model 3, we introduce the interaction terms of intervention beneficiaries who completed a Health Buddy® survey. We continue to observe no survival advantage among beneficiaries in the intervention group relative to the comparison group; however, we do observe a survival benefit among the original intervention beneficiaries who used the Health Buddy® device.

For the refresh population, we do not observe a statistically significant difference in the unadjusted intervention mortality rate for the refresh population as expected (Model 1, ***Table 6-8***). Nor do we observe a survival advantage for the intervention group when baseline covariates are added in Model 2. Beneficiaries age 75 and above and with medium or high baseline HCC risk scores are far more likely to die that those without these characteristics. And, as with the original

population, we observe a survival benefit among the refresh intervention beneficiaries who used the Health Buddy® device.

In developing the final multivariate regression models, regression diagnostics were conducted to confirm the underlying assumption of proportionality of the Cox hazard model. All time-dependent variables were not statistically significant, either collectively or individually, thereby supporting the proportionality assumption. However, other survival models failed to pass the proportionality test; most notably for models that included a measure of the beneficiary's concurrent HCC risk score calculated using claims data for the first 6 months of their demonstration experience. We included this variable in preliminary models to better capture the health status of the beneficiaries at the outset of the demonstration. In models that included the concurrent HCC risk score, we observed a survival advantage for intervention beneficiaries in the original and refresh populations as well as an incremental survival advantage among beneficiaries who used the Health Buddy® device, but implausibly large hazard ratios on the concurrent HCC risk score variable with considerable variation in the level of the hazard ratios for the intervention and Health Buddy® device interaction variables. Thus, the survival models were quite sensitive to specification. This does raise caution about interpretation of the survival benefit among the intervention beneficiaries who used the Health Buddy® device. It could be that the multivariate models reported may not be capturing adequately the factors associated with agreeing to use a Health Buddy® device so we may not have controlled completely for selection bias. However, our finding suggests that further investigation of the appearance of a survival benefit is desirable by analysis of other programs that use the Health Buddy® device or the HBC Phase II demonstration mortality experience.

6.3.4 Hospice

A major component of the HBC program was encouraging appropriate end-of-life-care planning, including use of the hospice benefit. We examine rates of hospice use between the intervention and comparison groups for both the original and refresh populations. ***Table 6-9*** provides the hospice rates and the mean and median days in hospice. We observe low use rates of the Medicare hospice benefit among the original and refresh intervention and comparison populations, ranging from 8% to 15% (statistically insignificant). Length of hospice is also quite similar between the intervention and comparison groups (no differences are statistically significant) for both the original and refresh populations.

Table 6-9
Rates of Hospice use and mean and median days of Hospice use among original and refresh HBC CMHCB demonstration beneficiaries that elected the Hospice benefit

Population	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	763	805	15%	15%	0.9	0.61	54	63	-9	0.56	16	18	-2	0.87
Refresh	1,038	1,041	9%	8%	0.5	0.68	66	43	23	0.11	14	11	3	0.22

SOURCE: RTI analysis of Medicare enrollment and eligibility data; Computer runs: hsp01 hospicetab1, hsptest 12MAY2010

6.4 Conclusions

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

- Did the HBC program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?

- Did the HBC program improve health outcomes by decreasing mortality?

During the course of the HBC demonstration, we observed increasing rates of all-cause and ambulatory care sensitive conditions (ACSC) hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations with one exception. All-cause hospitalizations declined within the refresh intervention group while the rate of all-cause hospitalizations increased within the comparison group for a rate of -154 hospitalizations per 1,000 beneficiaries D-in-D rate (p=0.02). This represents a 26% lower rate than what would have been expected[12]. Although we observed no other statistically significant differential rates of growth in all-cause or ACSC hospitalizations or ER visits or 90-day readmissions, we observed a trend toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures with a number of the D-in-D rates appearing to be of clinical significance although not statistically significant. Further, we do observe wide confidence intervals for several of the readmission estimates due likely to small sample sizes. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rate of the Medicare hospice benefit or in mean and median number of days in hospice.

We do observe a statistically significant lower rate of mortality in the original population's intervention group. Over the 38-month demonstration period for the original population, 35% of the original intervention group beneficiaries died while 40% of the comparison group beneficiaries died; a 5 percentage point lower rate of mortality in the intervention group (p=0.04). Over the 26-month demonstration period for the refresh population, 21% of the refresh intervention group beneficiaries died and 23% of the comparison group beneficiaries died; a 2 percentage point lower rate of mortality in the intervention group (p=0.33).

We estimated multivariate models of survival, whereby we controlled for potential imbalances in baseline beneficiary characteristics that may be related to mortality and not adequately accounted for in the development of a comparison group. When doing so, the observed survival benefit for the intervention group within the original population was no longer present. However, when we introduced into our model a variable that captures the impact of

[12] The percentage change in the D-in-D intervention rate is calculated by estimating the percent change in the comparison group's utilization between baseline and the demonstration period and applying the percent change to the intervention group's baseline rate. This produces an expected rate based upon the observed change in the comparison group. The percent change for the intervention rate is calculated using the expected rate as the baseline rate.

intervention beneficiaries using the Health Buddy® device, we observed an incremental increase in survival benefit among both the original and refresh populations' intervention beneficiaries who used the Health Buddy® device. Because we did not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the survival benefit. However, given this important finding, additional study is warranted.

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 38 months that the Health Buddy® Consortium's (HBC) CMHCB demonstration program was in operation (or 26 months for the refresh sample). The evaluation questions we address are:

- What were the Medicare costs per beneficiary per month (PBPM) in the base year versus the first 38 or 26 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 38- or 26-month period compare with the fees that were paid out? How close was the HBC program in meeting budget neutrality?
- 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?
- 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 have been conducted for financial reconciliation by Actuarial Research Corporation (ARC) under contract to CMS. ARC determined savings based on the demonstration's terms and conditions negotiated between CMS and the HBC. A detailed explanation of the differences is provided in ***Section 7.9***.

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 the HBC program achieved budget neutrality using RTI's costing methods. ***Section 7.6*** displays stratified PBPM costs and savings by high cost and high risk categories to test for possible imbalances in the intervention and comparison groups. ***Section 7.7*** examines regression-to-the-mean (RtoM) effects. ***Section 7.8*** uses multivariate regression to control for any imbalances between intervention and comparison groups that might affect t-tests of mean differences in PBPM growth rates. ***Section 7.8*** provides

a detailed explanation of the differences in methods used by ARC for financial reconciliation and RTI for the evaluation. The chapter concludes in ***Section 7.10*** with a summary of key findings.

7.2 Data and Key Variables

7.2.1 Sample 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 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 a managed care plan). Claims represent utilization anywhere in the United States, not just the target area of the HBC program. Medicare costs are based on eligible claims submitted during the full demonstration period plus 12 months prior to the start date. A 9-month "run-out" period after the demonstration ended assures a complete set of costs.

7.2.2 Constructing PBPM costs

All financial analyses were conducted on a PBPM cost basis, or the ratio of eligible Medicare costs to eligible months. The baseline period is defined as 365 days (or 1 year) prior to the HBC program's start date. The 38-month demonstration period for the original population includes 1,156 days (38 months x 30.42 days/month) after the start date. The refresh population covers 26 months, or 790 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[13]; and
- 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 38 or 26 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

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

$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 HBC's original population (763 with eligible days in both the base and intervention period) stratified by beneficiaries' number of eligible days in the demonstration period (1,095 maximum). There were no beneficiaries with fewer than 30 eligible days. Beneficiaries with 31-60 days of eligibility averaged PBPM costs of $4,000 compared with beneficiaries eligible for a year or more who averaged PBPM costs of $1,250. Beneficiaries with truncated eligibility averaged monthly costs 3 times greater than those with much longer eligibility. Less than 3% of the sample was eligible less than 3 months. (See ***Section 7.3.2*** for statistics on PBPM variation.) Maximum intervention period PBPM costs were $19,722.

