January 2010

Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Care Level Management (CLM)

Final Report

Prepared for

David Bott, Ph.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.
Carol Urato, M.A.
Terry Eng, M.S.
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: CARE LEVEL MANAGEMENT (CLM)

by Nancy McCall, Sc.D.
Jerry Cromwell, Ph.D.
Carol Urato, M.A.
Terry Eng, M.S.

Scientific Reviewer: Janet Mitchell, Ph.D.

Federal Project Officer: David Bott, Ph.D.

RTI International*

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

January 2010

This project was funded by the Centers for Medicare & Medicaid Services under contract no 500-00-0024 TO#25. The statements contained in this report are solely those of the authors and do not necessarily reflect the views or policies of the Centers for Medicare & Medicaid Services. RTI assumes responsibility for the accuracy and completeness of the information contained in this report.

*RTI International is a trade name of Research Triangle Institute.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

The purpose of this report is to present the findings from RTI International's evaluation of Care Level Management's (CLM's) Medicare Care Management for High Cost Beneficiaries (CMHCB) demonstration program. 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 CLM 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), CLM was held at risk for its monthly management fees based on the performance of the full population of eligible beneficiaries assigned to its intervention group and as compared with all eligible beneficiaries assigned to its comparison group. Beneficiary participation in the CMHCB demonstration was voluntary and did not change the scope, duration, or amount of Medicare FFS benefits received. All Medicare FFS benefits continued to be covered, administered, and paid for by the traditional Medicare FFS program. Beneficiaries did not pay any charge to receive CMHCB program services.

Our evaluation focuses upon three broad domains of inquiry:

- Implementation. To what extent was CLM able to implement its program?
- Reach. How well did CLM engage its intended audiences?
- Effectiveness. To what degree did CLM 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

CLM launched its CMHCB demonstration program October 1, 2005. During the first year of operations, CLM requested that a subset of its existing original population be removed (carved-out). CMS also offered CLM a refresh population at twelve months post launch that would partially offset attrition due primarily to death. CLM worked with CMS to redefine the

criteria for identifying the refresh population. The CLM refresh population went live on September 1, 2006.

Of CLM's original intervention group beneficiaries, 63% verbally consented to participate in the CMHCB demonstration at some point during the intervention period, 23% refused to participate, and 14% were not contacted or were unable to be located. CLM revised their criteria for selecting beneficiaries for their refresh population in order to have a population that they believed would benefit more from the services that they offered and would, therefore, be more likely to want to participate, and that could be contacted (that is, they had a phone number on file with the Social Security Administration [SSA]). With the selection criteria changes, there was modest improvement in their participation rate. Overall, 66% of the refresh intervention beneficiaries consented to participate at some point during the 16-month period. The percent that refused to participate went up slightly (24%), but the percent that were not contacted or were unable to be contacted decreased to 10%.

During the second year of operations, CLM received a termination letter with a 90-day notice from CMS. CMS was concerned that the quarterly monitoring reports showed no cost savings, and CLM did not develop a viable plan to change its financial course. The CLM program ended February 29, 2008, or 29 months, after initiation of the original population and 18 months after the start of the refresh population.

CLM negotiated a per beneficiary per month payment of $295 for the duration of the demonstration for both the original and refresh populations. The net savings requirements for those CMOs that complete a 36-month demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. However, because CLM's demonstration was terminated on February 29, 2008, the net savings requirement was pro-rated based on the number of months after the effective start date for each cohort for which the demonstration was terminated. In CLM's case, the termination date was 29 months after the start date of the original cohort and 18 months after the start date of the refresh cohort. Therefore, at the time of its termination, CLM was contractually obligated to achieve a 4% savings in Medicare payments among the original intervention group and 2% for the refresh intervention group (regardless of participation in the CLM program) compared with the comparison group and to cover the CLM program fees collected. In addition, CLM had the opportunity to share a portion of any savings that were achieved beyond its obligations. Payments associated with the Medicare Part D benefit were not to be included in these calculations.

E.2 Overview of the CLM CMHCB Demonstration Program

The overarching goal of the CLM program was to provide home-based care and 24/7 access to a personal visiting physician (PVP) to patients with multiple chronic conditions, who were at high risk for multiple hospital admissions. As originally envisioned, PVPs would provide acute episodic care, such as a home hospitalization consisting of daily visits to the patient's home for 2 or more days to administer treatments such as IV antibiotics. PVPs were also to provide a "bridge to home," by helping patients who were being cared for in more intensive venues to transition to lower levels of care. PVPs were not intended to replace community-based primary care physicians (PCPs); rather, PVPs served as adjuncts to the patients' PCPs and provided care

with a focus on chronic issues and established care goals for each chronic condition that a patient had, ensuring that the patient agreed with the care goals.

Nurse practitioners were available to support the PVPs by conducting follow-up visits and responding to urgent calls. Each nurse practitioner supported approximately five PVPs. In general, nurse practitioners conducted visits with lower acuity patients, who were identified in collaboration with the PVPs. Nurse case managers provided support to doctors in their efforts to care for patients and contact patients directly by telephone to implement physician care plans and address issues that arose during the time between PVP visits. Nurse care managers served as patient advocates and coordinated care with patients, family members, community-based primary care physicians, specialists, and home health nurses. Each nurse care manager was responsible for providing support to a single PVP and his or her respective patients.

CLM's demonstration program evolved during the course of the demonstration to respond to some of the challenges that it faced after implementation. CLM explored ways to modify its program to provide cost-effective care in rural areas where patients were geographically dispersed and physicians would have to drive long distances to conduct eight home visits per day, the desired number from an operations perspective. In addition, CLM found it difficult to recruit physicians who were willing to live in these areas.

Institutionalized beneficiaries were not excluded from CLM's original intervention population. However, CLM found the process to engage administrators of nursing homes and gain support for allowing the CLM PVPs to provide care for eligible residents required significant CLM staff time so much so that CLM opted to exclude this population from the refresh that it had negotiated with CMS.

Using information gleaned from the first year of operations, CLM made several enhancements to its clinical model. CLM replaced its panel management scheme with a population-based management approach that involved reorganizing patient care teams to include more nursing support. In particular, physicians started covering a caseload of patients with the support of 1.5 to 2 nurses, rather than just 1 nurse. This new arrangement was intended to connect stable patients who experienced a problem with clinical assistance more quickly because they would have an established relationship with a member of the care team that would provide support during an acute period.

CLM also reported that it transitioned to a three-dimensional risk stratification system in an attempt to better determine the frequency of physician visits and care management calls needed to meet the needs and preferences of participants. CLM felt that this approach addressed their belief that some people did best with more physician visits, while others preferred telephonic contact with a nurse care manager. Also during the latter part of the demonstration, CLM implemented a telemonitoring pilot. CLM believed that by routinely monitoring objective data it could be alerted when a patient was beginning to have problems, that is, when a PVP visit could effectively reduce the need for an emergency room visit or hospitalization. CLM reported that often by the time a patient called CLM about a problem, he/she was so sick that he/she had to go to the hospital.

E.3 Key Findings

In this section, we present key findings based upon the 29 months of CLM operations with its original population and 18 months with its refresh population. Our findings are based on the experience of approximately 34,000 ill Medicare beneficiaries assigned to an intervention or a comparison group. Six key findings on participation, intensity of engagement in the CLM 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.

Key Finding #1: Several vulnerable subpopulations of Medicare FFS beneficiaries were less likely to agree to participate in the CLM demonstration program.

Of all CLM intervention beneficiaries, 65% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. We found that Medicaid enrollees and institutionalized beneficiaries were less likely to be participants; both groups are costly and high users of acute care services. In general, participants tended to be healthier than nonparticipants using baseline characteristics including the prospective HCC score. However, beneficiaries with higher *concurrent* HCC scores based on the first 6 months of the demonstration were more likely to participate than healthier beneficiaries. This suggests that CLM made some inroads into engaging those with acute clinical deterioration. Further, as CLM's program matured, they appeared to be more successful engaging sicker and more costly beneficiaries based on baseline health status; however, those with Medicare/Medicaid dual enrollment and the institutionalized were still less likely to become participants. These findings suggest alternative recruiting and outreach strategies are needed to reach dual Medicare/Medicaid enrollees and beneficiaries who are institutionalized.

Key Finding #2: CLM's physician home intervention (PHI) was limited to less than one-third of their participating beneficiaries.

A cornerstone of CLM's program was physician home intervention (PHI) visits including home hospitalization services. We found some evidence that CLM made a focused effort to contact beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program. Yet, only 30% of their fully participating beneficiaries received a PHI visit during a 12-month period. When we add in other types of visits – including routine follow-up care, we do observe a higher level of physician interaction. However, 25% of beneficiaries received no physician visits. Telephone contact was the most dominant "frequent" form of contact.

Key Finding #3: CLM improved some aspects of beneficiary reported experience with care, level of physical activity, and self-reported physical health.

The beneficiary survey was designed to obtain assessments directly from beneficiaries about key outcomes of beneficiary ***experience of care, self-management***, and ***physical and mental function.*** We asked beneficiaries about the extent to which their health care providers helped them to cope with their chronic condition. We supplemented this item with questions related to two key components of the CLM 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.

Among the 19 outcomes covered by the survey, four statistically significant positive intervention effects were found—discussing treatment choices, communication with their health care providers broadly defined, 30 minutes of continuous physical activity, and, most notably, physical health. However, there was no improvement in overall beneficiary satisfaction that their health care team helped them deal with their chronic condition in spite of the positive intervention effects of two related experience with care measures—helpfulness and the quality of discussions with their health care team and improvement in self-reported physical health.

A positive intervention effect of 30 minutes of continuous physical activity did not translate into greater *confidence* on the part of the CLM population in exercising 2 to 3 times a week. The mean number of days of 30 minutes of exercise was 2.8 in the comparison group and 3.2 in the intervention group.

We used the RAND-12 scoring algorithm 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 CLM intervention population had a mean PHC score of 30 versus 28 in its comparison group, a statistically significant difference. Further, a 2 percentage point difference is generally viewed as a clinically meaningful difference. There was no difference in mental health functioning as a result of the CLM intervention. Nor was there was any difference in abilities to do activities of daily living.

Key Finding #4: CLM had a positive intervention effect on one of five quality of care process measures.

We have defined quality improvement for this evaluation as an increase in the rate of receipt of claims-derived, evidence-based process-of-care measures (e.g., serum cholesterol testing) and improvement in health outcomes as a reduction in the rate of hospitalizations, readmissions, and emergency room (ER) visits, and a reduction in mortality rates. We find no evidence of systematic improvement in quality of care in the CLM CMHCB demonstration program. Out of five measures, there was only one observed increase in rate of receipt of evidence-based care (influenza vaccination). We observe this increase in both the original and refresh population and during the last 12 months of CLM's operations.

Over the course of the demonstration, CLM had expected to increase rate of adherence to evidence-based care. However, during the last year of their demonstration program, we observe lower rates of adherence to the selected measures among their intervention beneficiaries than we do among the comparison group beneficiaries for all measures with the exception of influenza vaccination. We also observe between roughly one-third to one-half of intervention beneficiaries in both the original and refresh populations were not compliant during the last year of the CMHCB demonstration despite focused efforts by CLM to encourage beneficiaries to become

compliant with evidence-based care. As noted above, over 80% of intervention beneficiaries with COPD were not compliant with annual oxygen saturation assessment. These findings suggest that improving or sustaining adherence to guideline concordant care in a cohort of ill Medicare FFS beneficiaries was more challenging than originally envisioned.

Key Finding #5: During the last 12 months of the demonstration program, CLM had some success in reducing acute care utilization in both the original and refresh populations. CLM had no success reducing mortality in either the original or refresh population.

During the first half of program operations, CLM was not successful in reducing acute hospitalizations, ER visits, or 90-day readmissions in their original intervention population. But during the last 12 months of the demonstration, CLM was modestly successful in reducing the percent and rate of admissions for ten ambulatory care sensitive conditions (ACSCs). Most notably, CLM achieved a sizeable reduction in the all-cause readmission rate of -225 per 1,000 beneficiaries. During that same time period, CLM also had some modest success in reducing the acute care utilization in its refresh population. Rates of all-cause and ACSC hospitalizations declined, as did the percent with a readmission for all causes or an ACSC.

CLM was not successful reducing rates of ER visits in either time period for the original population or in the last 12 months of the demonstration for the refresh population. CLM had no success reducing mortality in either the original or refresh population.

Over the course of the first year of operations, CLM reported that they modified their program in an effort to identify, in real time, participants whom they believed would most benefit from their interventions by changing how they stratified beneficiaries into levels of visit urgency. CLM also reported that they reorganized their patient care teams to include more nursing support. CLM believed that this arrangement would allow patients to bond with the nurse care manager over time, whereas CLM had observed that the clinical specialists were not able to forge a sufficient bond as evidenced by the fact that some of their participants were going to the hospital rather than calling the clinical specialists when problems arose. CLM also reported that they felt the members of the refresh population had an illness profile that made them more appropriate to the CLM program. The data on health outcomes suggests for the original population that the program modifications had the desired effect of reducing some acute care utilization, but not ER visits; while the data on health outcomes for the refresh population shows that CLM was more successful reducing acute care utilization at an earlier stage in their demonstration period than for the original population.

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

No statistically significant savings were found for the intervention in either the original or refresh populations. Costs fell $41 slower (not faster as required) in the original intervention group (1.6% of comparison costs) but savings needed to exceed $133 (5%) to be considered statistically significant. Instead of offsetting its $295 monthly care management fee, Care Level Management may have increased Medicare's costs to $336 per beneficiary per month.

CLM performed slightly better with its refresh population as intervention costs increased $29 less than in the comparison group. This difference, however, was highly insignificant, as savings needed to be $143 to be considered statistically significant. Among refresh participants, alone, monthly Medicare costs did decline significantly (-$170) but were offset by large increases among non-participants who were one-third of the entire intervention population. It is not possible to rule out the selection of beneficiaries who are more responsive to the intervention in explaining success only with participants.

Multivariate regression was used to control for any imbalances in the intervention and comparison groups prior to the disease management intervention. No cost savings were found after adjusting for minor imbalances in the two groups.

Base year per beneficiary per month claims costs averaged $3,300 in both groups, a figure several times higher than in the general Medicare population. As a result, the comparison group exhibited extreme regression-to-the-mean effects[1] with costs declining $859, on average, during the demonstration period. At the same time average group costs were falling, initially lower cost beneficiaries saw their costs rising by several thousand dollars during the intervention period. The greater is the potential for regression-to-the-mean in either direction, the greater is the challenge facing care management groups in identifying the appropriate beneficiaries to target for intervention.

E.4 Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that CLM had limited success in improving key processes of care, beneficiary experience with care, self-management, or functional status, and reducing hospital admissions. CLM was most successful at reducing 90-day all-cause readmissions by -225 per 1,000 among its original beneficiaries. However, the overall set of modest improvements were achieved at substantial cost to the Medicare program in the form of monthly management fees ($58 million) with no demonstrable savings in program outlays on health services. Despite the limited gains, the lack of program savings to offset monthly management fees cannot justify the CLM model for chronically ill Medicare fee-for-service beneficiaries on cost effectiveness grounds.

What might explain the lack of success in CLM's demonstration program? One explanation may be the inability to accurately target beneficiaries at greatest risk of intensive, costly, service use (as distinct from the need for general care management). A cornerstone of CLM's program was physician home intervention (PHI) visits including home hospitalization services with a focus on ambulatory care sensitive conditions (ACSCs). Yet, only 30% of their fully participating beneficiaries for a 12-month period received a PHI visit. When we add in other types of visits—including routine follow-up care—we do observe a considerably higher level of physician interaction. However, 25% of beneficiaries received no physician visits and one-half of all beneficiaries received less than five visits during a 12-month period. Telephone contact was the most dominant "frequent" form of contact.

[1] Regression to the mean refers to high (low) initial costs gravitating to the mean cost over time.

In our multivariate regression modeling of likelihood of being in a high contact versus low contact group, we found some evidence that CLM made a focused effort to interact with beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program. Given CLM's high monthly management fee (almost $300 per month) and the population-based financial risk feature of this demonstration, the concentration of physician visits, in general, and PHI visits, in particular, suggests that CLM would have had to have been extremely successful in reducing costs associated with the beneficiaries they were targeting. Descriptive analyses showed that the rates of all-cause and ACSC hospitalizations during the demonstration were higher among beneficiaries who received PHI visits or had more than 20 contacts as compared to beneficiaries with no PHI visits or who received fewer than 20 contacts. The pattern was consistent across both the original and refresh populations.

Prior evaluations of Medicare care management programs that were primarily telephonic have not demonstrated savings sufficient to cover fees one-third the size of CLM's fee. CLM was successful in only modestly reducing hospitalizations during the last 12 months of demonstration operations, with no particularly greater success for ACSC hospitalizations. CLM was more successful at reducing readmissions but only among its original population. The lack of substantive improvements in acute care utilization broadly across their intervention population translated into limited financial savings. And, their targeting strategy was costly. Each contact cost was roughly $266, or over twice the national average payment for a face-to-face office visit with an established patient with the ***highest level of complexity*** under the Medicare Fee Schedule[2].

CLM's lack of success is not surprising in light of the extreme regression-to-the-mean (RtoM) behavior that we observed among their beneficiaries who had been selected based upon high prior costs or high prior rates of hospitalizations. Armed with data on beneficiary disease, utilization, and cost profiles in the base period, CLM focused first on those most likely to be major users of acute care services. Yet, many of these beneficiaries experienced declines in use and costs regardless of the intervention, as evidenced in the control group. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater the potential for RtoM, the greater the effort required to identify those lower cost, lower utilization, beneficiaries to avoid expensive hospitalizations in the future. Targeting the group who had a high number of hospitalizations in the base period focused extensive management resources on many "false positive" beneficiaries, who ultimately did not need nearly as many costly services as they did in the year prior to the demonstration.

The quixotic, immediate, nature of many elderly diseases calls for real time information on health status. Unable to predict future health status with any precision for those with initially stable, less costly, conditions, and lacking direct access to patients' medical records, the CLM physicians and nurses often began working with beneficiaries with incomplete information. Further, the CLM physicians and nurse care managers were not part of the beneficiaries' primary health care teams, further hindering their ability to directly interact with the beneficiaries' primary care providers and effectively help facilitate changes in medical care plans to mitigate

2 National non-facility price of $124.79 for HCPCS code 99215 for 2009.

deterioration in health status. It is not surprising that CLM was unable to successfully improve patient self-management.

Because targeting care management resources is so difficult with the elderly, and errors so costly, the way in which the clinical team communicates and interacts with them is extremely important. Yet, another possible reason why CLM was ineffectual has to do with the limitations of CLM's personal visit physician (PVP) and nurse care manager model. The PVP served only as an adjunct to the patients' primary care physicians with a stated goal of facilitating the relationship between the patient and his or her community-based provider with a focus on chronic issues. The PVP consulted with community-based providers if significant changes in treatment regimens were indicated and to ensure that they both were implementing a common care plan. Nurse care managers worked with the PVPs to telephonically help coordinate care services with patients, family members, community-based primary care physicians and specialists, and home health nurses. By complementing, not substituting, for the primary care physician, the PVPs and 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. Moreover, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments, and the most dominant form of contact was telephonic.

CHAPTER 1
INTRODUCTION TO THE MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION AND THE CARE LEVEL MANAGEMENT (CLM) 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 Care Level Management's Medicare 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:

- The Health Buddy Consortium (HBC), comprised of Health Hero Network, the American Medical Group Association, Bend Memorial Clinic, and Wenatchee Valley Medical Center
- Care Level Management (CLM)
- Massachusetts General Hospital and Massachusetts General Physicians Organization (MGH)
- Montefiore Medical Center (MMC)
- RMS Disease Management and its Key to Better Health program (KTBH)
- 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:

- Savings in the 0%-5% range will be paid 100% to CMS.
- Savings in the >5%-10% range will be paid 100% to CMO.
- Savings in the >10%-20% range will be shared equally between CMO (50%) and CMS (50%).
- 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:

- *Implementation.* To what extent were the CMOs able to implement their programs?
- *Reach.* How well did the CMOs engage their intended audiences?

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

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

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

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

RTI also conducted an assessment of beneficiary satisfaction with the CMHCB program and whether the program improved knowledge and self-management skills that led to behavioral change and improved health status among intervention beneficiaries. Program success for each of four beneficiary survey domains, satisfaction, care experience, self-management, and physical and mental health functioning, was evaluated by surveying intervention and comparison beneficiaries once at Month 15 of the intervention period. CLM'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 29 months of the CLM CMHCB program operations with its original population and 18 months with its refresh population. We start by reporting on the degree to which CLM 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 CLM and

participating beneficiaries from encounter data provided to RTI from CLM. We then report findings related to the effectiveness of CLM to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings.

1.2 CLM's CMHCB Program Design Features

1.2.1 CLM Organizational Characteristics

CLM was a privately owned, independent company that provided in-home physician care based on a mission of "Putting Patients First." CLM's corporate headquarters were located in Woodland Hills, California, with additional offices in northern California, Arizona, Texas, and Florida. CLM operated independently without formal partnerships with other health care service providers. The company retained the services of Milliman, a consulting and actuarial firm, to conduct analyses of Medicare claims and other market data to inform the choice of geographic locations for the CMHCB program and to conduct ongoing monitoring of intervention group characteristics, such as monthly Medicare costs. CLM had established partnerships with RAND, Ernst & Young, and Sullivan/Luallin to assist in evaluating the outcomes of CLM's CMHCB program.

The company was started by Dr. Henri Becker, a physician who started conducting home visits in 1995 to improve care for his patients who were hospitalized repeatedly as a result of complications associated with chronic illnesses, and Raouf Khalil, who ran a home health infusion company for 11 years and had a strong interest in patient satisfaction. The company founders met through a common patient who was particularly satisfied with the home-based care provided by Dr. Becker. CLM was established in 2000 to professionalize and institutionalize physician-based care in the home. The goal of the organization was to provide quality health care to patients who are at high risk for repeated hospitalizations due to medical and psychosocial issues associated with managing multiple chronic illnesses that were not being effectively addressed by the traditional health care system. CLM began providing care through its Personal Visiting Physician Delivery System (PVPDS) for its first contract in 2001.

CLM's management viewed its patient-centric model of care as the next major paradigm shift in health care, on par with the introduction of Medicare and Medicaid in 1963. Meeting patient needs was the focus of CLM's operations, illustrated by the fact that CLM patients were located at the top of the company's organizational chart. The company operated as a learning organization that continually reviewed program performance and made improvements to better serve its patients. Personal Visiting Physicians (PVPs) served as an adjunct to a patient's primary care physician, providing care in the home when the patient was unable to get to a physician's office. The CLM model allowed patients to see a physician in their homes, where they would be most comfortable, and provided patients with 24-hour access to a PVP via telephone and in-person, if needed. A staff of 70 physicians and 82 nurses were dedicated to meeting all the medical and psychosocial needs of patients aligned with patient preferences for care.

1.2.2 Market Characteristics

CLM's CMHCB program served Medicare Fee-for-Service (MFFS) beneficiaries in selected counties in California, Florida, and Texas. CLM launched its CMHCB program with

15,281 MFFS beneficiaries residing in California, Florida, and Texas in response to CMS's request that the company serve areas where it had an established infrastructure and experience serving the population. In addition, CLM was restricted from operating in Long Island, New York, because a different CMHCB program had been awarded in the area. CLM had also considered launching the program in Arizona; however, an existing CMS demonstration program conflicted with this choice of location. Despite the fact that there were existing CMS demonstrations in California and Texas, CLM was permitted to launch its CMHCB program in these areas. However, in order to achieve its goal of serving an intervention population of 15,000 beneficiaries, CLM had to expand its operations in northern California. CLM also expanded its Florida operations into the Orlando area for the CMHCB demonstration. As a result, participation in the CMHCB program increased the geographic footprint of CLM's operations to 25 times its size prior to the demonstration.

CLM believed that the CMHCB demonstration was an important opportunity to demonstrate the effectiveness of its Personal Visiting Physician Delivery System on a large scale, serving approximately 15,000 beneficiaries in three states. The program also provided CLM with a strategically important opportunity to work with CMS, which is financially responsible for a large proportion of the most vulnerable and frail individuals, the target population for the CLM model of care delivery.

Prior to 2004, CLM was serving approximately 1,800 patients via contracts with commercial insurance plans in four states (i.e., Arizona, California, Florida, and Texas), with an interest in expanding its operations into every state in the United States. The company's early experience in the commercial market demonstrated that contracts with individual health plans would provide access to a very small number of patients appropriate for the CLM care model. These small populations were geographically dispersed, requiring physicians to spend large amounts of time in the car between patient visits, thus limiting the financial viability of the home-based care model. Therefore, in March 2004, CLM began to focus all of its business development resources on building a relationship with CMS with the goal of accessing the large Medicare FFS market across the country. As a result, 85% to 90% of CLM's revenue during the demonstration period came from the CMHCB demonstration program, and the company expanded rapidly to serve the FFS beneficiaries participating in the program.

CLM was pleased to have the opportunity to adapt its intervention to the Medicare FFS population and implement the PVPDS in a large population of 15,000 beneficiaries to both demonstrate the effectiveness of the intervention and learn new ways to further improve its care model. CLM was optimistic about reaching the financial savings goals for the CMHCB program and improving care for beneficiaries served by the program based on the company's strong belief in the effectiveness of providing physician services in the home and the company's success providing such services for commercial clients.

As indicated in CLM's proposal for the CMHCB demonstration, the company's performance record included significant health care cost savings, improvements in patient functional status and quality of life, and high levels of patient satisfaction. Client health plans conducted analyses which demonstrated a net savings of 30% in institutional health care costs attributable to the CLM program. In addition, a random telephone survey of CLM patients conducted in November 2004 indicated that 89% of patients had experienced improvements in

activities of daily living, and all patients reported that their quality of life had improved. CLM measured patient satisfaction on a quarterly basis and found that 95% or more of patients were satisfied with the program in all six commercial networks.

CLM was eager to replicate these cost savings and improvements in health care quality among the CMHCB population and enter into a long-term relationship with CMS to serve the needs of high-risk Medicare beneficiaries with chronic conditions. Working as a provider for Medicare would help CLM realize its long-term vision of serving all states in the United States.

1.2.3 CLM Intervention and Comparison Populations

CLM worked with its CMS project officer and analysts from Actuarial Research Corporation (ARC) to develop a methodology for selecting the starting population for the CLM CMHCB program. Beneficiaries had to meet the following three inclusion criteria for eligibility in the CLM CMHCB demonstration program:

- were Medicare fee-for-service beneficiaries with a primary residence in designated counties of California, Florida, or Texas, with high costs in 2004 (i.e., top 5 percent of costs),
- had two or more hospitalizations in 2004, and
- had at least one diagnosis from a list provided by CLM, such as heart failure (HF).

The population was further restricted using the following exclusion criteria:

- were enrolled in or met criteria for any CMS demonstration,
- received hospice care,
- were enrolled in the end-stage renal disease (ESRD) benefit or received dialysis and/or kidney transplant,
- were enrolled in a Medicare Advantage (MA) plan,
- used Medicare as a secondary payer,
- were younger than 18 years of age,
- lacked Medicare Part A and Part B coverage as of September 1, 2005, or
- had at least one of the exclusionary diagnoses designated by CLM (e.g., liver cancer).

Beneficiaries who elected the hospice benefit following the launch of the program remained in the study population, so that CLM physicians could support patients during their transition into hospice. The remaining beneficiaries, after the inclusion and exclusion criteria were applied, were randomized into the intervention and comparison populations.

CLM expanded its service area in Florida into the Orlando area and significantly increased its operations in Northern California for the CMHCB demonstration to obtain enough beneficiaries to populate the intervention group with 15,194 beneficiaries and the comparison group with 6,084. These expansions were due in part to the fact that a large number of beneficiaries who would have otherwise been eligible for the CLM program were excluded due to the co-located CMS demonstrations in California and Texas. CLM also chose to include beneficiaries living in institutionalized settings which permitted CLM to have access to California beneficiaries with HF, the primary diagnosis characteristic of individuals who were likely to benefit from the PVPDS model.

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 CLM 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),
- developed ESRD, or
- died.

Carve-out group—After program implementation, CLM asked CMS to reconsider its intervention population and 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 its planned refreshment of its intervention and comparison populations. CLM found that the claims algorithm used to select beneficiaries for their program yielded a significant number of Medicare beneficiaries who they felt were not sick enough to truly benefit from physician home visits. These individuals often had high medical costs in 2004 due to a procedure performed to address an acute health condition, such as a hip replacement, and their health care costs had regressed to the mean by the launch of the CLM program. As a result, the CLM felt that the presence of such individuals in the intervention population would limit its ability to achieve its required financial savings. Further, many of these beneficiaries opted out of the program, because they did not feel like they needed the services provided. To address this issue, CLM hired Milliman to analyze patient claims and suggest an alternative claims-based algorithm to identify beneficiaries with greater disease severity for inclusion in the refresh population.