Table 7-1
HBC CMHCB demonstration period PBPM mean costs by eligible days, intervention group, original population

Eligible days[1]	N (%)	PBPM Costs	Range
< 10	0 (0.0%)	$–	$–
11–30	0 (0.0)	–	–
31–60	9 (1.2)	3,999	0–17,719
61–90	12 (1.6)	5,393	126–16,532
91–365	103 (13.5)	3,332	1–19,722
366+	639 (83.7)	1,250	11–10,950
Mean	763	3,274	0–19,722

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; N (%) = number of beneficiaries (percent of all eligible beneficiaries).

[1] Number of days beneficiary eligible for intervention.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (4/20/10).

Table 7-2 shows the unweighted cost effects of short-term eligible beneficiaries in the refresh intervention group. Again, short-eligibility beneficiaries were several times as costly per month as those with more than 1 year of eligibility. Maximum PBPM costs for the refresh population were $56,937. Note that mean reference costs are roughly one-half of the original population's costs. This is primarily due to not using a minimum cost threshold in the refresh group.

Table 7-2
HBC CMHCB demonstration period PBPM mean costs by eligible days, intervention group, refresh population

Eligible days[1]	N (%)	PBPM Costs	Range
< 10	1 (0.0)	$2,140	$0–2,140
11–30	10 (1.0)	2,836	0–8,534
31–60	15 (1.4)	6,818	29–28,001
61–90	9 (0.9)	5,195	0–24,926
91–365	115 (11.1)	3,777	27–56,937
366+	888 (85.5)	955	5–15,311
Mean	1,038	1,546	0–56,937

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; 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 CMS for financial reconciliation. While the 1% trim reduces the HBC's program'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 in the original population, or roughly 110-times less than beneficiaries with full eligibility through the entire demonstration period. This weighting method is equivalent to simply adding the beneficiary's $30,000 and 10 eligible days to total costs and days of fully eligible beneficiaries and then calculating the combined PBPM cost.

7.2.3 Monthly Fees

Demonstration Care Management Organizations (CMOs) proposed monthly fees when submitting their applications for the demonstration program to the CMS Office of Demonstrations. CMS then negotiated final fees as part of each CMO's agreed-upon contract terms and conditions. RTI benchmarked savings against each CMO's initially negotiated fee. For the HBC, its negotiated management fee was $120 for the original intervention group during the first year, $123.84 in year 2, and $127.80 in year 3. Fees for the refresh intervention group were $123.84 in year 2 and $127.80 in year 3. To be consistent with the calculation of gross savings, these two fees were weighted by the share of fee-bearing to all eligible months in the intervention group.

7.3 Analytic Methods

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

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

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

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

7.3.1 Tests of Gross Savings

Gross savings to Medicare is defined as the difference between the claims costs of the intervention and comparison groups. There are two ways to calculate these differences. Assuming that the selection process balanced the intervention and comparison populations, PBPM cost differences between the two groups can be based solely on the demonstration period. That is, the HBC program 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 HBC program began recruiting beneficiaries to the intervention. Also, because we wanted to conduct statistical tests of intervention effects, it was necessary to construct PBPM cost estimates at the beneficiary level and then use variation in the observations to produce confidence intervals around the estimates.

Recognizing that base year costs may be different between intervention and comparison populations, we used a mixed paired sample approach. First, we used each beneficiary's own mean PBPM costs in the base year just prior to the HBC program'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.[14] 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

[14] 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.[15] 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 HBC program.

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[16] of intervention savings uses the following confidence interval:

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

and the detectable threshold is

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

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

Table 7-3 and 7-4 show the variation that exists in the (unweighted) PBPM costs in the base year prior to the start date and the demonstration period for the HBC program'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 $8,785 in the original comparison group. The coefficient of variation (CV), or the standard deviation of beneficiary-level PBPM costs divided by the mean, is fairly large in the base year (standard deviations roughly 25% greater than mean costs). CVs in the original and refresh comparison groups increased slightly during the demonstration period

---

15 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.

16 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.

while they decreased slightly in the original intervention group but increased substantially in the refresh intervention group. Some of the variation is reduced after weighting observations when determining intervention savings later in this chapter.

The difference between median and mean PBPM costs indicates how skewed costs actually are. Mean costs are roughly double median costs in the original sample's base year with little change during the intervention period, indicating a strong right tail of very high costs. Costs were similarly skewed in the refresh population (***Table 7-4***). Note that 25% of refresh comparison beneficiaries had base year costs less than $147. Maximum values show how high PBPM costs can be before weighting, $30-57,000 per month. 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.

Table 7-3
HBC CMHCB demonstration program 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)	(805)	(763)	(805)	(763)
Minimum	$0	$13	$85	$0
<10%	93	159	231	287
<25%	217	293	451	531
Median	458	594	1,060	1,119
>75%	1,269	1,375	2,069	2,273
>90%	2,783	2,886	3,955	3,962
Maximum	8,785	11,166	31,242	19,722
Mean	1,003	1,150	1,731	1,792
CV	1.26	1.28	1.32	1.19

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; 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 (4/20/10).

Table 7-4
HBC CMHCB demonstration program 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)	(1,041)	(1,038)	(1,041)	(1,038)
Minimum	$0	$0	$0	$0
<10%	0	55	107	139
<25%	147	210	255	287
Median	455	602	763	697
>75%	1,351	1,486	1,857	1,626
>90%	2,874	2,949	3,693	3,458
Maximum	21,087	10,520	30,062	56,937
Mean	1,074	1,118	1,560	1,546
CV	1.57	1.27	1.60	2.06

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; 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 (4/20/10).

Because of the relatively large variances in the base year PBPM costs (CV[original comparison] = 1.26), coupled with adjustments for the repeated nature of the experimental design, the power afforded by the original sample sizes was low, i.e., about 10%.[17]

[17] Power for a comparison of two mean changes in PBPMs is given by $\phi[-1.96+(\nu n\Delta/\sigma_d \nu 2)]$ (Rosner, 2006, p. 336). $\sigma_d = [\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\rho\sigma_2]^{0.5}$, where subscript 1 and 2 pertain to variances in the intervention and comparison groups' PBPM costs, and ρ = correlation between observations between the base and intervention periods. The intervention and comparison groups' standard deviations in the base period were $1,307 and $1,203, respectively. Assuming a 0.33 intra-patient correlation, $\sigma_d = \$1{,}455$. If there were no increase in the comparison group's PBPM cost over time, then $\Delta = 0.05(\$966) = \$48.30$ (see Table 7-5). The treatment n = 763. Thus, $Power = \phi[-1.96+(\$48.30 \bullet 27.62/1{,}455 \bullet 1.41 = 0.648) = 1-\phi[1.31] = 0.10$. With the HBC original intervention group, we had a 10% likelihood of finding a significant difference if the true mean change in the intervention group's PBPM cost was $48.30 less than the change in the comparison group's PBPM cost.

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 full demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. Thus, to avoid paying back any fees with a 5% savings requirement in the original sample,

$$PBPM_I <= 0.95PBPM_c - MF \qquad (7.6a)$$

or as a fraction of the comparison PBPM cost,

$$PBPM_I/PBPM_c <= 0.95 - (MF/PBPM_c), \qquad (7.6b)$$

where $PBPM_I$, $PBPM_C$ = average monthly costs in the intervention and comparison groups, MF = the average monthly fee.

For example, if a CMO's monthly fee were 5% of the comparison PBPM cost, then intervention PBPM costs would have to be 90% or less of monthly comparison costs to avoid paying back fees. Debt obligation per intervention beneficiary month is the positive difference:

$$PBPM_I - [0.95PBPM_c + MF] . \qquad (7.6c)$$

RTI's conclusion regarding budget neutrality will differ from those of CMS during financial reconciliation, given the way we adjust for unequal base period costs, how fees are calculated, the lack of an outlier trim, and a few other minor differences. Because we use statistical confidence intervals to judge the extent of gross savings, we test whether a CMO achieved any savings at all: the Z-test against zero savings.