CLM requested that the Hierarchical Condition Category (HCC) score be recalculated using calendar year 2005 claims data and a June 5, 2006 Enrollment Data Base (EDB) check of

eligibility information. Beneficiaries from the original starting population would be retained if their HCC score was 2.75 or greater or if they had an HCC score less than 2.75 but had a diagnosis of selected clinical conditions such as peripheral vascular disease, ischemic heart disease, hypertensive heart and/or kidney disease, heart failure, chronic obstructive pulmonary disease (COPD), and asthma. Of the starting intervention population of 15,326, 23%, or 3,572 beneficiaries, were identified as having an HCC risk score less than 2.75 and not meeting the diagnostic criteria for retention.[3] Of the starting comparison population of 6,014, 23%, or 1,365 beneficiaries, were identified as having an HCC risk score less than 2.75 and not meeting the diagnostic criteria for retention.

Refresh population—CLM worked with its CMS project officer and analysts from ARC to develop a methodology to develop the refresh populations for the intervention and comparison groups. There were three inclusion criteria for eligibility for the refresh population. Beneficiaries must have:

- been Medicare FFS beneficiaries with a primary residence in designated counties of California, Florida, or Texas that had claims in calendar year 2005,
- had a 2005 HCC risk score >2.749, and
- had two or more hospitalizations in 2005.

Once these beneficiaries were selected, the following exclusion criteria were applied:

- beneficiaries in HCC groups 51 (drug/alcohol psychosis), 52 (drug/alcohol dependence), 54 (schizophrenia) and 55 (major depressive, bipolar, and paranoid disorders);
- hospitalizations with specific diagnosis-related group (DRG) codes (e.g., acute major eye infections, kidney and urinary tract infections) were not to be counted toward the admission requirement. In addition to the DRG exclusion, inpatient claims that had discharge date equal to admission date were excluded, as well as inpatient claims where the admission date matched the discharge date of a prior claim; or
- beneficiaries who were institutionalized during the last three months of 2005.

The population was further restricted using the following exclusion criteria based on July 12, 2006, EDB information. The program excluded any beneficiaries who

- were enrolled in or met criteria for any CMS demonstration,
- received hospice care,
- were enrolled in the ESRD benefit or receipt of dialysis and/or kidney transplant,

[3] The starting populations were slightly lower than the originally selected populations due to loss of eligibility between time of selection for the demonstration and the start of operations by CLM.

- were enrolled in a Medicare Advantage (MA) plan,
- used Medicare as a secondary payer,
- were less than 18 years of age, or
- lost Medicare Part A and Part B coverage.

As a result, there were 26,990 beneficiaries remaining in the potential refresh population. One issue that had to be handled in the selection of the refresh population was that the characteristics of the beneficiary population was dynamic, so that those determined eligible using EDB data as of July 12, 2006, could lose eligibility as the September 1, 2006, start date approached. Therefore, the refresh population was rescreened for eligibility as of August 1, 2006, using data pulled from the EDB on August 1, 2006. This information was used to identify any members of the refresh population who became ineligible between July 12, 2006, and August 1, 2006. In addition, CMS received telephone numbers for this population from the Social Security Administration (SSA); beneficiaries without known telephone numbers were also removed at the request of CLM. This left 19,073 beneficiaries in the eligible population as of August 1, 2006, who were randomized into an intervention population of 13,623 and a comparison population of 5,450. The randomization was done on a state-by-state basis by ranking the eligible beneficiaries by Medicare Health Insurance Claim (HIC) number. For each group of seven HICs, five beneficiaries were placed in the intervention group and two beneficiaries were placed in the comparison group.

1.2.4 CLM Operations

CLM launched its CMHCB demonstration program October 1, 2005. During the first year of operations, CLM requested that a subset of its existing original population be removed (carved-out). CMS also offered CLM a refresh population at twelve months post launch that would partially offset attrition due primarily to death. CLM worked with CMS to redefine the criteria for identifying the refresh population. The CLM refresh population went live on September 1, 2006. During the second year of operations, CLM received a termination letter with a 90-day notice from CMS. CMS was concerned that the quarterly monitoring reports showed no cost savings, and CLM did not develop a viable plan to change its financial course. The CLM program ended February 29, 2008, or 29 months, after initiation of the original population and 18 months after the start of the refresh population.

CLM received monthly management fees for its intervention population beneficiaries. During the 9-month outreach period for the original population, CLM received fees for all beneficiaries except those who had actively opted out of the program and those without at least one day of eligibility during each month. After the initial 9-month outreach period, CLM accrued management fees for only for those beneficiaries who verbally consented to participate and only during periods of participation. Participation continued until a beneficiary became ineligible for the CMHCB program or opted out of services provided by CLM. Participants could drop out of the program at any time and begin participation again at any time, as long as they were eligible. Beneficiaries who declined participation could be re-contacted by the CLM after a sentinel event, such as a hospitalization or an emergency room visit. For the refresh cohort, CLM received a monthly fee for 7,500 beneficiaries during the 6-month outreach period if they did not opt out,

while a fee was paid for the refresh beneficiaries only if they became participants during the last 12 months of operations.

CLM negotiated a per beneficiary per month payment of $295 for the duration of the demonstration for both the original and refresh populations. The net savings requirements for those CMOs that complete a 36-month demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. However, because CLM's demonstration was terminated on February 29, 2008, the net savings requirement was pro-rated based on the number of months after the effective start date for each cohort for which the demonstration was terminated. In CLM's case, the termination date was 29 months after the start date of the original cohort and 18 months after the start date of the refresh cohort. Therefore, at the time of its termination, CLM was contractually obligated to achieve a 4% savings in Medicare payments among the original intervention group and 2% for the refresh intervention group (regardless of participation in the CLM program) compared with the comparison group and to cover the CLM program fees collected. In addition, CLM had the opportunity to share a portion of any savings that were achieved beyond its obligations. Payments associated with the Medicare Part D benefit were not to be included in these calculations.

1.2.5 Overview of the CLM CMHCB Demonstration Program

RTI conducted a site visit to the CLM's corporate headquarters located in Woodland Hills, California 10 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 CLM and reporting early experiences with program implementation and engagement of eligible beneficiaries, providers, and CMS. Prior to the conduct of a second site visit, CLM received a CMHCB demonstration program termination notice from CMS. RTI conducted a telephone interview with key program staff to learn about changes in the CLM program since the last contact with RTI and about the process of closeout of its program and contract with CMS. The description of CLM's CMHCB demonstration program and its activities in this report reflects CLM's impressions and interpretation of its 29-month experience and does not necessarily reflect RTI's or CMS's perspective on these issues.

The overarching goal of the CLM program was to help participants take an active role in their health and receive timely access to appropriate health and social services. The CLM intervention model provided patient-centric home-based physician care to care for patients with multiple chronic conditions to therefore avoid acute exacerbations that would lead to emergency room visits and hospitalizations. First, we describe the continuum of services CLM reported they provided to CMHCB participants, and then we explain the relationship between PVPs and community-based primary care physicians (PCPs) and the quality improvement infrastructure and information system to support the CLM physicians and nurses who delivered these services.

Personal Visiting Physician Delivery System (PVPDS)—CLM's mission was to provide home-based care and 24/7 access to a PVP to patients with multiple chronic conditions, who were at high risk for multiple hospital admissions. CLM's model represented the institutionalization and professionalization of home-based physician care. The organization's president anticipated that this model of health care delivery would become common practice for

the top 1%-2% of individuals with the greatest health care needs and would result in the development of a new specialty in medicine for physicians who provide home-based care.

CLM's PVPDS was supported by clinical resources, such as nurse care managers, nurse practitioners, and systems, such as electronic medical records, that allowed PVPs to effectively care for its patients. A key element of this model was bidirectional communication between the patient and the PVP—physicians made appointments to see its patients on a routine basis and patients were asked to call their PVPs when problems arose. The model depended heavily on the PVP's ability to bond with his or her patients so that they were comfortable contacting the PVP when they began to experience the early symptoms of an exacerbation (i.e., a time when a PVP could manage the condition by providing care in the home and avoid unnecessary hospitalizations). This ongoing relationship with patients also allowed PVPs to understand the issues that put patients at risk for acute health events and initiate interventions that decreased these risks. Despite the intention to develop an ongoing relationship, the PVP explicitly was not intended to replace the patient's primary care physician.

Below, we describe the process used to assess patients and the types of care provided by the PVPDS, followed by a discussion of the staffing model used to implement the CMHCB program.

Assessment—Once a Medicare FFS beneficiary agreed to participate in the CMHCB program, the enrollment specialist registered this information in the Siebel contact management system, which alerted a nurse care manager that she had a new patient. The nurse care manager called the patient to conduct a brief assessment of his or her care needs to establish the baseline acuity of the patient and scheduled an initial visit with the PVP, usually within 72 hours. During the initial visit, the PVP got acquainted with the patient and conducted an initial in-home assessment to determine the patient's actual acuity, which was entered into CLM Central, the electronic medical record. The PVPs used a 16-item tool to assess patient acuity that was developed by CLM to ensure the reliability and consistency of acuity ratings. Examples of items on the tool include the number of admissions and emergency room visits within the last 6 months, the presence of unmet social and emotional needs, the expected number of PVP visits that would be needed in the next 30 days, and the presence of compliance, psychiatric or ongoing home health issues. As a result, patients were classified into five levels of acuity associated with five levels of care designated by CLM. The PVPs assessed patient acuity on a monthly basis.

Levels of care—The goal of the PVPDS was to help patients move from higher to lower levels of acuity by addressing as many risk factors associated with each patient as possible. For patients with high acuity (i.e., levels 4 and 5), CLM reported that the PVPs would provide acute episodic care, such as a home hospitalization consisting of daily visits to the patient's home for 2 or more days to administer treatments such as IV antibiotics. PVPs were also to provide a "bridge to home," by helping patients who were being cared for in more intensive venues to transition to lower levels of care. For example, a patient discharged home from an emergency room into the care of a PVP rather than get admitted to a hospital.

Care management—At the various levels of care, PVPs, nurse practitioners, and nurse care managers provided comprehensive care management to address the following issues:

- patient adherence to treatment regimens,
- facilitation of the relationship between the patient and his or her community-based provider,
- coordination of care services (e.g., make referrals and/or appointments for specialist or rehabilitative care as needed),
- end-of-life planning, including advance directives and transition to hospice,
- home safety,
- socioeconomic issues,
- psychosocial issues, and
- medication management, including review of prescription regimens to assess duplication of medications, potential for adverse drug events, and patient compliance.

Personal Visiting Physicians (PVPs)—PVPs were not intended to replace community-based primary care physicians; rather, PVPs served as adjuncts to the patients' PCPs. As part of a patient's health care team, PVPs helped patients find a PCP if they did not have an established relationship with a provider in the community. PVPs provided care with a focus on chronic issues and established care goals for each chronic condition that a patient had, ensuring that the patient agreed with the care goals. PVPs discussed the potential causes and complications associated with health conditions using terms patients could understand. PVPs also considered patient preferences, financial resources, and physical ability to undergo medical interventions.

PVPs consulted with community-based providers if significant changes in treatment regimens were indicated and to ensure that they were both implementing a common care plan. CLM's electronic medical record allowed PVPs to fax information collected during each home visit to community-based providers. When patients were hospitalized, PVPs typically conducted social visits to patients to maintain relationships with patients and obtain information about the timing of discharge from the hospital. These visits did not involve clinical care; PVPs did not review patient charts. PVPs scheduled a home visit as soon as conveniently possible and appropriate following hospitalizations to review discharge plans, including changes in medication, with patients.

Nurse practitioners—Nurse practitioners were available to support the PVPs by conducting follow-up visits with patients and administering treatments, such as intravenous antibiotics and wound care. Nurse practitioners could also respond to urgent calls and see patients on an urgent basis if they were available and in the geographic area where the visit was needed. Each nurse practitioner supported approximately five PVPs. In general, nurse practitioners conducted visits with lower acuity patients, who are identified in collaboration with the PVPs.

Nurse case managers—Nurse case managers provided critical support to doctors in their efforts to care for patients and contact patients directly by telephone to implement physician care plans and address issues that arose during the time between PVP visits. Nurse care managers served as patient advocates and coordinated care with patients, family members, community-based primary care physicians, specialists, and home health nurses. Each nurse care manager was responsible for providing support to a single PVP and his or her respective patients.

When a patient first enrolled in the CLM program, the beneficiary was assigned to a PVP and his or her associated nurse care manager. The nurse care manager learned about new patient assignments based on information documented in the Siebel system by the enrollment specialist, which automatically transmitted enrollments to CLM Central. If an enrollment specialist determined that a patient needed to be seen urgently, he or she alerted the nurse case manager via telephone. Subsequently, the nurse care manager called the patient to introduce him- or herself, answer questions about the program, and, in the Los Angeles area, schedule an initial visit with the PVP. In other areas, PVPs made their own appointments. Then, the nurse contacted the PVP to let him or her know that he or she had a new patient. Following the first appointment, the nurse case manager called the patient to ask about his or her experience with the PVP.

During the early months of CLM's program implementation, nurse care managers focused on building relationships with the patients during telephone contact between PVP visits, so that patients would be comfortable calling the nurses if health problems arose. Patients at highest risk were to receive calls on a weekly basis, whereas those at moderate and low risk were to receive calls on a monthly or bimonthly basis. Prior to beginning their care manager roles, nurses received training to help prepare them for such conversations. Orientation to the program included information such as the mission, vision, principles, and culture of CLM; health insurance (including Medicare and Medicaid); policies and procedures (e.g., guidelines related to the Health Insurance Portability and Accountability Act); and telephone communication skills.

CLM continued to evolve its PVPDS program in order to respond to some of the challenges that it faced after implementation. One area of focus was the creation of specifications for a "dynamic" [or continuous] refresh population. The request for a dynamic refresh was motivated by CLM's desire on how it would obtain its intervention population. CLM felt that claims data, especially because of the lag in claims data, were not the best resource to identify beneficiaries that could benefit from the CLM program. Claims data were useful to some extent to identify an initial pool of beneficiaries and for periodic refreshes. However, in addition to claims data, the program wanted to use methods such as assessments conducted by physicians and nurses. For example, a call by a nurse could help to determine whether a particular person was appropriate, in real time, for the CLM level of intervention. CLM also wanted more frequent refreshes from claims even though it was not an ideal way of selecting beneficiaries. Without a dynamic refresh CLM felt that the static ITT population would continuously lose the very ill, who were the primary target of the CLM intervention, to death and would not have adequate numbers of other beneficiaries in the static ITT population becoming sufficiently ill to yield the necessary cost savings given their fee structure. CMS had indicated during the initial demonstration design negotiations and reiterated during negotiations for the refresh population that clinical judgment for selection of the intervention group and matched controls were not an option.