In addition to Z-tests of mean cost differences between the entire intervention group and the comparison group, we also tested for differences in PBPM cost growth rates between intervention beneficiary participants and nonparticipants relative to the comparison group. If the intervention had more success with those beneficiaries it actively engaged, then savings should be greater for participants than nonparticipants.

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 25 months (i.e., the sixth month of the base year to the 19th mid-period month of the demonstration). γ provides a test of the 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 38-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
HBC CMHCB demonstration PBPM cost growth 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	763	$1,044	47.3	$1,398	51.2	$353**	63.6
Participants	346	1,155	78.8	1,438	72.7	283**	98.2
Nonparticipants	417	936	54.9	1,358	72.2	422**	82.4
Comparison	805	966	42.4	1,436	56.0	470**	61.9
Differences							
I – C	—	78	63.4	-38	76.1	-117	88.7
Participants - C	—	189*	81.5	2	95.5	-188	112.1
Nonparticipants - C	—	-30	71.7	-78	94.6	-47	105.5
Participants - Nonparticipants	—	-220*	94.4	80	102.5	-140	127.1

NOTE: HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; 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(4/20/10).

Overall. The weighted base year average PBPM cost was $78 (8%) more (*p* = insig) in the intervention group versus the comparison group ($1,044 versus $966). The intervention-comparison difference in PBPM Medicare costs reversed to -$38 (*p*=insig) in the demonstration period ($1,398 versus $1,436). Intervention beneficiaries, who were 8% more costly on a weighted basis at baseline, became 3% less costly, on average, than the comparison group after 38 months. Between the base year and the end of the 38-month demonstration period, the average comparison group PBPM cost increased significantly by $470 (*p* <.01), while the intervention group's PBPM average Medicare costs rose more slowly by $353 (*p* <.01). Consequently, the intervention group's PBPM cost rose -$117 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 45% (346/763 - 1). Participant costs in the HBC intervention group were 20% higher ($189; p<.05) than in the comparison group in the base period. Nonparticipants were $30 less costly (*p* = insig). Participant costs were essentially identical to comparison costs over the demonstration period. Nonparticipants became -$78 less costly (*p* = insig) during the demonstration period versus -$30 less initially. Thus, the -$117 slower growth in intervention PBPM costs appears to be due in large part to slower growth in the participant group (-$188; *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 26-month demonstration period for the refresh population. The weighted base year average PBPM cost was $18 more (*p* = insig) in the intervention versus comparison group ($1,064 versus $1,046). The intervention-comparison gap in PBPM Medicare costs reversed in the demonstration period ($1,155 versus $1,211). The average comparison group's PBPM cost increased $165 (p<.01) while the intervention group's PBPM average Medicare costs increased $92 (*p*=insig). As a result, the intervention group's PBPM cost increased -$73 slower (*p* = insig) compared with the comparison group's PBPM cost. Intervention beneficiaries, who were 1.7% more costly at baseline, were 4.6% less costly than the comparison group, on average, after 19 months between the mid-points of the baseline and demonstration periods.

The participation rate, based on beneficiaries used in the refresh cost analysis, was 40% (413/1,028 – 1). Participants in the base period in the HBC intervention group were $87 more costly (*p* = insig) than comparison group beneficiaries and nonparticipants were -$34 less costly (*p* = insig). Participants became $160 less costly (p<.056) during the demonstration period. Nonparticipants became $23 more costly (*p* = insig) during the demonstration period. Consequently, the participant group's PBPM cost rose -$247 more slowly (p<.05) than the comparison group's cost while the nonparticipant group's PBPM cost rose $56 faster (*p*=insig) than the comparison group's PBPM cost.

Table 7-6
HBC CMHCB demonstration 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	1,028	$1,064	41.2	$1,155	58.1	92	63.9
Participants	413	1,133	68.6	1,051	61.9	-82	82.3
Nonparticipants	625	1,012	51.0	1,233	88.8	221*	91.9
Comparison	1,041	1,046	50.8	1,211	48.2	165**	60.6
Differences							
I - C	—	18	65.3	-56	75.6	-73	88.1
Participants - C	—	87	89.3	-160	83.9	-247*	106.8
Nonparticipants - C	—	-34	76.8	23	92.8	56	106.0
Participants - Nonparticipants	—	121	83.3	-183	117.4	-303*	128.9

NOTE: HBC = Health Buddy® Consortium; CMHCB = Care Management for High Cost Beneficiaries; 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 (4/20/10).

7.5 Savings and Budget Neutrality

7.5.1 Original Population

Table 7-7 presents summary statistics on savings from the HBC's 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 HBC's monthly fee is reported also as a percentage of the comparison group's PBPM cost. Over the course of the 38-month intervention, average monthly costs increased $353 in the intervention group and $470 in the comparison group. The result was a -$117 relative decrease in PBPM cost growth in the intervention group. This negative difference implies *gross savings* at a rate of 8.1% of the comparison group's demonstration period PBPM cost. However, savings were statistically insignificant.

Table 7-7
Average PBPM gross savings, fees, and budget neutrality status, original population: HBC

Description	PBPM cost change
Intervention group	$353
Comparison group	470
Difference	-$117
Gross (dis)saving %[1]	-8.1%
Minimal Detectable Savings[2]	
Absolute	-$174
% of comparison PBPM[3]	-12.1%
Monthly Fee	
Absolute[4]	$58
% of comparison PBPM[3]	4.0%
Net Fee	
Absolute[5]	-$59
% of comparison PBPM[3]	-4.1%
Return on Investment (RoI)[6]	2.01

NOTES: HBC = Health Buddy® Consortium; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison demonstration PBPM (= $1,436). 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 ($1,436) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of $120, $123.84, $127.80 fees paid in outreach period and thereafter through month 38. Weights = fee-eligible members.

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

[6] RoI = Gross savings difference/Absolute Monthly Fee.

SOURCE: Medicare 2004-2008 Part A&B claims; PBPM cost changes and detectable savings: Table 7-5; monthly fees: Fees and members taken from ARC Final Reconciliation for Health Buddy® Phase I, April 8, 2010, Tables 3, 5 and 6.

With only 763 beneficiaries in the intervention group and only 805 in the comparison group, the minimal detectable savings threshold was -$174 at the 95% confidence level. This threshold rate was 12.1% of the comparison group's PBPM cost, implying that the intervention would have had to achieve this level of savings to be considered statistically reliable in repeated patient samples.[18]

The HBC's average monthly fee was $58 when averaged over all intervention beneficiaries, which amounted to 4% of the comparison group's PBPM cost during the demonstration period. It was so low because it was paid on a low percentage of all intervention beneficiaries. Thus, the HBC 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. An actuarial analysis that ignores statistical significance would show the HBC's intervention savings of -$117, with a net fee to Medicare of -$59 instead of +58, and a Medicare return on investment of 2.01. Because we cannot say with confidence that the savings are not zero, possibly due to small sample sizes, it is possible that the intervention's RoI is zero.

7.5.2 Refresh Population

Table 7-8 presents summary statistics on savings from the HBC intervention with the refresh sample. Over the course of the 26-month intervention, average monthly costs increased $92 in the intervention group and $165 in the comparison group. The result was a -$73 lower relative increase in PBPM costs in the intervention group. This negative difference implies *gross savings* at a rate of 6% of the comparison group's PBPM cost.

With roughly 1,000 beneficiaries in each study group, the minimal detectable refresh savings threshold was -$173 at the 95% confidence level. This rate is -14.3% of the comparison group's PBPM cost, implying that the intervention would have had to achieve this level of savings to be considered statistically reliable in repeated samples. Ignoring the fact that the -$73 in intervention savings was not statistically different from zero, the net fee to Medicare was reduced from $65 per beneficiary per month to -$8, resulting in a net Medicare cost of -0.7% of the comparison group's average monthly outlay on claims. Based on actuarial methods, Medicare's return on investment was 1.12, implying net savings (albeit statistically insignificant) of $0.12 on every dollar of Medicare fees paid out. However, the refresh RoI could also be zero in a future intervention.

[18] If minimal savings were based just on differences in PBPM costs during the demonstration period, the intervention would have to achieve a 10.3% savings rate (76.1(1.96)/$1,436) based on RTI's weighting methodology.

Table 7-8
Average PBPM gross savings, fees, and budget neutrality status, refresh population: HBC

Description	PBPM cost change
Intervention group	$92
Comparison group	$165
Difference	-$73
Gross (dis)saving %[1]	-6.0%
Minimal Detectable Savings[2]	
Absolute	-$173
% of comparison PBPM[3]	-14.3%
Monthly Fee	
Absolute[4]	$65
% of comparison PBPM[3]	5.4%
Net Fee	
Absolute[5]	-$8
% of comparison PBPM[3]	-0.7%
Return on Investment (RoI)[6]	1.12

NOTES: HBC = Health Buddy® Consortium; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison PBPM (= $1,211). 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 ($1,211) in demonstration period.

[4] Absolute Monthly Fee = Weighted average of $120, $123.84, $127.80 fees paid in outreach period and thereafter through months 13-38. Weights = fee-eligible members.

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

[6] RoI = Gross savings difference/Absolute Monthly Fee.

SOURCE: Medicare 2004-2008 Part A&B claims; PBPM cost changes and detectable savings: Table 7-6; monthly fees: Fees and members taken from ARC Final Reconciliation for Health Buddy® Phase I, April 8, 2010, Tables 3, 5 and 6.

7.6 Imbalances between Intervention and Comparison Populations

Because the HBC's comparison group was not based on random sampling, it is possible that material imbalances remained between intervention and comparison groups simply by chance. If the distribution of high cost and high risk beneficiaries differs between the HBC's intervention group and its comparison group, then demonstration period PBPM cost comparisons could be biased against the intervention, if it had a disproportionate number of high risk, more cost-increasing, beneficiaries. We created four, mutually exclusive, high-low cost-risk groups. The high-cost threshold was set at $16,000/month, or the top 25% of cases in either group based on their costs the year prior to selection for the program. 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. Because of the roughly one year lag between randomization of the original population into intervention and comparison groups and the official base year, intervention beneficiaries, compared with comparison beneficiaries, were more likely to be high cost and male. They were less likely to be over age 85, disabled, living in an urban area, and in a SNF prior to the demonstration period. Except for the sharp differences in urban location, the two refresh groups were nearly balanced in the base year.

Table 7-9
Frequency distribution of beneficiary characteristics, intervention and comparison groups, base year, original population: HBC

Characteristics	Intervention (%)	Comparison (%)
Cost-Risk Group		
High-cost > =$ 16,000	9.5%	7.5%
Both	13.7	14.8
High-risk: HCC > 2.0	11.5	13.1
Neither	65.3	64.7
Age Group		
<65	7.1	11.1
65-69	13.4	12.5
70-74	22.0	17.2
75-79	25.1	20.4
80-84	19.7	20.0
85+	12.7	18.8
Gender		
Female	46.6	54.3
Male	53.4	45.7
Race		
Minority	2.6	2.9
White	97.4	97.1
Medicaid Eligible		
No	98.6	98.2
Yes	1.4	1.8
Disabled		
No	92.9	88.9
Yes	7.1	11.1
Urban residence		
No	99.7	76.3
Yes	0.3	23.8
Long-term care		
No	100.0	100.0
Yes	0.0	0.0
SNF		
No	94.5	92.0
Yes	5.5	8.0

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. HBC = Health Buddy® Consortium; HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (4/20/10).

Table 7-10
Frequency distribution of beneficiary characteristics, intervention and comparison groups, base year, refresh population: HBC

Characteristics	Intervention (%)	Comparison (%)
Cost-Risk Group		
High-cost > =$ 16,000	12.0%	11.3%
Both	14.6	13.4
High-risk: HCC > 2.0	7.8	8.6
Neither	65.5	66.7
Age Group		
<65	8.4	9.4
65-69	14.5	14.2
70-74	20.9	20.3
75-79	20.2	20.5
80-84	19.3	16.2
85+	16.6	19.4
Gender		
Female	46.2	48.4
Male	53.8	51.7
Race		
Minority	3.3	3.5
White	96.7	96.5
Medicaid Eligible		
No	98.7	98.0
Yes	1.3	2.0
Disabled		
No	91.5	90.6
Yes	8.5	9.4
Urban residence		
No	100.0	62.4
Yes	0.0	37.6
Long-term care		
No	100.0	99.9
Yes	0.0	0.1
SNF		
No	93.1	93.3
Yes	6.9	6.7

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. HBC = Health Buddy® Consortium; HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (4/20/10).

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 $781 in the base year compared with $2,438 for high-cost only beneficiaries (3.1 times greater) and both high-cost and high-risk beneficiaries ($3,148; 4.0 times greater). Both high-cost intervention groups experienced large declines in their PBPM costs while the high risk–only group's PBPM cost almost doubled. The comparison group showed similar patterns of cost levels and trends. Costs in the base period's neither high-cost nor high-risk group rose the fastest, which is suggestive of RtoM effects. Focusing on the difference in trends at the bottom of ***Table 7-11***, despite large negative values, we find no statistically significant differences between the original intervention and comparison group growth rates.

Table 7-11
PBPM costs by cost and risk group, intervention and comparison groups, base and demonstration periods, original population: HBC

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)	(130; 17%)	—	(66: 9%)	—	(90: 12%)	—	(477; 63%)	—
Base Year	$3,148	167.4	$2,438	128.8	$781	34.0	447	14.5
Demonstration	2,102	150.2	1,447	172.4	1,484	149.9	1,228	59.5
Difference	-1,046**	224.7	-991**	209.9	703**	151.8	780**	59.6
% Change	-33%	—	-41%	—	90%	—	174%	—
Comparison (N)	(131; 16%)	—	(59; 7%)	—	(108; 13%)	—	(507; 63%)	—
Base Year	3,023	126.3	2,390	119.5	694	33.6	387	13.8
Demonstration	2,262	155.3	1,707	194.5	1,745	214.2	1,154	57.6
Difference	-761	189.2	-683	229.3	1,052	213.1	767	58.2
% Change	-25%	—	-29%	—	152%	—	198%	—
Difference-in-Differences	-285	292.8	-308	311.2	-348	270.6	13	83.2

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. HBC = Health Buddy® Consortium; 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 $16,000 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 (4/20/10).

7.6.2.2 Refresh Population

Table 7-12 presents different results on PBPM cost trends by the four cost-risk groups for the refresh population. The three high cost or high risk refresh groups showed costs rising faster, not slower, in the intervention group. All of the cost savings came from the neither group (-$191; p<.05). The large standard errors for the refresh population are noteworthy. We had little power, except in the neither group, to detect savings rates even as large as several hundred dollars per month given the small group sizes and high cost variance from year to year.

Table 7-12
PBPM costs by cost and risk group, intervention and comparison groups, base and demonstration periods, refresh population: HBC

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)	(171; 16%)	—	(120; 11%)	—	(85: 8%)	—	(678; 64%)	—
Base Year	$3,205	130.8	$2,279	85.1	$775	28.7	397	13.7
Demonstration	2,118	173.3	1,182	139.6	1,474	188.8	897	55.9
Difference	-1,087**	260.6	-1,098**	167.9	700**	186.4	500**	56.9
% Change	-34%	—	-48%	—	90%	—	126%	—
Comparison (N)	(156; 15%)	—	(107; 10%)	—	(100; 10%)	—	(662; 65%)	—
Base Year	3,514	216.7	2,593	125.6	686	34.6	335	12.7
Demonstration	2,042	173.3	1,317	179.5	1,207	159.7	1,027	48.8
Difference	-1,472**	241.8	-1,276**	197.3	521**	160.9	692**	50.0
% Change	-42%	—	-49%	—	76%	—	207%	—
Difference-in-Differences	385	357.9	178	257.5	179	244.6	-191*	75.7

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. HBC = Health Buddy® Consortium; 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 $16,000 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 (4/20/10).