CLM also explored ways to modify its PVPDS program to provide cost-effective care in rural areas where patients were geographically dispersed and physicians would have to drive long distances to conduct eight home visits per day, the desired number from an operations perspective. In addition, CLM found it difficult to recruit physicians who were willing to live in these areas. Examples of options that were considered included PVPs or nurse care managers providing support via telephone or PVPs spending a certain number of weeks per year in a rural area to serve the patient population on a limited basis.

Institutionalized beneficiaries were not excluded from CLM's intervention population. Community relations managers visited administrators of nursing homes to describe the CLM PVP model and to gain support for allowing the CLM PVPs to provide care for eligible residents. Following these initial visits community relations managers also had to gain the support of the facility's director of nursing, the attending physician, and the beneficiary before the PVP could begin working with the patient. This process to enroll beneficiaries living in institutionalized settings was arduous and required significant CLM staff time so much so that CLM opted to exclude this population from the refresh that it had negotiated with CMS.

Using information gleaned from the first year of operations about which elements of the CLM program were effective, CLM made several enhancements to its clinical model, which became effective in early 2007. CLM replaced its panel management scheme with a population-based management approach that involved reorganizing patient care teams to include more nursing support. In particular, physicians started covering a caseload of patients with the support of 1.5 to 2 nurses, rather than just 1 nurse. This new staffing arrangement was intended to allow CLM to enroll a larger number of beneficiaries more rapidly. Specifically, nurse care managers conducted the initial assessments with patients and triaged patients (the enrollment center had initially performed this role). The enrollment center then focused exclusively upon explaining the program to beneficiaries. Nurses within each care team provided telephonic care management to stable patients that had been previously classified as "stand-by monitoring"—rather than maintain a staff of individuals in Arizona that conducted these calls and presented cases to the appropriate clinical team if a problem arose. This new arrangement was intended to connect stable patients who experienced a problem with clinical assistance more quickly because they would have an established relationship with a member of the care team that would provide support during an acute period.

CLM also reported that it enhanced its risk stratification process in the fall of 2006 by introducing two levels of risk and then revised the process again in early 2007 to incorporate three levels of risk. The first revision had the following two indices:

- a severity index that encompassed all aspects of a beneficiary's risk (medical as well as psychosocial); 3 = most severe, most risk unaddressed, 1 = risk that had been addressed; and

- an urgency index that classified patient issues in three categories based on the time sensitivity of a response: whether a patient needed a visit on the same day or required an emergency room visit if he/she could not be seen by a PVP; whether a patient could be managed telephonically or might need a visit; or whether a patient did not require a visit on the same day.

CLM reported that they transitioned to a three-dimensional risk stratification system to in an attempt to ensure that it provided the right care at the right time by the right provider. This added a dimension of management to the severity and urgency indices above. This dimension helped determine the frequency of physician visits and care management calls needed to meet the needs and preferences of participants. CLM felt that this approach addressed their belief that some people did best with more physician visits, while others preferred telephonic contact with a nurse care manager. Approximately 50% of the participants were in the telephonic group while the other 50% were in the active management group. The death rate among the active management group was quite high, therefore CLM anticipated that the percentage of individuals in the telephonic group would increase during the demonstration period.

During the latter part of the demonstration, CLM implemented a telemonitoring pilot. CLM believed that by routinely monitoring objective data it could be alerted when a patient was beginning to have problems, that is, when a PVP visit could effectively reduce the need for an emergency room visit or hospitalization. CLM reported that often by the time a patient called CLM about a problem, he/she was so sick that he/she had to go to the hospital. CLM observed that some participants had difficulty using these devices, and others got bored with the repetitive nature of the questions and stopped using the device. Therefore, CLM implemented a pilot of a home monitoring program at the beginning of July 2007, that provided beneficiaries with appropriate equipment (e.g. blood pressure monitors, pulse oximeters, scales) and beneficiaries self-reported measurements obtained using these devices during telephone calls with care managers. This approach provided an additional opportunity for participants to bond with care managers. CLM used the objective data collected to monitor patient needs for an in-person visit in remote areas, where a "same day" visit was often not feasible.

In a telephone call prior to termination, CLM identified three issues it would have done differently, which it believed may have changed the outcome of its demonstration:

1. CLM would have discharged people on telephonic intervention sooner. CLM was responsible for everyone in the intervention group and tried to intervene with all beneficiaries by maintaining contact with them. CLM felt that some of the beneficiaries were not appropriate for the CLM program or they did not want to participate. CLM felt that it should have dropped them sooner than it did; in hindsight, CLM believed that it should have disenrolled them from the program and stopped collecting their CMS fees.
2. CLM believed that it should have stood firm and negotiated two rates: a lower rate for the telephonic intervention people and a higher rate for beneficiaries who were actively visited. In its view, a two-tier rate, or perhaps a monthly fee that was an average of the two, would have been better for the program. However, at the time CLM was negotiating a fee, CLM had no data to support a two-tier fee calculation because it had no idea of the proportion of the telephonic versus active visit group. For its other clients, CLM was able to conduct ongoing assessments and enroll patients in the program in times of crisis, so everyone receiving interventions was deemed appropriate for their level of care. Lacking the ability to enroll and disenroll beneficiaries in the CMS demonstration, in its view, put the program at a disadvantage.

3. In retrospect, CLM also realized that its model was not a good fit for beneficiaries in rural or other areas that were difficult to staff with health care providers. Covering such a wide geographical area turned out to be a problem; CLM tried to cover as much area as geographically possible in California. For example, in Fresno, which is a growing city, CLM could not find many physicians who were willing to relocate and serve the area. As a result, CLM hired only a few physicians that attempted to cover huge areas and serve a set of patients who were dispersed geographically. CLM felt that its model worked best in densely populated areas.

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 CLM CMHCB demonstration period and provider satisfaction with the CLM CMHCB program culled from interviews with physicians during the site visit. In Chapter 4, we provide the results of our analyses of participation levels in the CLM 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. Supplements to chapters 2, 4, and 7 are available from the CMS Project Officer upon request.

CHAPTER 2
EVALUATION DESIGN AND DATA

2.1 Overview of Evaluation Design

2.1.1 Gaps in Quality of Care for Chronically Ill

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

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

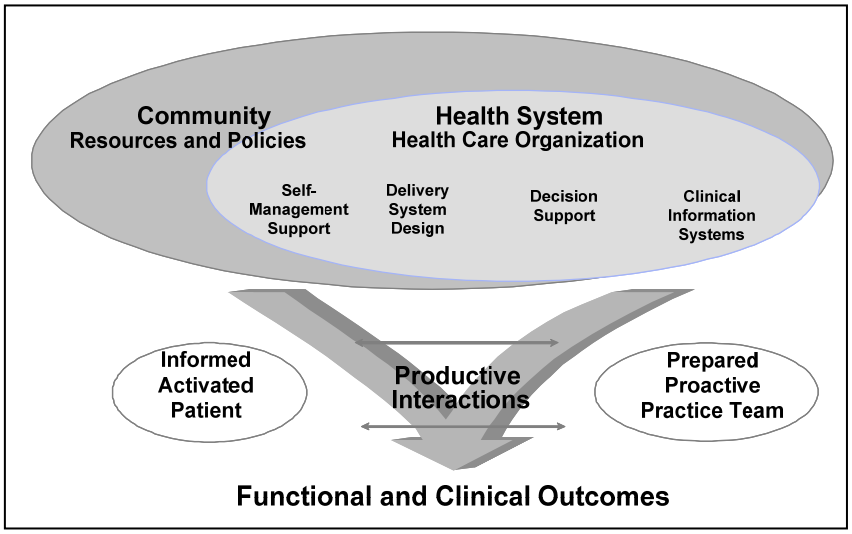
2.1.2 Emerging Approaches to Chronic Care

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 Centers for Medicare & Medicaid Services (CMS) case management demonstration programs in the early 1990s, which required physician consent for patient participation, resulted in increased beneficiary satisfaction but failed to achieve any improvement in health outcomes, patient self-care management, or cost savings (Schore, Brown, and Cheh, 1999). In 2002, CMS selected 15 demonstration programs of varying sizes and intervention strategies as part of the Medicare Coordinated Care Demonstration (MCCD). None of the 15 programs produced any statistical savings in Medicare outlays on services relative to the comparison group, and two had higher costs (Peikes et al., 2009).[4] There were a few, scattered quality of care improvement effects. Two programs did show some promise in reducing hospitalizations and costs, suggesting that care coordination might at least be cost neutral. A major reason given for the lack of success in both Medicare savings and better health outcomes is attributed to the absence of a true transitional care model in which patients were enrolled during their hospitalizations. Studies have shown that approach to significantly reduce admissions within 30/60 days post-discharge, when patients are at high risk of being readmitted (Coleman et al., 2006; Naylor et al., 1999; Rich et al., 1995).

2.1.3 Conceptual Framework and CMHCB Demonstration Approaches

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

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

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

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

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

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

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

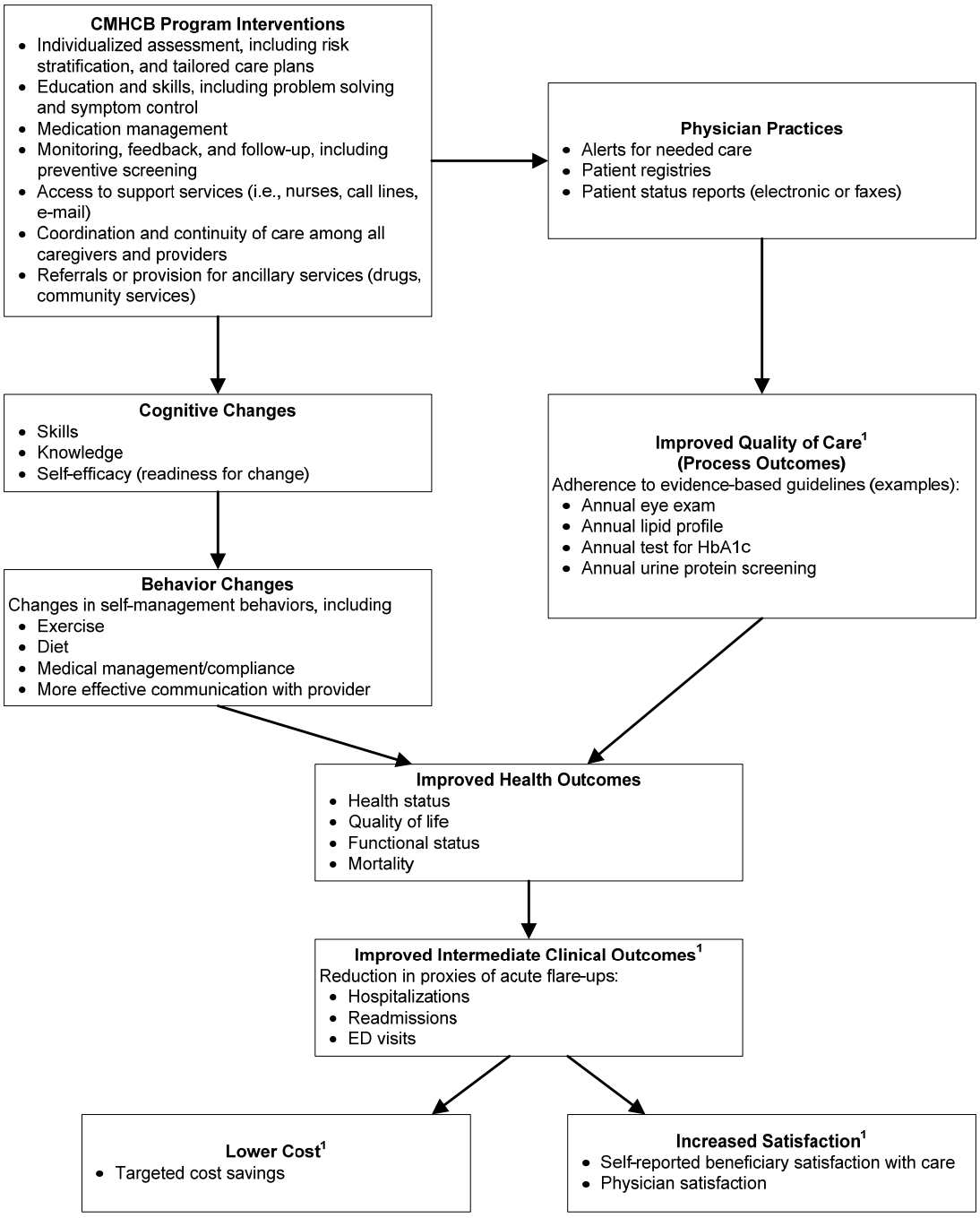
- *Monitoring, feedback, and follow-up*. Activities in this domain include ongoing biomonitoring of beneficiaries by placing scales or other equipment in their homes or by having the beneficiaries self-report their weights, blood sugars, or other measures. When data on preventive services, screenings,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

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

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

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

Figure 2-2
Conceptual framework for the CMHCB programs



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

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

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

In this report, we present our findings with respect to the degree to which CLM was able to engage its randomized intervention population and achieve four outcomes. ***Table 2-1*** presents a summary of research questions and data sources, organized by three evaluation domains: Reach, Implementation, and Effectiveness. The CLM 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 CLM able to implement its programs?				
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 CLM?	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 CLM engage physicians and physician practices in their programs?	Yes	No	No	No
REACH: How well did CLM 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 CLM 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 CLM program?	No	Yes	Yes	No
4. To what extent were the intended audiences exposed to the CLM 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 CLM demonstration intervention versus a low level of intervention?	No	Yes	Yes	No
EFFECTIVENESS: To what degree was CLM 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 CLM 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 CLM 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 CLM program result in greater engagement in health behaviors?	No	No	No	Yes
3. Did the CLM 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 CLM 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 CLM program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?	No	No	Yes	No
3. Did the CLM 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 29 or 16 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 29- or 16-month period compare with the fees that were paid out? How close was CLM 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; CLM = Care Level Management; 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-randomized" scheme, wherein CMS assigns eligible beneficiaries to an intervention or comparison stratum before gaining their consent to participate. In fact, comparison beneficiaries are not contacted at all. Further, beneficiaries opting out of the intervention are assigned to the intervention group, even though they will receive no CMO services. These refusals are included in the same stratum as those receiving care coordination services on an ITT basis.

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

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

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

Our proposed model specification to explain any particular performance variable, Y, is

$$Y = \alpha + \beta_1 I + \beta_2 T + \beta_3 I \bullet T + \beta_4 I \bullet T \bullet P + \beta_5 X + \varepsilon \qquad (2.1)$$

where

α = the intercept term, or reference group;

I = 0,1 intervention indicator;

T = a vector of monthly indicators for the demonstration (1–36);

P = count of days beneficiaries are in intervention group; = 0 for comparison beneficiaries;

X = a vector of beneficiary covariates; and

ε = a regression error term.