7.7 Regression-to-the-Mean

Tables 7-13 and 7-14 demonstrate the extensive RtoM 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 $1,003 in the comparison group's base period in the original population **(*Table 7-13***), with an overall increase of $728. 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 $1,240 while those with initial costs greater than $4,000 experienced average decreases of $2,486. Mean costs in both periods are well above median costs and indicate a strong skewness in PBPM costs.

Table 7-13
Regression-to-the-Mean in comparison group PBPM costs, original population: HBC

Base year PBPM cost level	N	Base year PBPM cost	Demonstration period PBPM cost	Change
< $250	238	$134	$1,374	$1,240
251-500	183	361	1,380	1,019
501-750	80	618	1,919	1,302
751-1,000	52	866	2,170	1,304
1,001–1,250	44	1,119	1,894	775
1,251-1,500	37	1,361	1,738	377
1,501-1,750	29	1,653	2,734	1,081
1,751-2,000	18	1,884	1,925	41
2,001-2,250	21	2,150	2,902	751
2,251-2,500	8	2,356	1,977	-379
2,501-2,750	12	2,612	2,204	-408
2,751-3,000	16	2,867	1,960	-908
3,001-3,250	9	3,101	1,386	-1,715
3,251-3,500	10	3,328	2,207	-1,116
3,501-3,750	4	3,600	1,257	-2,343
3,751-4,000	7	3,872	2,054	-1,818
> 4,000	37	5,118	2,632	-2,486
Mean	805	1,003	1,731	728
Median	—	458	1,060	602

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (4/20/10).

Regression-to-the-mean also is quite strong in the refresh population (***Table 7-14***). Unweighted mean costs increased $486 due mostly by much larger cost increases for beneficiaries with base year costs under $750 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 7-14
Regression-to-the-Mean in comparison group PBPM costs, refresh population: HBC

Base year PBPM cost level	N	Base year PBPM cost	Demonstration period PBPM cost	Change
< $250	355	$84	$1,233	$1,149
251-500	188	358	1,091	733
501-750	103	612	1,905	1,293
751-1,000	65	868	1,354	486
1,001–1,250	46	1,100	1,659	558
1,251-1,500	48	1,369	1,099	-270
1,501-1,750	28	1,637	1,387	-250
1,751-2,000	34	1,879	3,231	1,352
2,001-2,250	24	2,136	2,321	185
2,251-2,500	25	2,386	1,726	-660
2,501-2,750	10	2,665	3,768	1,102
2,751-3,000	18	2,858	2,249	-610
3,001-3,250	10	3,138	937	-2,201
3,251-3,500	14	3,345	2,577	-769
3,501-3,750	18	3,641	1,890	-1,751
3,751-4,000	6	3,889	1,573	-2,316
> 4,000	49	6,519	3,132	-3,387
Mean	1,041	1,074	1,560	486
Median	—	455	763	308

NOTES: Observations unweighted. HBC = Health Buddy® Consortium; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 2004-2008 Part A & B claims; COSTRUN2 (4/20/10).

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. Coefficients can be interpreted as the change between each beneficiary's average demonstration and base year PBPM costs. The Intervention variable's coefficient (column 1) is the cost-saving impact associated with beneficiaries who did not use the Health Buddy® device. Beneficiaries who used a Health Buddy® device and completed at least one survey were distinguished by a dichotomous variable (Inter_surv = 1). Its coefficient is the additional cost-saving impact among intervention beneficiaries who had the device and used it at least once.

In the first column of results controlling only for each beneficiary's base period PBPM cost, the Intervention coefficient of -63 is statistically insignificant implying no reliable success in slowing beneficiary cost increases for beneficiaries not using the Health Buddy® device. The Inter_surv coefficient is positive and also highly insignificant implying no additional savings for beneficiaries using the device. (The overall Intervention coefficient excluding Inter_surv was -57,[19] implying that the cost savings of the intervention group as a whole was also insignificant.)

The base period PBPM cost coefficient (0.241; $p < .01$), when combined with the intercept coefficient, implies substantial RtoM effects on costs (0.241 - 1 = -0.759, the RtoM effect). Imagine two comparison group beneficiaries, one with a relative low ($500) and another with a relatively high ($3,000) PBPM cost in the base period. The predicted PBPM cost of the initially "low cost" comparison beneficiary would increase 2.5-fold during the intervention period, while the "high cost" beneficiary's PBPM cost would decline by almost 40%.[20] Whereas cost differences were 6:1 in the base period, they would now be compressed to 1.5:1.

RtoM effects are quite substantial but not in one direction. Including only high cost beneficiaries in the intervention group would clearly have produced even greater declines in comparison group PBPM costs during the demonstration. Major cost increases did occur among initially lower cost beneficiaries, as evidenced in ***Table 7-13***.

The second regression model controls for which cost-risk group the beneficiary was in during the base period as well as several other beneficiary characteristics. The Intervention and Inter_surv coefficients are essentially unaffected when applying the controls and still insignificant. This is true even though two of the three cost-risk groups experienced much larger

---

[19] The Intervention coefficient, alone, of -57 differs from the sum of the Intervention and Inter_surv coefficients, i.e., -46, because of the correlation of the use of the Health Buddy® device and beneficiary base period PBPM costs.

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

PBPM[base]	PBPM[demo]	PBPM Change	%Change
$500	$1,261	$761	+152%
$3,000	$1,863	-$1,137	-38%

cost increases than the neither group. Minor changes in the two intervention coefficients are due to relatively minor imbalances between the intervention and comparison groups. The PBPM base coefficient is even smaller, implying more RtoM within each of the cost-risk groups.

Table 7-15
HBC demonstration regression results: Intervention gross savings controlling for base period PBPM costs and beneficiary characteristics: Original population

Independent Variable	**Model 1** PM_ Demo Coefficient	**Model 1** PBPM_ Demo t-stat	**Model 2** PBPM_ Demo Coefficient	**Model 2** PBPM Demo t-stat
Intercept	1,140**	15.7	1,203	8.3
Intervention	-63	0.7	-3	0.0
Inter_surv	17	0.2	12	0.1
PBPM_Base	0.241**	8.1	0.107	2.0
High-cost&high risk	—	—	701	4.1
High-cost	—	—	134	0.8
High-risk	—	—	383	3.3
Male	—	—	118	1.6
Minority	—	—	-371	1.6
Age 65-69	—	—	-41	0.3
70-74	—	—	-116	0.8
75-79	—	—	-234	1.6
80-84	—	—	-83	0.6
85+	—	—	-222	1.4
Medicaid	—	—	453	1.5
Urban	—	—	181	1.5
SNFB	—	—	4	0.0
R^2	.041	—	.066	—
N	1,567	—	1,567	—

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. PBPM = per beneficiary per month; Inter_surv = 1 if beneficiary had a Health Buddy® device and completed at least one survey; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

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

HBC = Health Buddy® Consortium.

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

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

High-Cost & High Risk: PBPM > $16,000 and HCC > 2.0 in base year.

High-Cost: PBPM > $16,000 and HCC< 2.0.

High-Risk: PBPM < $16,000 and HCC > 2.0.

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

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1 (4/20/10); final/cost5 (5/18/10).

7.8.2 Refresh Population

In the first results column of the refresh population in ***Table 7-16,*** controlling only for each beneficiary's base period PBPM cost, the Intervention coefficient of 12 is insignificant, implying no statistical cost trend differences between intervention beneficiaries not using the Health Buddy® device and the comparison group, all else being equal. The base period PBPM cost coefficient (0.261; $p < .01$), when combined with the intercept coefficient, again implies substantial RtoM of costs in the refresh sample (= 0.261 - 1 = -0.739, the RtoM effect). The Inter_surv coefficient is -229, however, which is statistically significant at the 95% confidence level. This implies that for the refresh population's beneficiaries, having used the Health Buddy® device appears to have lowered the rate of Medicare cost increases during the demonstration period.

The second regression model controls for cost-risk group and other patient characteristics determined during the base period. The Intervention coefficient remains insignificant and the Inter_surv coefficient (-224) remains statistically significant. Thus, the cost-saving effect of using the Health Buddy® device does not appear to be an artifact of any differences in measurable beneficiary characteristics between beneficiaries with the appliance and the entire comparison group.

Two of the three cost-risk refresh groups continue to show higher cost increases than the neither group. The PBPM_base coefficient declines somewhat, implying more RtoM within each of the cost-risk groups. Male beneficiaries appear to have slower cost growth. Disabled eligibility status had to be dropped from the model because of its high correlation with the age<65 reference group. No age effects were found relative to the under-65 group.

Table 7-16
HBC demonstration regression results: Intervention gross savings controlling for base period PBPM and beneficiary characteristics: Refresh population

Independent variable	**Model 1** PBPM_ Demo Coefficient	**Model 1** PBPM_ Demo t-stat	**Model 2** PBPM_ Demo Coefficient	**Model 2** PBPM_ Demo t
Intercept	950**	14.1	1,003**	6.9
Intervention	12	0.2	-7	0.1
Inter_surv	-229*	2.1	-224*	2.0
PBPM_Base	0.261**	10.6	0.197**	5.0
High-cost&high risk	—	—	559	3.5
High-cost	—	—	-102	0.7
High-risk	—	—	310*	2.3
Male	—	—	-140	1.9
Minority	—	—	-142	0.7
Age 65-69	—	—	-62	0.4
70-74	—	—	20	0.1
75-79	—	—	4	0.0
80-84	—	—	-7	0.1
85+	—	—	15	0.1
Medicaid	—	—	204	0.7
Urban	—	—	-15	0.1
SNFB	—	—	-4	0.0
R^2	.066	—	.067	—
N	2,078	—	2,078	—

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. PBPM = per beneficiary per month; Inter_surv = 1 if beneficiary had a Health Buddy® device and completed at least one survey; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

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

HBC = Health Buddy® Consortium.

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

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

High-Cost & High-Risk: PBPM > $16,000 and HCC > 2.0 in base year.

High-Cost: PBPM > $16,000 and HCC< 2.0.

High-Risk: PBPM < $16,000 and HCC > 2.0.

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

SOURCE: Medicare 2004-2008 Part A & B claims; Cost4b1(4/20/10); final/cost5 (5/18/10).

7.9 Comparison of Cost Savings: RTI versus ARC

The cost analyses conducted by RTI differ from those conducted by Actuarial Research Corporation (ARC) for financial reconciliation. ARC determined savings based on the demonstration's terms and conditions negotiated between CMS and the HBC. ARC first determined average per beneficiary per month (PBPM) costs of intervention and comparison group beneficiaries by summing Medicare outlays for all eligible claims in the numerator and dividing by all eligible months in each group in the denominator. ARC then performed the same calculation for the year preceding the start of the intervention. In summing beneficiary outlays, if a beneficiary's total costs in the intervention's first year was in the top 1% of costs, his/her costs were capped at the 1% threshold. To adjust for any random discrepancy in base year costs, ARC then multiplied the comparison group's PBPM cost during the 36-month intervention period by the ratio of intervention-to-comparison PBPM costs in the base year. A similar top 1% threshold cap was used. Thus, if intervention PBPM costs were greater than comparison costs in the base year, comparison PBPM costs were increased, proportionally, in the demonstration period, thereby factoring out of the final reconciliation any base year cost differences, on average.

ARC's sole task was to determine, using eligible claims, whether the HBC met CMS's contract terms and conditions. This is a straightforward actuarial exercise answering the question: What financial success did the HBC have during the demonstration period? The task assigned RTI was not only to determine whether savings were achieved but how confident could CMS be in the HBC replicating its performance in the future in other venues. This is a critical question on a technical level given the relatively small beneficiary samples involved in the HBC demonstration program. It also is a critical policy question given the potential federal financial risk involved in expanding the HBC approach to a larger, possibly national, program.

Generalizing results to likely future performance necessarily requires testing the hypothesis that any savings in a sample of beneficiaries during a particular time period could have been caused by chance with no long-run cost-saving implications. ARC's actuarial method precludes statistical testing. To answer the key financial questions, RTI conducted a range of analyses 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. Any imbalances in beneficiary characteristics that might produce inter-temporal differences in medical utilization or costs are factored out using first-differencing at the beneficiary level of analysis.

- Second, by calculating changes in PBPM costs at the beneficiary level using "paired" base-demonstration period PBPM costs, we can produce statistical *t*-tests and multivariate regression estimates that isolate the differences in spending growth rates between intervention and comparison groups.

RTI's estimation of savings also differs from ARC's in that

- 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*.

RTI did not trim outliers to avoid biasing the analysis against an intervention that was particularly successful in reducing costs of high-cost beneficiaries. Instead of capping top 1% outliers, RTI weighted PBPM mean costs and standard errors by each beneficiary's eligible fraction of days, or exposure to the intervention. This was done separately for the intervention and comparison groups. For the 36-month demonstration period, RTI's weighting method is equivalent to ARC's ratio-of-sums approach. The difference comes in implicitly weighting each beneficiary's base year costs by how long they were exposed to the actual intervention. Consequently, early demonstration dropouts (usually due to death) will have their base period PBPM costs down-weighted by RTI compared with ARC's base period costs. As early intervention dropouts are likely sicker and more costly in the previous base period, our mean base year costs are 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. Because RTI's methodology is applied consistently to both intervention and comparison groups, the net impact on relative differences in base period costs between RTI and ARC is ambiguous, as shown below.

ARC's untrimmed PBPM costs in the demonstration period are essentially identical to RTI's untrimmed estimates: Original intervention sample ($1,399 ARC; $1,398 RTI); comparison sample ($1,439 ARC; $1,436 RTI). ARC does not report actual trimmed base period costs by study group, only the ratio. ARC estimated that base period HBC PBPM costs were 13% higher than in the comparison group in the original sample which raised the comparison group's demonstration PBPM cost from $1,396 to $1,577 ($206 higher than the $1,371 intervention PBPM cost). RTI's untrimmed difference in base period costs was only 8.1%, resulting in an adjusted comparison group mean of $1,552, or $154 lower than RTI's $1,398 intervention HBC mean.

Investigating further the impact of the difference in RTI and ARC methods, we conducted a simulation of what ARC's cost savings would have been if they had not trimmed costs and/or not given intervention drop-out beneficiary's full weight in the base year. Table 7-17 shows the impact of ARC's data trimming and weighting approach on its final reconciliation calculations. Before any adjustments, ARC's statistics (ARC, 2010, Table 3) implied $893,051 in gross savings for the original sample and $1,209,029 for the refresh sample. ARC estimated gross savings of $4.572 million with trimming and applying its base year ratio adjustment. If RTI's base year ratio (1.081) of mean costs is used instead of ARC's, savings in the original sample decline by one-third to $3,057,418. Further eliminating trimming of high cost outliers raises gross savings to $3,479,726, as trimming worked against HBC in the original sample. RTI's refresh base year adjustment (1.017) actually slightly increases gross savings relative to ARC's 1.015 adjustment but is more than offset by the effect of first not trimming costs (compare $1,688,363 with ARC's $1,961,909 in refresh savings).

Table 7-17
Simulated ARC Gross Savings Using RTI Methods

Gross Savings Adjustments	Original Sample	Refresh Sample
ARC no trims or base adjustment	$893,051	$1,209,029
ARC trim & base adjustment	$4,572,116	$1,961,909
ARC trim w RTI base adjustment	$3,057,418	$2,021,798
ARC no trim w/ RTI base adjustment	$3,479,726	$1,688,363

ARC determined "excess" savings as the difference between gross savings, net of accrued fees, minus the 5% minimum savings required in the original sample (2.5% in the refresh sample). Net savings in excess of required savings were $1.537 million in the original sample (ARC, 2010, Table 2), allowing the HBC to retain all accrued fees. Excess savings were a negative $217,407 in the refresh sample. However, the HBC was allowed to keep all their fees for the refresh population because excess savings in the original sample were more than adequate to cover the negative excess in the refresh sample (see ARC, 2010, p. 8).