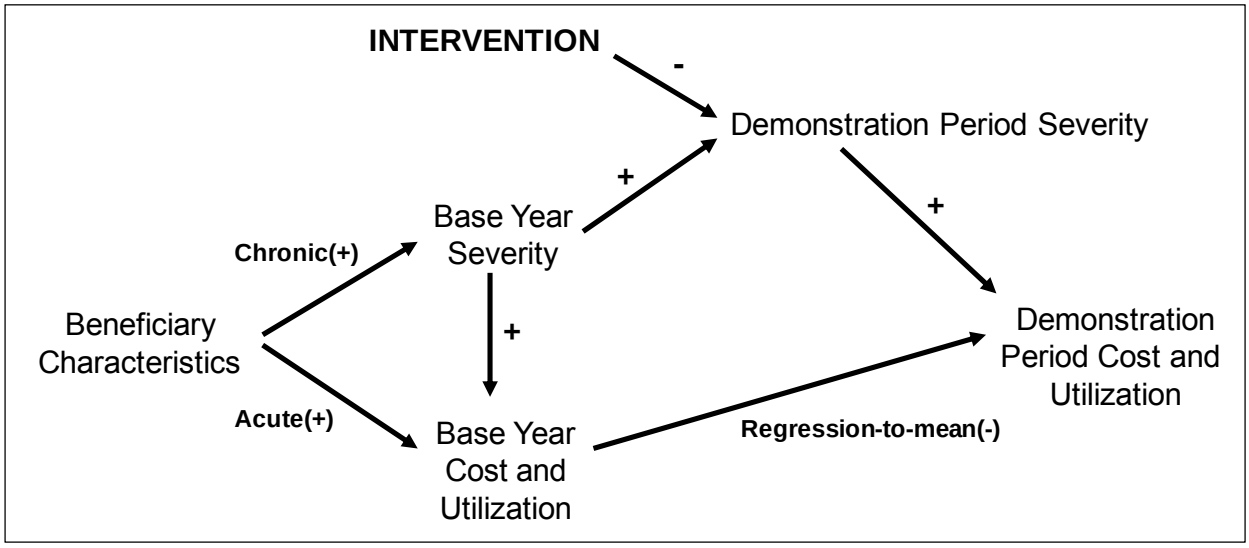
This model uses four sets of variables in analysis of variance format to capture differences between intervention and comparison beneficiaries. The β_1 coefficient provides a test of the difference between the intervention group and comparison group in the base period for a particular outcome variable. (The reference comparison group mean value is in the α intercept.) If preprogram random assignment is successful, β_1 will be approximately zero before controlling for beneficiary-specific (X) factors. The β_2 coefficient tests for temporal changes among beneficiaries, while the β_3 interaction coefficient tests whether the intervention group's performance profile differs over time from the comparison group's performance. The β_4 coefficient(s) capture individual participant deviations from the average intervention profile due to differing durations in the program. As noted earlier, a beneficiary's status of eligibility for the CMHCB demonstration may change multiple times during the 3-year period. The vector of β_5 coefficients controls for beneficiary-specific covariates influencing individual differences in the dependent variable of interest. Including X covariates should set the estimated β_1 equal to 0 if predemonstration randomization is contravened in some way.

Program effects during the demonstration are reflected in the interaction coefficients. The null hypothesis is that the coefficients for both β_3 and β_4 are zero, implying no CMHCB program impact. Estimates that are significant at the 95% confidence level imply distinct program effects, either overall for the CMHCB initiative (β_3) or depending on duration in the program (β_4). These effects may be graphed over the course of the 3 years relative to the comparison group to show at a glance the time path of intervention effects. The model may also be expanded to conduct analyses across beneficiary subpopulations and including 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

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



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.

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 CLM CMHCB demonstration program. As noted in Chapter 1, we also conducted a survey of CLM 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 Actuarial Research Corporation (ARC). The participant status information originates from CLM and was submitted to ARC. This file was updated quarterly and logged status changes among the intervention groups by CLM. 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.

- *High cost finder file*. RTI used this file, produced by ARC, to identify the group into which each CLM beneficiary was randomized—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 CLM comprised of all randomized original and refresh beneficiaries, excluding the carve-out beneficiaries from the original population. RTI used these files to determine daily eligibility based on CLM 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 CLM's 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 extract to obtain demographic characteristics at the time of randomization (September 1, 2005) for CLM's original population.

  - RTI conducted an EDB extract to obtain demographic characteristics at the start date (September 1, 2006) for CLM'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 CLM's start date and claims with processing dates that span the full intervention period and 9 months thereafter (or claims run out).

- *CMO beneficiary intervention data files*. Quarterly, CLM sent RTI limited beneficiary-level intervention files that contained counts of intervention activities, such as the number of competed calls to participants, the number of in-person visits, etc. 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 for CLM

Ineligibility reasons	Description
Death	Ineligible beginning on day following date of death.
ESRD	Ineligible beginning on day of ESRD enrollment. Eligible on day following ESRD disenrollment.
MA plan	Ineligible on day of MA plan enrollment when GHO contract number does not equal the contract number for CLM. 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: CLM = Care Level Management; ESRD = end-stage renal disease; MA = Medicare Advantage; GHO = Group Health Organization.

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

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 randomization, September 1, 2005 for the original population and the refresh go-live date (September 1, 2006) for the refresh population. Medicaid enrollment was determined at any time during the baseline period and was also determined using the EDB.

Table 2-3
Analysis periods used in the CLM 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				
10/1/05	2/29/08	29	10/1/04	9/30/05
Refresh Population				
9/1/06	2/29/08	18	9/1/05	8/31/06

NOTES: CMHCB = Care Management for High Cost Beneficiaries; CLM = Care Level Management.

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, for the participation analyses, we examine the full intervention period. 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 CLM was active in the demonstration. The numerator is the number of days the beneficiary is eligible during that time period. CLM participated in the demonstration for 29 months, so the number of days in the denominator for each original population beneficiary in CLM is 882 (CLM end date minus CLM 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 882, or 0.476.

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 the CMHCB population was also required to have used services to be selected for randomization.

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

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

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

physician and/or OPD claims had to have occurred on different days. The diagnosis codes used to identify these clinical conditions are in ***Supplement 2A***.

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

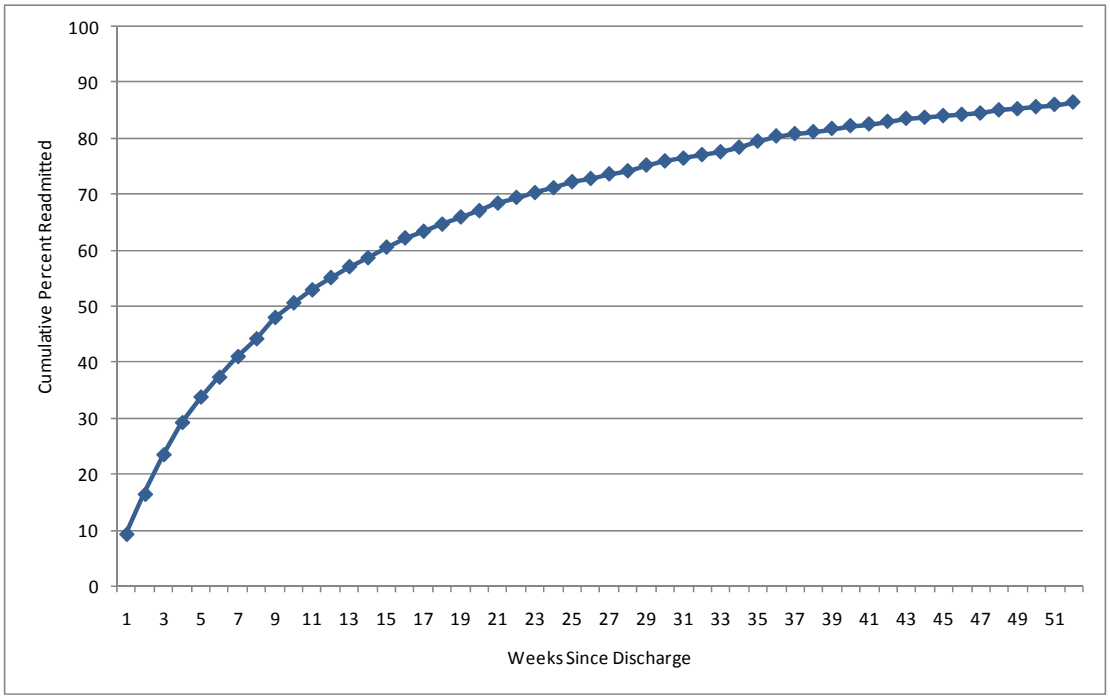
Only claims that occurred during periods of eligibility were included in the utilization measures. For both the demonstration and baseline periods, claims were included if services were started during days that the beneficiary met CLM's CMHCB program eligibility criteria, as determined from the ARC daily 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 upwards of 50% of subsequent admissions in our analyses[5].

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

Figure 2-4
Percent with readmission for any diagnosis: CLM'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 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 3 through 14 and months 15 through 26 and included readmissions through months 17 and 29, respectively. The intervention period for the refresh population examined admissions during months 4 through 15 and readmissions through month 18. 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 29-month

original intervention period and the 18-month refresh intervention period. Payments associated with end-stage renal disease (ESRD) services were excluded from the demonstration program payment amounts as beneficiaries lost CMHCB program eligibility at the point they became eligible for the ESRD benefit.

7) ***Guideline Concordant Care***. We define quality of care as adherence to evidence-based, guideline-concordant care and selected four process-of-care measures as the focus of our evaluation for this report. The measures and relevant disease population (when applicable) are as follows:

   - rate of influenza shots during influenza season (September through February) for adults,
   - rate of annual HbA1c testing for beneficiaries with diabetes in the baseline period,
   - rate of annual low-density lipoprotein cholesterol (LDL-C) testing for beneficiaries with diabetes or ischemic vascular disease in the baseline period, and
   - rate of annual oxygen saturation assessment for beneficiaries with COPD in the baseline period.

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 CLM. 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 four process-of-care measures were included regardless of CMHCB 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 CLM in case the services occurred after exposure to the CMHCB intervention and during the intervention period. One could envision that CLM 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 CLM, we constructed counts of the number of contacts with the participants—either telephonically or in-person—as well as total contacts (both). Additionally, we constructed counts by type of provider—physician or nurse coach.

CHAPTER 3
BENEFICIARY AND PHYSICIAN SATISFACTION

3.1 Beneficiary Satisfaction

The Care Management for High Cost Beneficiaries (CMHCB) programs' principal strategy to improve quality of care while reducing costs is by empowering Medicare beneficiaries to better cope with their chronic disease(s) and manage their care. The programs do this in three ways: (1) by enhancing beneficiary knowledge of their chronic condition through educational and coaching interventions, (2) by improving beneficiary communication with their care providers, and (3) by improving beneficiary self-management skills. Successful interventions should alter beneficiaries' use of medications, eating habits, and exercise, as well as promoting more effective interaction with their primary health care providers. The CMHCB programs hypothesized that lifestyle changes and better communication with providers would mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions[6].

The 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 the Centers for Medicare & Medicaid Services (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

---

6 In our survey, we examine satisfaction more broadly than satisfaction with a particular member of their health care team or a particular member of the CLM 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 CLM intervention on the beneficiaries' assessment of satisfaction that their full health care team is helping them to cope with their chronic conditions.

Ambulatory Consumer Assessments of Health Plans Survey (CAHPS®) measures of communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

3.1.1.1 Measures of Experience and Satisfaction with Care

The impact of the care management organization (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 o**r 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 CMO interventions on the outcomes appearing in the conceptual model. The research design for this evaluation involved only a single-up survey. 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 Hierarchical Condition Category (HCC) risk score and per beneficiary per month (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 help to further level the playing field when evaluating the impact of the CMHCB 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 Enrollment Database (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. CLM's survey was conducted between June 11, 2007, and October 10, 2007.

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

The target was 300 completed surveys for the intervention and comparison populations. From the sample frame for each group, we randomly selected 300/.7 = 429 beneficiaries. The response rate for CLM was 62%. 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 Care Level Management

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

Table 3-1
Medicare Health Services Survey: estimated intervention effects for experience and satisfaction with care,
CLM
(N = 504)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Helping to cope with a chronic condition (1 to 5)	3.84	3.71	0.19	N/S
Number of helpful discussion topics (0 to 5)	2.23	2.04	0.20	N/S
Discussing treatment choices (1 to 4)	3.21	3.05	0.23	*
Communicating with providers (0 to 100)	77.7	72.9	6.55	**
Getting answers to questions quickly (0 to 100)	68.2	63.7	4.90	N/S
Multimorbidity Hassles score (0 to 24)	3.23	3.51	-0.44	N/S

NOTES: CLM = Care Level Management; 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.

N/S means not statistically significant.

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

Overall experience: helping beneficiary to cope with chronic condition—The average score for the key satisfaction outcome item that assessed how well the health care team helped beneficiaries cope with their illness was 3.8 for the intervention group, or about midway between "very good" and "good" ratings. The average score for the comparison group was about 3.7. The difference is not statistically significant. More than 62% of CLM beneficiaries rated their experience as "excellent" or "very good," and approximately another one-third selected "good." It is not uncommon among the elderly to report high satisfaction ratings. For that reason, the mean scale score was used in the analyses to capture transitions between all response categories. Beneficiaries with higher PBPM Medicare expenditures reported significantly higher ratings on the overall satisfaction outcome than beneficiaries with lower PBPM expenditures.

Across the six measures of experience and satisfaction with care, we observe two statistically significant positive intervention effects. Beneficiaries in CLM's demonstration program reported significantly higher scores for communication with health care team and in discussions of beneficiary treatment choices. For four other measures of experience and satisfaction with care, we found that the effects were in the desired positive direction but not statistically significant.

3.1.3.2 Self-Management

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 CLM.

Table 3-2
Medicare Health Services Survey: Estimated intervention effects, self-management, CLM (N = 504)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Percent receiving help setting goals	62.7	54.0	6.1	N/S
Percent receiving help making a care plan	59.6	51.9	3.9	N/S
Self-efficacy ratings				
Take all medications (1 to 5)	4.32	4.20	0.20	N/S
Plan meals and snacks (1 to 5)	3.90	3.83	0.15	N/S
Exercise 2 or 3 times weekly (1 to 5)	3.62	3.40	0.20	N/S
Self-care activities				
Prescribed medications taken (mean # of days)	6.70	6.64	0.06	N/S
Followed healthy eating plan (mean # of days)	5.00	5.17	-0.03	N/S
30 minutes of continuous physical activity (mean # of days)	3.21	2.78	0.63	*

NOTES: CLM = Care Level Management; ANCOVA = Analysis of Covariance.