Applying RTI's no-trim and alternative weighting of base period costs show substantially lower gross (and excess) savings: Original: $3,48 million ($469,930); Refresh: $1,688,363 (-$787,160). RTI's excess savings, while still positive for the original sample, are 70% less than ARC's because of the smaller adjustment for base year cost differences. RTI's negative excess savings for the refresh group are 3.6-times greater than ARC's negative estimate of excess savings. Consequently, using RTI's methodology, excess savings in the original sample would not have covered all of the negative excess savings in the refresh sample. Instead of keeping all $2.77 million in fees, using RTI's method, the HBC would have retained $967,000 ($1.284 million + (-$.787 million + $.479 million).

Thus, while the HBC did achieve actuarial savings, the level of savings is sensitive to the decision to trim high-cost outliers and to giving early drop-out beneficiaries full weight in calculating base year costs. RTI's adjustments to ARC's approach produce lower gross savings for both original and refresh groups, albeit still positive. Lower RTI savings, determined at the beneficiary level, failed to meet the high level needed to achieve statistical significance, possibly because the HBC had the smallest sample sizes of any of the demonstration programs.

7.10 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. Approximately 1,000 beneficiaries in each of the original and refresh intervention and comparison groups limited our power to detect significant savings. Gross savings had to be 12.1% in the original intervention population and 14.3% in the refresh intervention population to be considered significant at the 95% confidence level.

No statistically significant savings were found for the intervention in the original population. Costs rose -$117 slower in the original intervention group (8.1% of comparison

costs), but savings needed to exceed 12.1% to be considered statistically significant. Medicare's return on investment was 2.01 only if we assume that the HBC's gross savings were statistically significant.

The HBC program, overall, did not perform better with its refresh population as gross savings averaged only -$63 (6.0% of comparison monthly costs). Medicare's return on investment was 1.12 if these savings were considered to be significant. However, beneficiaries who had a Health Buddy® device and completed at least one survey were found to have achieved significantly lower cost increases relative to the comparison group (and the rest of the intervention group). However, it could be that the multivariate models may not adequately control for the factors associated with agreeing to participate and use a Health Buddy® device so we may not have completely factored out selection bias effects.

A few material imbalances between intervention and comparison groups were found across many cost, severity, and other patient characteristics in the base period. However, controlling for imbalances had little effect on our overall final conclusion of no statistically significant savings. Using the Health Buddy® device, however, in the refresh population is suggestive of true cost savings.

The HBC CMHCB demonstration involved a select group of high cost, severely ill beneficiaries. As a result, the comparison group exhibited both rapidly rising costs during the intervention period ($470 in the original and $165 in the refresh groups) as well as extreme RtoM effects. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by over $1,000. Over the same time period, beneficiaries with costs over $3,000 saw their average costs decline by $1,500-$2,500. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings.

This page intentionally left blank.

CHAPTER 8
KEY FINDINGS FROM THE HEALTH BUDDY® CONSORTIUM'S 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 Health Buddy® Consortium's (HBC) Care Management for High Cost Beneficiaries (CMHCB) demonstration program. Our evaluation focuses upon three broad domains of inquiry:

- **Implementation.** To what extent was the HBC able to implement its program?
- **Reach.** How well did the HBC engage its intended audience?
- **Effectiveness.** To what degree was the HBC 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 38 months of the HBC operations with its original population and 26 months with its refresh population. Our findings are based on the experience of approximately 3,600 ill Medicare beneficiaries split across 4 groups for analysis purposes (original and refresh intervention and comparison groups) limiting statistical power somewhat to detect differences. Eight key findings on participation, intensity of engagement in the HBC program, 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 fee-for-service (FFS) beneficiaries. The CMHCB demonstration program holds the HBC financially responsible for financial savings but does not hold the HBC financially responsible for quality of care improvements.

Key Finding #1: The HBC program was able to engage beneficiaries who were at higher risk of acute clinical deterioration as measured by the concurrent HCC score.

Of the HBC original intervention beneficiaries, 45% verbally consented to participate in the CMHCB demonstration at some point during the intervention period; 40% of the refresh population agreed to participate. For the HBC program, we find that beneficiaries with medium and high concurrent HCC scores were more likely to be participants. Beneficiaries with higher prospective HCC scores and baseline Charlson comorbidity scores were less likely to be participants. This suggests that the HBC program was less able to engage the historically sicker Medicare beneficiaries but more able to engage those at higher risk of acute clinical deterioration as measured by the concurrent HCC score.

Key Finding #2: Thirty-six percent of the intervention population agreed to use the Health Buddy® device.

A cornerstone of the HBC's program was the Health Buddy® device and interactions with care managers to address gaps in knowledge or self-management of their chronic diseases. Of the roughly 1,800 intervention beneficiaries, 668 beneficiaries (36%) agreed to participate in the program and used the device to complete at least one survey. Among the beneficiaries that did agree to participate in the HBC program, use of the Health Buddy® device was high (88%). Under an intent-to-treat model, active engagement of less than one-half of the total number of intervention beneficiaries requires that the HBC program has a large intervention effect on the beneficiaries with whom the HBC program staff members are actively engaging to achieve the desired outcomes.

Key Finding #3: The HBC program did not substantially affect beneficiary reported experience with care, level of physical activity, and self-reported physical health. Among the 19 outcomes covered by the survey, the HBC intervention resulted in a higher frequency of medication compliance for beneficiaries in the intervention group relative to the comparison group.

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 HBC CMHCB intervention: helpfulness of discussions with their health care team and quality of communication with their health care team. In addition, the survey instrument collected information about beneficiary self-care frequency and self-efficacy related to medications, diet, and exercise and Clinician and Group Adult Primary Care Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

The HBC demonstration program employs strategies to improve quality of care for high cost Medicare beneficiaries while reducing costs by empowering Medicare beneficiaries to better manage their care and mitigate acute flare-ups in the chronic conditions. 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, the HBC intervention resulted in a higher frequency of medication compliance for beneficiaries in the intervention group relative to the comparison group.

Key Finding #4: Rates of compliance with 3-of-4 quality-of-care process measures were high at baseline providing limited opportunity for improvement. The general trends during the demonstration were stable or decreasing rates of compliance in both the intervention and comparison groups.

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. Within the original and refresh intervention and comparison populations, we generally observe stable or negative trends in the rates. The original intervention group's rates tended to fall more than its comparison group's rates in 6-of-8 measurements; while the refresh intervention group's rates tended to fall less than its comparison group's rates in 3-of-4 measurements. The difference-in-differences (D-in-D) rates per 100 beneficiaries ranged from 1 to -7 per 100 beneficiaries for the original population and 8 to -7 per 100 beneficiaries for the refresh population. Of these differences, there is one that is statistically significant. The rate of receipt of the influenza vaccine among the refresh intervention beneficiaries declined by 1 percentage point while the rate of receipt among the refresh comparison beneficiaries increased 6 percentage points. Thus, the D-in-D change is -7 per 100 beneficiaries.

Key Finding #5: Rates of acute care utilization increased during the demonstration in the original and refresh intervention and comparison groups with one exception; all-cause hospitalizations declined within the refresh intervention group while the rate of all-cause hospitalizations increased within the comparison group. Although we observe no other statistically significant differential rates of growth in acute care utilization, we do observe a trend toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures. We do not observe differential use of the Medicare hospice benefit.

During the course of the HBC demonstration, we observed increasing rates of all-cause and ambulatory care sensitive conditions (ACSC) hospitalizations, ER visits, and 90-day readmissions in both the intervention and comparison groups and for both the original and refresh populations with one exception. All-cause hospitalizations declined within the refresh intervention group while the rate of all-cause hospitalizations increased within the comparison group for a rate of -154 hospitalizations per 1,000 beneficiaries D-in-D rate (p=0.02). This represents a 26% lower rate than what would have been expected. Although we observed no other statistically significant differential rates of growth in all-cause or ACSC hospitalizations or ER visits or 90-day readmissions, we observed a trend toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures with a number of the D-in-D rates appearing to be of clinical significance although not statistically significant. Further, we do observe wide confidence intervals for several of the readmission estimates due to small sample sizes. We did not find any statistically significant differences between the intervention and comparison beneficiaries in either the original or refresh populations in their take-up rate of the Medicare hospice benefit or in mean and median number of days in hospice.