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

N/S means not statistically significant.

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

Computer program: CreqD2

Setting goals and making a care plan—The survey included two questions that asked if someone from their health care team helped set goals or a plan to take care of their health problems. Sixty three percent of CLM beneficiaries in the intervention group reported receiving help setting goals compared with 54 %, although the difference is not statistically significant. Similarly, 60% of CLM beneficiaries in the intervention group reported receiving help making a

care plan compared with 52 %, where the difference is also not statistically significant. These results reveal that CLM's intervention did not significantly affect either the percentage of beneficiaries who had received help to set goals for self-care management, or the proportion of beneficiaries reporting that they had help from their health care team in making health care plans. In terms of receiving help with care plan, Hispanic CLM beneficiaries and beneficiaries with higher PBPM expenditures were significantly less likely to receive such a service.

Self-efficacy ratings—To assess self-efficacy, respondents were asked how sure they were that they could perform each of three specific behaviors. Overall, CLM beneficiaries typically reported relatively high levels of self-efficacy with mean ratings averaging from 3.4 to 4.3 (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 (4.3 for the intervention group versus 4.2 for the comparison group), and the lowest scores were for getting exercise 2 or 3 times per week (3.6 for the intervention group versus 3.4 for the comparison group). For CLM, we observe no significant intervention effects for any of the self-efficacy measures. In terms of other characteristics, African American CLM beneficiaries, CLM beneficiaries without additional insurance coverage, proxy respondents, and those with higher baseline HCC scores were less likely to express confidence about planning their meals and snacks, and those with higher PBPM expenditures were more likely to express confidence about this outcome. Females, proxy respondents, and those with higher baseline HCC scores were also less likely to express confidence about 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 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 beneficiaries said they took their medications as prescribed ranged from 6.7 to 6.6 out of 7 days, but the mean number of days that beneficiaries said they had 30 minutes of continuous physical activity ranged from 3.2 to 2.8 days. For self-care activities, we observe one positive intervention effect for CLM for the frequency of maintaining 30 minutes of continuous exercise. No statistically significant intervention effects were found for CLM for adhering to prescription medications and dietary guidelines. In terms of other characteristics, African American and proxy CLM respondents and respondents with higher baseline HCC scores were less likely to comply with their prescribed medications, and those with higher PBPM expenditures were more likely to follow their medication regiment. Mail CLM respondents were less likely to adhere to health eating plans, and females, proxy respondents, and respondents with higher baseline HCC scores were significantly less likely to engage in physical activity.

3.1.3.3 Physical and Mental Health Functioning

Physical and mental function—***Table 3-3*** displays the mental and physical functioning outcomes for CLM. On average, CLM participants reported better mental health than physical health functioning, which is consistent with the general Medicare population. The mean PHC score for the intervention group was 30, significantly higher than the score for the comparison group, which was 28. The ANCOVA results revealed that this is the only one statistically significant intervention effect for physical and mental functioning outcomes. The mean MHC

score for the intervention group was 38.8 and the PHQ-2 score of 1.9, compared with 37.6 and 2.1 for the comparison group. For mental health outcomes, there was no difference in mental health functioning as a result of the CLM intervention. Among other characteristics, PHC scores for CLM beneficiaries were significantly lower for females, proxy respondents, those with lower baseline HCC score and for those who completed the survey by mail. MHC scores were also significantly lower for females and proxy respondents. CLM beneficiaries with higher PBPM expenditures had significantly higher MHC scores. Because higher PHQ scores indicate greater depressive symptoms, the depressive symptoms declined with age, years of education, and increase in PBPM spending but increased significantly for proxy respondents.

Table 3-3
Medicare Health Services Survey: Estimated intervention effects, physical and mental health function, CLM
(N = 504)

Outcome	Intervention mean	Comparison mean	ANCOVA-adjusted intervention effect	Stat. sig.
PHC score (physical health, mean =50, std=10)	30.3	28.4	2.1	*
MHC score (mental health, mean =50, std=10)	38.8	37.6	1.7	N/S
PHQ-2 score (depression, 0 to 6)	1.93	2.14	-0.26	N/S
Number of ADLs difficult to do (0 to 6)	2.72	2.79	-0.03	N/S
Number ADLs receiving help (0 to 6)	1.68	1.67	0.00	N/S

NOTES: CLM = Care Level Management; ANCOVA = Analysis of Covariance; PHC = Physical Health Composite; MHC = Mental Health Composite; PHQ-2 = Patient Health Questionnaire 2; ADLs = activities of daily living.

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

N/S means not statistically significant.

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

Computer program: CreqD2

Activities of daily living—A related measure of beneficiary functioning is the ability to perform basic ADLs. On average, CLM respondents reported limitations on 2.7-2.8 ADLs and received help with an average of 1.7 ADLs. There was no significant difference in ADL difficulties or help as a result of the intervention. For CLM members, when other characteristics are held constant, females reported significantly more ADL limitations than males, African Americans reported more than members of other races, and proxy respondents reported more than self-respondents. As expected, those with higher baseline HCC scores also reported significantly higher levels of functional impairment. Survey respondents in CLM with higher PBPM spending reported significantly fewer ADL limitations and fewer ADLs for which they received help. CLM members who had additional health insurance coverage reported fewer ADL

limitations than those who only had Medicare. In terms of needing help with ADLs, the patterns are similar: females, proxy respondents, and members with higher baseline HCC score reported needing help on a significantly higher number of ADLs. Those with additional health coverage also reported needing help with fewer ADLs.

3.1.4 Conclusions

The CMHCB demonstration employs strategies to improve quality of care for high cost Medicare beneficiaries while reducing costs by empowering Medicare beneficiaries to better manage their care. CLM hypothesized that lifestyle changes and better communication with providers will 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, four statistically significant positive intervention effects were found: discussion of treatment choices, communication with providers, 30 minutes of continuous physical activity, and most notably, physical health.

3.2 Physician Satisfaction

RTI made one site visit to meet with the CLM program staff during the demonstration period. The site visit was conducted in July 2006, 9 months after initiation of the CLM demonstration program. During this visit, RTI evaluators consulted with the senior management of CLM and key program staff. They also spoke by telephone with two randomly selected, community-based physicians. RTI had planned to conduct a more comprehensive evaluation of physician satisfaction with the CLM demonstration program during its second site visit to be held approximately 18 months after program launch. A second site visit was not conducted because the CLM demonstration program was terminated by CMS in early 2008.

In this section, we begin by describing the outreach efforts of the CLM program to community-based physicians and sharing beneficiary information with those physicians. We conclude with an assessment of the value of the CLM program to the interviewed physicians.

3.2.1 Care Level Management Outreach to Physicians

A major goal of the CLM program was to use Personal Visiting Physicians (PVPs) to provide intensive case management for patients in their homes and nursing facilities on an urgent as well as on a routine basis 24 hours a day, 7 days a week. The PVP was the case manager in the Personal Visiting Physicians ™ Delivery System (PVPDS). PVPs were supported by Personal Care Advocate Nurses (PCANS) who were based in nearby regional offices.

CLM's process for conducting outreach with primary care physicians (PCPs) evolved over the initial months of program operations. CLM developed a list of all physicians associated with claims during 2004 for their intervention group. CLM sent a letter from CMS introducing the CLM CMHCB demonstration program to all physicians on this list. CLM chose to focus additional outreach efforts on providers who were identified by participating beneficiaries as their primary care physicians.

Initially, CLM reached out to community providers with PVPs, contacting physicians after they had conducted their initial assessment of the patients. Because community physicians were often unfamiliar with the program, they often responded to PVP calls with resistance and were unwilling to cooperate with the program. When this occurred, community relations managers were sent to visit the offices of the primary care physicians to explain the benefits of the program, which often resulted in provider support of the program. CLM adjusted this procedure so that enrollment specialists asked beneficiaries to identify their primary care providers once they agreed to participate in the CMHCB program, and a community relations manager proactively contacted the provider to explain the PVPDS model. Once the community relations manager successfully reached a provider, the PVP associated with the provider's patient was alerted so that he or she could contact the PCP.

3.2.2 Sharing of Information/Ongoing Relationship with Providers

One of CLM's goals was to support patient relationships with their community-based PCPs and specialists. For example, PVPs encouraged patients to make and keep appointments with PCPs and made referrals to specialists, as appropriate. PVPs also used a variety of communication channels to inform PCPs of changes in patient health status. Specifically, PVPs faxed summaries of home visits to PCPs, as well as called PCPs when urgent issues arose. Similarly, CLM requested clinical information about patients from PCPs, such as laboratory results, to facilitate care coordination efforts; however, CLM was sensitive to the fact that PCPs are busy and tried not to make burdensome requests. Since each PCP had only one to three patients who were eligible for the CLM CMHCB program, they generally received information about all patients from the same PVP.

Physicians who expressed concern about the extra burden associated with working with CLM were offered the opportunity to enter into a co-management agreement with CLM and receive a monthly $70 per patient co-management fee.[7] The agreement required PCPs to respond to information requests within 1 day, inform their office staff and colleagues about CLM, and "interact meaningfully and collaboratively in a consultative manner, with the CLM physician to successfully coordinate the beneficiary's care."[8] As of July 2006, PCPs of 196 patients were receiving co-management fees from CLM. CLM noted that there was a limited but positive difference between the level of responsiveness of community physicians with co-management agreements and those without; not all physicians with co-management agreements were fully supportive of the program, but all were at least minimally receptive.

3.2.3 Physician Assessment of the CLM Demonstration Program Value

Community relations managers also made presentations to a small number of physician groups, such as those associated with independent practice associations; however, most contacts with physicians were conducted on a one-on-one basis. Between October 1, 2005, and June 30,

---

7 The co-management fee is offered to physicians only for patients being actively managed by CLM PVPs. Therefore physicians are not eligible to receive this fee for patients who are managed by telephone only (i.e., standby status).

8 CLM Co-Management Agreement.

2006, community relations managers contacted 2,483 physicians and classified physician responses to the program into the following categories: *supportive, receptive, neutral*, and *nonsupportive*. Almost 17% of providers were supportive of the program and believed that the PVPDS was a great idea. Providers who were particularly supportive of the program requested to enroll additional patients from their practices in the program. Fifty-nine percent of physicians were receptive to the program (i.e., they believed the delivery model would save money); however they were not committed to a model where another physician provided care to their patients. One out of every five providers was neutral and did not care if their patients wanted to participate or whether they received information from CLM about their patients. Only 4% of physicians contacted (n = 105) were nonsupportive, expressing concern about losing their patients and frustration with CMS.

RTI interviewed two community physicians who were randomly selected and found that both were aware of the CLM program. The physicians believed that home visits for chronically ill patients were potentially valuable. In fact, one of the physicians actually conducted home visits with his patients, so this physician did not feel that the CLM program added value for his patients. The second physician cared for approximately 700 Medicare patients and had found that even if he had time to conduct house calls, Medicare often denied claims for these visits. Further, he found that Medicare did not usually reimburse for office visits that focused on preventive health care. Therefore, he was pleased to see Medicare supporting preventive care through the CMHCB demonstration program. With two patients in the program, this physician received the co-management fee for one of his patients. He was impressed by the level of detail included in the CLM PVP's chart notes and was happy to know that the PVP would call him if his patient's condition changed significantly.

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

4.1 Introduction

Our participation analysis is designed to critically evaluate the level of engagement by the care management organization (CMO) 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 Care Management for High Cost Beneficiaries (CMHCB) 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 CMO 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 CLM 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 CLM program?

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

- What beneficiary characteristics predict a high level of CLM demonstration intervention versus a low level of intervention?

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 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 July 2006 site visit, Care Level Management (CLM) reported that they had engaged 54% of their CMHCB intervention population, a little lower than the company's goal of 64% participation (Brody and Bernard, 2006). 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 CMO and its randomized beneficiaries. The CLM intervention had a variety of telephonic and in-person elements (e.g., it facilitated patient relationships with physicians and helped patients to comply with physician care plans, hospital discharge planning support, support patient adherence to medication

regimens, education related to self-management activities to decrease risk for acute exacerbations of chronic diseases, and targeted care management support for nursing home residents). Therefore, we examine the number of telephonic and in-person contacts between CLM and their participants. For each participating beneficiary, CLM provided the Centers for Medicare & Medicaid Services (CMS) with a count of the number of completed telephone calls, visits, and written communications with care managers as well as calls with physicians. CLM also provided information on the nature of the contacts (e.g., discharge planning, care planning, end-of-life planning).

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 Actuarial Research Corporation (ARC) in the *Participant Status* file to align with dates of eligibility for the CMHCB 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 CMHCB program) and the percentage of beneficiaries participating for more than 75% of their eligible months.[9] These latter two sets of numbers provide an estimate of the number of beneficiaries with whom the CLM had the greatest opportunity to intervene. Because beneficiaries lose eligibility for various reasons over time (e.g., loss of Part A or Part B benefits, or due to death), we report counts of full-time equivalents (FTEs) or numbers of intervention and comparison beneficiaries weighted by the fraction of the demonstration period each beneficiary was eligible. Only beneficiaries who had at least 1 day of eligibility in both the baseline and demonstration periods are included in these analyses.

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

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

where P_i = the probability that the *i*th individual will consent to participate, βX_i = an index value for the *i*th individual based on the person's specific set of characteristics (represented by the vector), and e = the base of natural logarithms. The probability of a beneficiary being in the participant group is thus explained by the variables.

---

[9] 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), dies, or develops end stage renal disease (ESRD).

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 days). The participation model investigates whether group membership is influenced by beneficiary demographic attributes, clinical characteristics, and utilization and cost factors previously defined in Chapter 2. The demographic variables included in the model are defined as follows from the Medicare enrollment database (EDB) and determined as of the date of randomization for the original population (September 1, 2005) and the start date for the refresh population (September 1, 2006):

- 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;
- urban, a dichotomous variable, set at 1 for beneficiaries with ZIP codes within metropolitan statistical areas; 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 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 Charlson score medium and high, two dichotomous variables set at 1 if the Charlson index score was 1 (medium) and 2 or greater than (high); Charlson score of 0 is the reference group;
- baseline per beneficiary per month (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 CLM original demonstration program was greater than or equal to $1,671 and less than $4,034 (medium) and $4,034 or greater (high); PBPM cost less than $1,671 is the reference group for the original population. Baseline PBPM costs greater than or equal to $780 and less than $3,287 were assigned to the medium group and $3,287 or greater to the high category; PBPM cost less than $780 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 0.62 but less than 1.95 (medium) and greater than or equal to 1.95 (high); concurrent HCC score less than or equal to 0.62 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 0.886 and 2.0 and values greater than or equal to 2.0 were assigned to the high category; a concurrent HCC score less than or equal to 0.886 is the reference group.