Key Finding #6: We observe a lower rate of mortality among intervention beneficiaries that used the Health Buddy® device.

We do observe a statistically significant lower rate of mortality in the original population's intervention group. Over the 38-month demonstration period for the original population, 35% of the original intervention group beneficiaries died while 40% of the comparison group beneficiaries died; a 5 percentage point lower rate of mortality in the intervention group (p=0.04). Over the 26-month demonstration period for the refresh population,

21% of the refresh intervention group beneficiaries died and 23% of the comparison group beneficiaries died; a 2 percentage point lower rate of mortality in the intervention group (p=0.33).

We estimated multivariate models of survival, whereby we controlled for potential imbalances in baseline beneficiary characteristics that may be related to mortality and not adequately accounted for in the development of a comparison group. When doing so, the observed survival benefit for the intervention group within the original population was no longer present. However, when we introduced into our model a variable that captures the impact of intervention beneficiaries using the Health Buddy® device, we observed an incremental increase in survival benefit among both the original and refresh populations' intervention beneficiaries who used the Health Buddy® device. Because we did not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the survival benefit. However, given this important finding, additional study is warranted.

Key Finding #7: Medicare cost growth was slower in the intervention group in both the original and refresh populations, but neither trend was statistically significant.

No statistically significant savings were found for the intervention group in the original population. Costs rose $117 slower in the original intervention group (8.1% of the comparison group's costs), but savings needed to exceed 12.1% to be considered statistically significant. The HBC's trend in gross savings averaged -$73 in the refresh intervention group (6.0% of the comparison group's monthly costs), but savings needed to exceed 14.3% to be statistically significant. Insignificance may have been due to small numbers of intervention beneficiaries: 763 (original population); 1,028 (refresh population). A few material imbalances were found in cost, severity, and other patient characteristics between the original and refresh intervention and comparison groups in the base period. Still, controlling for imbalances had little effect on our overall final conclusion of no detectable statistically significant savings.

Actuarial Research Corporation (ARC), under separate contract to CMS, conducted an actuarial reconciliation of financial performance of the HBC program and also found gross savings for the intervention. ARC-determined savings differed from savings reported by RTI in three ways. First, ARC capped high-cost beneficiaries at the top 1% threshold. RTI did not cap outliers because we did not want to inadvertently bias results against the intervention if it was particularly successful in reducing costs of the very high-cost beneficiaries. Second, ARC adjusted for base period differences in intervention-comparison group costs without taking beneficiary eligibility during the demonstration period into account. RTI down-weighted base period costs for beneficiaries with shorter demonstration period exposure. Third, ARC made no independent assessment of the statistical reliability of their cost estimates. RTI conducted all analyses at the individual beneficiary level to be able to test the reliability of savings.

Simulation analyses showed that ARC's level of savings was sensitive to its outlier trimming and its estimates of base year average costs. Without trimming and using RTI's method for calculating base year costs, ARC's gross savings would have been $3.5 million in the original sample instead of $4.6 million and $1.7 million in the refresh sample instead of $2.0 million. Using ARC's gross savings based on RTI methods would have resulted in the HBC

retaining $967,000 in fees instead of $2.8 million. That savings are still positive using a modified ARC approach and RTI's statistical approach suggest than the HBC's intervention is an approach worthy of continued study.

Key Finding #8: Beneficiaries in the refresh population using the Health Buddy® device exhibited a slower rate of cost growth.

Although the HBC program performance summarized in other findings is based on the entire intervention population, we were interested in whether beneficiaries using the Health Buddy® device had a slower rate of cost growth. Controlling for age, gender, minority status, and other beneficiary characteristics, those using the Health Buddy® device in the refresh population exhibited slower cost growth of over $200, significant at the 5% level of confidence. No difference was found in the original population. Because we could not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the cost savings and not the Health Buddy® device itself. Nevertheless, the lower rate of growth in Medicare costs and the lower observed rate of mortality supports continued study of the cost effectiveness of using monitoring devices in the home.

8.2 Conclusion

Based on extensive quantitative analysis of performance using statistical tests at standard 5% confidence levels, we did not detect improvement in key processes of care, beneficiary self-reported experience with care, self-management, and functional status, or use of the Medicare hospice benefit. The HBC program was successful in reducing the rate of all-cause hospitalizations within its refresh intervention group with a trend (not statistically significant) toward lower rates of growth within the original and refresh intervention populations for two-thirds of the acute care utilization measures. We also observed an incremental increase in survival benefit among the original and refresh populations' intervention beneficiaries who used the Health Buddy® device relative to the comparison group (and the rest of the intervention group). Although PBPM costs rose slower in the original and refresh intervention groups relative to the comparison groups, statistically significant savings were not achieved in the *overall* intervention groups. Nevertheless, we observed significantly lower cost increases among refresh intervention beneficiaries who used the Health Buddy® device.

What might explain the lack of *overall* program effectiveness? One factor may be relatively small sample sizes and lack of statistical power. Only 763 and 1,028 intervention beneficiaries were available for analysis in the original and refresh groups and comparable numbers in the corresponding comparison groups. In addition, wide variation in beneficiary costs over time made precise estimates of program success difficult with such small samples. Responding to the HBC's request, CMS selected a very costly, complex set of Medicare beneficiaries for their intervention and comparison groups. Mean per beneficiary per month base year claims costs (weighted by fraction of time eligible for the intervention) were approximately $1,000 in both groups, a figure considerably higher than in the general Medicare population. Further, we observed extreme regression-to-the-mean (RtoM) behavior among the HBC's selected beneficiaries. Beneficiaries incurring less than $500 monthly in Medicare costs saw their average PBPM costs rise by over $1,000. Over the same time period, beneficiaries with

monthly costs over $3,000 saw their average costs decline by $1,500-$2,500. The large churning of beneficiaries from lower (higher) to higher (lower) cost groups over time adds considerable statistical noise to the test of savings.

A second factor may be the HBC's beneficiary recruitment strategy. Given the HBC program's monthly management fee (roughly $120 per month) and the population-based design of this demonstration, engagement of less than 50% of the intervention population required the HBC program to have been extremely successful with the participating beneficiaries.

And, a third factor 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 HBC program's fee. A cornerstone of the HBC's program was health coaching interactions with care manager nurses in response to alerts generated by the Health Buddy® device. Nearly all participating beneficiaries using the Health Buddy® device received at least one call from a care manager and nearly 60% received more than 20 calls. This is a relatively high contact rate compared to other care management programs that we have evaluated. However, the Health Buddy® nurse care managers often were not in direct proximity to their beneficiaries' primary care physicians, thereby potentially affecting their interactions with the beneficiaries' primary providers, changing medical care plans, or mitigating deterioration in health status. The care manager served primarily as an adjunct to the patients' primary physicians. Interviewed physicians felt that care management would be more effective and efficient if care managers were colocated with primary care physicians. Further, not all intervention beneficiaries had primary care physicians in the two study sites, therefore the care managers had to interact with community-based providers with whom they had little or no prior relationship. During our site visits, the care managers cited several challenges working with these physicians, in particular, because of communication barriers. Lastly, by complementing, not substituting, for the primary care physician, the nurse care managers were not directly determining whether a patient was admitted to a hospital or what service intensity the beneficiaries would receive during the demonstration period.

Yet, we do observe an incremental increase in survival benefit and lower cost increases among intervention beneficiaries who used the Health Buddy® device. As noted before, because we could not directly compare Health Buddy® device users with a matched comparison group instead of the entire comparison group, it is possible that unmeasured characteristics explain the survival benefit and cost savings and not the Health Buddy® device itself. These two substantive findings require further evaluation by analysis of the HBC Phase II demonstration experience. It will be important to explore with the HBC what beneficiary characteristics they believe lead them to agree to use the Health Buddy® device. With this information, we may be able to develop an alternative comparison group that more closely aligns with the subset of beneficiaries that use the Health Buddy® device.

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.