4.2.2 Level of Intervention Analysis Methods

On a quarterly basis, CLM reported the number and nature of contacts with participating beneficiaries at the beneficiary level. We use these data to develop estimates of the level of intervention provided to CLM participants. A cornerstone of CLM's program was physician home intervention (PHI) visits. CLM provided intensive care management through a distributed network of Personal Visiting Physicians (PVPs) who see patients urgently and routinely in their home and nursing facilities and are available 24 hours a day, 7 days a week (Brody and Bernard, 2006). The PVP is the case manager in the Personal Visiting Physicians ™ Delivery System (PVPDS). PVPs are supported by Personal Care Advocate Nurses (PCANs) who are based in nearby regional offices and who provide care coordination and maintain regular phone contact with beneficiaries. Four PHI types of visits were provided for home care for ambulatory care sensitive conditions, home hospitalizations, care that addresses barriers to traditional care delivery system, and end-of-life care. In addition, data on physician visits made to participating beneficiaries for a number of other reasons were provided: acute intervention for a high complexity–level medical condition; acute intervention for a low complexity–level medical condition; follow-up after an acute visit; regularly scheduled maintenance visit with treatment changes; regularly scheduled maintenance visit without treatment changes; initial visit primarily for purposes of introducing the program and conducting an initial physical assessment; and visit for purpose of performing a history and physical examination. Counts of telephonic encounters with either physicians or PCANs were also provided.

Using the encounter data submitted by CLM, we constructed counts of the number of contacts with participants, in total and by three stratifications: PHI visits, all physician visits, and telephone contacts with either a physician or a nurse. We also 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 CLM and 0 for beneficiaries who had a low level of contact. Beneficiaries who had a medium level of contact with CLM were excluded from 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 from month 4 through the end of the demonstration program and
- multiple intervention period hospitalizations set at 1 if the beneficiary had more than one hospitalization from month 4 through the end of the demonstration program.

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

We report levels of intervention with the original intervention population starting with month 7 and ending with month 18. We excluded the first 6 months of the demonstration period from this analysis to reduce confounding intervention contacts with engagement contacts. Because beneficiaries could have intermittent periods of eligibility and participation, we restricted inclusion in this analysis to beneficiaries who were eligible for and participating in the CLM CMHCB demonstration program for each month from month 7 to month 18. This is the subset of beneficiaries with whom CLM would have had the maximum opportunity to intervene. Beneficiaries who died during this period but were fully eligible and participating up to their deaths were also included. The number of intervention beneficiaries that met these criteria was 4,041.

4.3 Findings

4.3.1 Participation Rates for the CLM 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 period and exclude the beneficiaries that were carved out of the original population. The results are based on the full demonstration period for both the original and refresh population. The number of months for the full demonstration period for CLM is 29 months for the original population and 18 months for the refresh. However, CLM did not provide participation data for the last 2 months of operation, so all participation analyses cover a time

period that is 2 months shorter (27 months for the original population and 16 months for the refresh).

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

1) *Number of beneficiaries.* The number of beneficiaries is equal to all beneficiaries who had at least 1 day of eligibility in the 1-year baseline period *and* had at least 1 day of eligibility in the period tabulated.
2) *Full-time equivalents.* FTEs defined here are the total number of beneficiaries weighted by the number of days eligible in the intervention period divided by the total number of days in the intervention period. For example, a beneficiary in CLM had a total of 27 months (or 822 days) of possible enrollment. If they died after 90 days, their FTE value would be 90/882 or 0.109 FTEs. If someone was eligible for all 27 months, then his or her value is 1. The sum of this value across all beneficiaries gives us the total FTE value reported in the tables below.
3) *Number fully eligible.* The number fully eligible is the number of beneficiaries that had no gap in CMHCB program eligibility during the demonstration period.

The ratio of FTEs to the total number of eligible beneficiaries in the original intervention population is 0.81 for the entire intervention period (months 1-27) compared with a higher ratio for months 4-15 (.90) and months 16-27, the last 12 months (.92). These differences in ratios illustrate the effect of subsetting to beneficiaries in the different time periods. The ratio is higher in the latter two 12-month periods because we dropped from the eligible population anyone that died or lost eligibility in the first 3 months. Loss of eligibility was primarily due to mortality. Beneficiaries also became ineligible for participation in the CMHCB program if they joined a Medicare Advantage (MA) plan, developed end-stage renal disease (ESRD), lost Medicare Part A or B eligibility, Medicare became a secondary payer, or they moved out of the service area.

Twenty-five percent of the original intervention beneficiaries and 36% 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. Eligibility was higher for participants and lower for nonparticipants. CLM's nonparticipant group was eligible only 65% of the days—much lower than the 90% of days for participants. Also, the participant group had a much higher rate of beneficiaries being fully eligible for the entire intervention period (75%) compared with 48% for the nonparticipant group.

Table 4-2 displays eligibility data for the refresh population, which is a little larger than the original population due to the removal of the carve-out population from the original population. The ratios of total number of beneficiaries to FTEs were the same for the full 16 months (0.84) and the last 12 months (0.88) for both the intervention and comparison populations. However, the percent of beneficiaries that were fully eligible for the time periods is higher among participants (78%) than nonparticipants (54%) or the comparison group (71%).

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

Characteristics	Months 1-27	Months 4-15	Months 16-27
Intervention group			
Number eligible[1]	11,516	10,851	9,014
Full time equivalent[2]	9,274	9,801	8,251
Number fully eligible	7,445	8,918	7,540
Participants			
Number eligible	7,289	6,357	6,094
Full time equivalent	6,524	6,027	5,669
Number fully eligible	5,437	5,578	5,225
Participants > 75%			
Number eligible	3,817	4,002	5,074
Full time equivalent	3,575	3,787	4,686
Number fully eligible	3,107	3,516	4,306
Non-participants			
Number eligible	4,227	4,444	2,857
Full time equivalent	2,750	3,767	2,550
Number fully eligible	2,008	3,338	2,291
Comparison group			
Number eligible	4,561	4,308	3,558
Full time equivalent	3,673	3,888	3,257
Number fully eligible	2,931	3,521	2,966

NOTES:

FFS = fee-for-service; CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries.

CLM provided participation information for 27 of the 29 months of their original intervention period.

[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/CLM/9mo/tables/tabCLM-1.sas 08JUL2009

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

Characteristics	Months 1-16	Months 5-16
Intervention group		
Number eligible[1]	13,102	11,954
Full time equivalent[2]	11,024	10,508
Number fully eligible	9,216	9,227
Participants		
Number eligible	8,670	8,301
Full time equivalent	7,903	7,560
Number fully eligible	6,802	6,715
Participants > 75%		
Number eligible	3,955	6,136
Full time equivalent	3,696	5,552
Number fully eligible	3,262	4,935
Non-participants		
Number eligible	4,432	3,555
Full time equivalent	3,121	2,917
Number fully eligible	2,414	2,490
Comparison group		
Number eligible	5,240	4,784
Full time equivalent	4,424	4,224
Number fully eligible	3,703	3,708

NOTES:

FFS = fee-for-service; CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries.

CLM provided participation information for 16 of the 18 months of their original intervention period.

[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/CLM/9mo/tables/tabCLM-1.sas 08JUL2009

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

Participation rates for the CLM original population. Of all CLM original intervention group beneficiaries, 63% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. We previously reported (McCall et al., 2008) that, among the carve-in population, 38% consented in the initial 6-month engagement period and we observe a large increase in CLM's enrollment over the entire intervention period. Only 44% of beneficiaries were continuous participants (***Table 4-3***), which equates to 70% of participants. Of CLM beneficiaries, 23% refused to participate. The percent not contacted or unable to be located was 14%.

Participation rates are 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 CLM's intervention beneficiaries, 43% participated for more than 75% of their eligible months, which mirrors the continuous participant percentage. ***Table 4-3*** also reports the number of participants over time (for months 6 and 16 and month 27, the last month of CMHCB program operation for which we have participation data). The number of participants increased from months 6 to 16, but then declined over time as would be expected given the attrition due to loss of eligibility or death.

Participation rates for the CLM refresh population. CLM revised their criteria for selecting beneficiaries for their refresh population in order to have a population that they believed would benefit more from the services that they offered and would, therefore, be more likely to want to participate, and that could be contacted (that is, they had a phone number on file with the Social Security Administration [SSA]). With the selection criteria changes, there was modest improvement in their participation rate ***(Table 4-4)***. Overall, 66% of the refresh intervention beneficiaries consented to participate at some point during the 16-month period. Of those, 56% were continuous participants, which equates to 85% of participants. The percent that refused to participate went up slightly (24%), but the percent that were not contacted or were unable to be contacted decreased to 10%.

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

Characteristics	Statistic
Number of intervention months	27
Participation rate (entire demonstration period)	63%
Length of participation	
Continuous participation after engagement	44%
After initial participation, became a continuous non-participant	13%
Intermittent participation	7%
Nonparticipation (never agreed)	37%
Refused to participate when contacted	23%
Not contacted/unable to be contacted	14%
Beneficiaries participating more than 75% of days	43%
Number of participants in selected months[1]	
Month 6	4,235
Month 16	5,224
Month 27 (last month)	4,823

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries.

CLM provided participation information for 27 of the 29 months of their original intervention period.

[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/clm/9mo/tables/tableCLM-2.sas 08JUL2009

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

Characteristics	Statistic
Number of intervention months	16
Participation rate (entire demonstration period)	66%
Length of participation	
Continuous participation after engagement	56%
After initial participation, became a continuous nonparticipant	8%
Intermittent Participation	2%
Nonparticipation (never agreed)	34%
Refused to participate when contacted	24%
Not contacted/unable to be contacted	10%
Beneficiaries Participating more than 75% of days	45%
Number of participants in selected months[1]	
Month 6	5,842
Month 12	7,154
Month 16 (last month)	6,352

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries.

CLM provided participation information for 16 of the 18 months of their original intervention period.

[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/clm/9mo/tables/tableCLM-2.sas 08JUL2009

4.3.2 Characteristics of the CLM Intervention and Comparison Populations

In addition to evaluating the level of initial engagement by CLM, 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.

Because the date of randomization and the go-live date for each CMO was a month or less apart, 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 randomization 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. ***Tables 4-5 and 4-6*** provide the percent of beneficiaries by these characteristics for the intervention and comparison populations for both the original and refresh populations. Because these tables do not display any participation information, they report characteristics for all beneficiaries in the full 29 months of the original intervention period and 18 months of the refresh intervention period and the weighting is adjusted to reflect the extended time periods.

Characteristics of the CLM original population—Beneficiaries for both the intervention and comparison groups were eligible based on being in the top 5% of costs in 2004, having two or more hospitalizations, and meeting specific diagnostic criteria. The population was further restricted during a carve-out process to those beneficiaries with an HCC score of 2.75 or greater. If a beneficiary had an HCC score less than 2.75 but met specific diagnostic criteria he or she was retained. The carve-out restrictions were applied to both the intervention and comparison populations. We observe both cost and HCC score equivalency between the revised intervention and comparison groups ***(Table 4-5)***. The mean HCC score for both the intervention and comparison groups was 2.8, meaning that beneficiaries selected for the demonstration were, on average, predicted to be nearly three times more expensive than the average fee-for-service (FFS) beneficiary.

Based on beneficiary characteristics, there are a number of statistically significant albeit modest differences between the intervention and comparison populations at baseline. The intervention group had higher rates of beneficiaries who were eligible for Medicare because they were disabled, under the age of 65, or a minority. These characteristics are often proxies for poor health status. We also observe some modest variation in baseline rates of chronic conditions. The intervention group had slightly lower rates of conditions such as hypertension, chest pain, and dizziness, syncope, and convulsions. However, there is no apparent pattern in the differences in chronic conditions between the intervention and comparison groups. Out of a large number of comparisons, one would expect to find a small number of the comparisons statistically significant by chance.

Characteristics of the CLM refresh population—Beneficiaries for both the intervention and comparison groups were eligible if they had an HCC risk score greater than or equal to 2.75 and two or more hospitalizations. We observe differences in the distribution of the HCC scores, with the intervention group having a higher percentage of beneficiaries in the medium category and a lower percentage in the high category compared to the comparison group ***(Table 4-6)***. However, the mean HCC score for both the intervention and comparison groups was 3.8, meaning that beneficiaries selected for the demonstration were, on average, predicted to be nearly four times more expensive than the average FFS beneficiary.

Table 4-5
Characteristics of the CLM CMHCB demonstration program intervention and comparison populations: Original population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Total number of beneficiaries	11,516	4,561	—	—
Full time equivalent	9,153	3,623	—	—
Beneficiary characteristics				
Aged-in (vs. disabled)	87.5	89.1	-1.6	*
In Medicaid (vs. not in Medicaid)	8.9	9.4	-0.5	N/S
Male (vs. female)	46.3	49.5	-3.2	**
Urban (vs. rural)	97.6	97.9	-0.3	N/S
Age				
Mean	74.1	74.4	-0.4	N/S
<65	13.5	12.2	1.4	*
65-69	16.2	15.7	0.5	N/S
70-74	18.6	18.5	0.1	N/S
75-79	19.0	20.7	-1.7	*
80-84	16.2	17.2	-1.0	N/S
85+	16.5	15.8	0.7	N/S
Race				
White	73.2	74.0	-0.9	N/S
African American	9.0	7.9	1.1	*
Other	17.7	18.0	-0.3	N/S
Unknown	0.2	0.1	0.0	N/S
Health status				
Recalculated HCC score				
Mean	2.8	2.8	0.0	N/S
Low: > 1.35 and < 2.00	38.2	36.7	1.5	N/S
Medium: > 2.00 and < 3.10	26.4	27.3	-0.9	N/S
High: > 3.10	35.4	36.1	-0.7	N/S
Baseline PBPM low	35.4	33.6	1.8	*
Baseline PBPM medium	33.6	33.9	-0.3	N/S
Baseline PBPM high	31.0	32.5	-1.5	N/S
Charlson comorbidity index—mean	0.9	1.0	0.0	N/S

(continued)

Table 4-5 (continued)
Characteristics of the CLM CMHCB demonstration program intervention and comparison populations: Original population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Chronic conditions				
HF	27.2	27.2	0.0	N/S
Coronary artery disease	54.0	55.4	-1.3	N/S
Other respiratory disease	33.2	33.2	0.1	N/S
Diabetes without complications	33.1	33.3	-0.2	N/S
Diabetes with complications	13.1	13.4	-0.3	N/S
Essential hypertension	68.1	70.5	-2.4	**
Valve disorders	9.4	9.3	0.1	N/S
Cardiomyopathy	6.2	7.0	-0.8	N/S
Acute & chronic renal disease	18.1	18.2	-0.1	N/S
Renal failure	9.1	8.7	0.4	N/S
Peripheral vascular disease	8.7	10.0	-1.3	*
Lipid metabolism disorders	37.1	38.4	-1.3	N/S
Cardiac dysrhythmias & conduction disorders	32.7	33.6	-0.9	N/S
Dementias	3.6	3.5	0.0	N/S
Strokes	10.7	10.9	-0.2	N/S
Chest pain	20.5	22.1	-1.6	*
Urinary tract infection	23.7	23.5	0.2	N/S
Anemia	28.2	29.0	-0.8	N/S
Malaise & fatigue (including CFS)	7.1	7.3	-0.2	N/S
Dizziness, syncope, convulsions	17.1	18.5	-1.4	*
Disorders of joint	12.1	12.4	-0.3	N/S
Hypothyroidism	13.0	13.1	0.0	N/S

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; I = intervention population; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

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

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

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

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: /vol1/project/07964/025 hiccup/pgm/larsen/programs/clm/9mo/tableCLM-3x.sas 28AUG2009.

Based on beneficiary characteristics, there were no statistically significant differences between the intervention and comparison populations at baseline. We also observe some variation in baseline rates of chronic conditions. The intervention group had a slightly lower rate of diabetes without complications and slightly higher rates of urinary tract infections and malaise/fatigue. There was no apparent pattern in the differences in chronic conditions between the intervention and comparison groups.

4.3.3 Characteristics of Participants in the CLM Original and Refresh Populations

In this section, we report the beneficiary characteristics that predict participation in the CLM CMHCB demonstration program for both the original and refresh populations. First, we report the same characteristics from ***Tables 4-5 and 4-6*** by participation status (any participation during the intervention period, participation more than 75% of eligible months, and no participation) and test for differences between the any participation and no participation group (see ***Tables 4-7 and 4-8***). Within the original population ***(Table 4-7)***, in general, beneficiaries who participated were in better health than those who did not (as measured by having lower rates of Medicaid enrollment, not being aged 85 or older, and having lower baseline HCC and PBPMs). Participants were also more likely to be white compared to nonparticipants. Some noted differences by baseline chronic conditions are that participants were significantly less likely to have heart failure, renal failure, dementia, stroke, and urinary tract infections compared to nonparticipants. ***Table 4-8*** presents results for the refresh population. There were similar patterns for the beneficiary characteristics; there were lower percentages of disabled and Medicaid beneficiaries and a higher percentage of whites among the participants. At the same time, the participants had a lower HCC score indicating modestly better health status. There continued to be no pattern to the differences by baseline chronic conditions: participants had higher percentages of heart failure, lipid metabolism disorders, cardiac dysrhythmias, and hypothyroidism and lower rates of dementia and urinary tract infections compared to nonparticipants.

Next, for both the original and refresh populations we report 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 (***Table 4-9***). Within the original population, the general picture that emerges is one whereby beneficiaries who were in better health (as measured by not being disabled or aged 85 or older, staying alive for the entire intervention period, not being enrolled in Medicaid, having lower baseline PBPM costs, having lower prospective HCC scores, and not being institutionalized) tended to be more likely to participate than those who were disabled or aged 85 or older, died during the intervention period, were enrolled in Medicaid, had higher baseline PBPM costs, had higher prospective HCC scores, or were institutionalized. Additionally, minority beneficiaries and rural beneficiaries had lower rates of participation than white beneficiaries and urban beneficiaries. Institutionalized and rural beneficiaries were two groups that CLM stated that they had trouble engaging. The overall participation rate in the first 6 months was much higher for the refresh population (64%), but the overall relative picture of participation by beneficiary characteristics is similar.

Table 4-6
Characteristics of the CLM CMHCB demonstration program intervention and comparison populations: Refresh population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Total number of beneficiaries	13,104	5,240	—	—
Full time equivalent	10,812	4,339	—	—
Beneficiary characteristics				
Aged-in (vs. disabled)	70.1	69.1	1.0	N/S
In Medicaid (vs. not in Medicaid)	4.3	4.6	-0.3	N/S
Male (vs. female)	42.3	42.5	-0.1	N/S
Urban (vs. rural)	97.1	97.1	0.0	N/S
Age				
Mean	74.5	74.2	0.3	N/S
<65	11.7	12.2	-0.5	N/S
65-69	10.2	10.5	-0.3	N/S
70-74	14.4	13.9	0.4	N/S
75-79	15.4	15.2	0.2	N/S
80-84	14.0	14.6	-0.6	N/S
85+	15.1	14.2	0.9	N/S
Race				
White	60.3	60.3	0.0	N/S
African American	9.2	8.7	0.5	N/S
Other	11.1	11.5	-0.3	N/S
Unknown	19.4	19.6	-0.1	N/S
Health Status				
Recalculated HCC score				
Mean	3.8	3.8	0.0	N/S
Low: > 1.35 and < 2.00	15.6	16.3	-0.7	N/S
Medium: > 2.00 and < 3.10	26.1	22.9	3.2	**
High: > 3.10	58.3	60.8	-2.5	**
Baseline PBPM low	34.7	35.3	-0.6	N/S
Baseline PBPM medium	34.5	35.2	-0.7	N/S
Baseline PBPM high	30.9	29.6	1.3	N/S
Charlson comorbidity index—mean	1.5	1.5	0.0	N/S

(continued)

Table 4-6 (continued)
Characteristics of the CLM CMHCB demonstration program intervention and comparison populations: Refresh population

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Chronic conditions				
HF	41.0	40.6	0.4	N/S
Coronary artery disease	51.6	51.5	0.0	N/S
Other respiratory disease	40.3	40.2	0.1	N/S
Diabetes without complications	41.7	43.6	-1.8	*
Diabetes with complications	21.2	21.8	-0.6	N/S
Essential hypertension	76.2	76.5	-0.3	N/S
Valve disorders	15.0	14.9	0.1	N/S
Cardiomyopathy	13.5	13.6	-0.1	N/S
Acute & chronic renal disease	31.6	30.7	0.9	N/S
Renal failure	15.9	15.4	0.5	N/S
Peripheral vascular disease	12.2	12.7	-0.5	N/S
Lipid metabolism disorders	47.1	45.9	1.2	N/S
Cardiac dysrhythmias & conduction disorders	43.7	43.7	-0.1	N/S
Dementias	3.4	3.1	0.3	N/S
Strokes	11.4	11.3	0.1	N/S
Chest pain	21.9	21.9	-0.1	N/S
Urinary tract infection	25.5	23.7	1.9	*
Anemia	38.9	37.9	1.0	N/S
Malaise & fatigue (including CFS)	13.4	12.0	1.4	*
Dizziness, syncope, convulsions	22.5	22.6	-0.2	N/S
Disorders of joint	14.3	14.9	-0.6	N/S
Hypothyroidism	17.8	16.7	1.0	N/S

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; I = intervention population; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

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

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

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

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: /vol1/project/07964/025 hiccup/pgm/larsen/programs/clm/9mo/tableCLM-3x.sas 28AUG2009.

Table 4-7
Characteristics of the CLM CMHCB demonstration program intervention population by participation status: Original population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Total number of beneficiaries	7,289	3,817	4,227	—	—
Full time equivalent	6,524	3,817	2,750	—	—
Beneficiary characteristics					
Aged-in (vs. disabled)	87.8	88.5	86.8	1.1	N/S
In Medicaid (vs. not in Medicaid)	8.5	7.0	9.9	-1.4	*
Male (vs. female)	46.4	46.5	46.0	0.4	N/S
Urban (vs. rural)	97.7	97.9	97.2	0.5	N/S
Age					
Mean	73.8	73.8	74.7	-0.9	**
<65	13.2	12.3	14.1	-0.9	N/S
65-69	16.9	17.8	14.4	2.5	**
70-74	19.3	19.6	16.8	2.5	**
75-79	19.3	19.5	18.3	1.0	N/S
80-84	16.2	16.1	16.3	-0.1	N/S
85+	15.1	14.7	20.1	-5.0	**
Race					
White	75.1	76.7	68.6	6.4	**
African American	9.1	8.9	8.6	0.5	N/S
Other	15.6	14.1	22.7	-7.1	**
Unknown	0.2	0.2	0.1	0.1	N/S
Health status					
Recalculated HCC score					
Mean	2.8	2.7	3.0	-0.2	**
Low: > 1.35 and < 2.00	38.8	39.9	36.2	2.6	*
Medium: > 2.00 and < 3.10	27.2	26.8	24.7	2.4	*
High: > 3.10	34.1	33.3	39.1	-5.0	**
Baseline PBPM low	36.3	36.9	33.1	3.3	**
Baseline PBPM medium	33.5	33.6	33.7	-0.2	N/S
Baseline PBPM high	30.1	29.5	33.2	-3.1	**
Charlson comorbidity index—mean	0.9	0.9	1.0	0.0	N/S

(continued)

Table 4-7 (continued)
Characteristics of the CLM CMHCB demonstration program intervention population by participation status: Original population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Chronic conditions					
HF	26.6	25.2	29.2	-2.7	**
Coronary artery disease	54.6	55.9	52.3	2.4	*
Other respiratory disease	33.5	33.1	32.9	0.6	N/S
Diabetes without complications	32.7	30.5	34.0	-1.2	N/S
Diabetes with complications	13.4	13.2	12.3	1.1	N/S
Essential hypertension	67.6	66.9	69.4	-1.8	N/S
Valve disorders	9.4	9.7	9.5	-0.1	N/S
Cardiomyopathy	6.3	6.5	6.0	0.3	N/S
Acute & chronic renal disease	18.1	18.3	18.2	-0.1	N/S
Renal failure	8.6	8.3	10.5	-1.9	**
Peripheral vascular disease	9.1	8.5	7.8	1.2	*
Lipid metabolism disorders	38.5	38.4	33.2	5.3	**
Cardiac dysrhythmias & conduction disorders	32.8	33.3	32.8	0.0	N/S
Dementias	2.7	2.4	5.7	-2.9	**
Strokes	9.9	9.7	12.9	-3.0	**
Chest pain	20.6	20.6	20.2	0.4	N/S
Urinary tract infection	21.7	21.5	28.8	-7.1	**
Anemia	27.9	27.3	29.3	-1.4	N/S
Malaise & fatigue (including CFS)	7.1	7.2	7.2	-0.1	N/S
Dizziness, syncope, convulsions	16.7	16.5	18.2	-1.5	N/S
Disorders of joint	12.2	12.7	12.0	0.2	N/S
Hypothyroidism	13.3	13.9	12.3	1.0	N/S

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; P = participating; NP = nonparticipating; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

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

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

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

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/clm/9mo/tables/tableCLM-4x.sas 28AUG2009.

Table 4-8
Characteristics of the CLM CMHCB demonstration program intervention population by participation status: Refresh population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Total number of beneficiaries	8,670	3,955	4,432	–	–
Full time equivalent	7,903	3,955	3,121	–	–
Beneficiary characteristics					
Aged-in (vs. disabled)	71.7	72.6	66.4	5.2	**
In Medicaid (vs. not in Medicaid)	3.6	3.6	6.1	-2.5	**
Male (vs. female)	42.9	42.2	41.0	1.9	N/S
Urban (vs. rural)	97.4	98.6	96.2	1.2	**
Age					
Mean	74.5	74.4	74.7	-0.2	N/S
<65	11.8	12.2	11.2	0.7	N/S
65-69	10.7	11.0	8.8	1.9	**
70-74	14.8	15.5	13.1	1.6	*
75-79	15.6	15.3	14.9	0.7	N/S
80-84	14.6	14.4	12.4	2.2	**
85+	14.8	15.0	16.1	-1.3	N/S
Race					
White	63.0	64.4	53.6	9.4	**
African American	8.9	8.9	9.9	-1.0	N/S
Other	10.4	10.1	12.9	-2.5	**
Unknown	17.8	16.6	23.7	-5.9	**
Health status					
Recalculated HCC score					
Mean	3.7	3.8	3.9	-0.1	*
Low: > 1.35 and < 2.00	15.4	14.8	15.8	-0.4	N/S
Medium: > 2.00 and < 3.10	26.2	25.4	25.5	0.7	N/S
High: > 3.10	58.4	59.8	58.7	-0.3	N/S
Baseline PBPM low	33.0	31.5	38.6	-5.6	**
Baseline PBPM medium	35.8	36.0	30.8	5.0	**
Baseline PBPM high	31.2	32.6	30.6	0.5	N/S
Charlson comorbidity index—mean	1.5	1.5	1.5	0.0	N/S

(continued)

Table 4-8 (continued)
Characteristics of the CLM CMHCB demonstration program intervention population by participation status: Refresh population

Characteristics	Any participation Rate per 100[1,2]	> 75% participation Rate per 100[1,2]	Never participated Rate per 100[1,2]	P vs. NP Rate per 100[1,2]	p[3]
Chronic conditions					
HF	41.2	40.4	41.1	0.1	N/S
Coronary artery disease	52.5	51.6	49.2	3.4	**
Other respiratory disease	40.9	40.7	39.0	2.0	N/S
Diabetes without complications	41.5	40.4	42.4	-1.0	N/S
Diabetes with complications	21.0	21.3	21.9	-0.9	N/S
Essential hypertension	76.4	76.4	75.6	0.9	N/S
Valve disorders	15.4	14.7	14.0	1.4	N/S
Cardiomyopathy	13.7	13.5	13.3	0.4	N/S
Acute & chronic renal disease	31.9	31.2	31.0	0.9	N/S
Renal failure	15.6	14.8	17.2	-1.6	N/S
Peripheral vascular disease	12.3	11.9	12.1	0.2	N/S
Lipid metabolism disorders	48.5	48.5	42.9	5.6	**
Cardiac dysrhythmias & conduction disorders	44.5	45.5	41.7	2.8	*
Dementias	2.7	2.5	5.3	-2.6	**
Strokes	11.3	11.4	11.9	-0.5	N/S
Chest pain	22.4	23.2	20.2	2.2	*
Urinary tract infection	24.4	24.6	29.0	-4.6	**
Anemia	38.8	38.3	39.6	-0.8	N/S
Malaise & fatigue (including CFS)	13.4	13.5	13.6	-0.2	N/S
Dizziness, syncope, convulsions	22.1	22.3	23.3	-1.2	N/S
Disorders of joint	14.6	14.1	13.2	1.4	N/S
Hypothyroidism	18.3	18.7	16.4	1.8	*

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; P = participating; NP = nonparticipating; C = comparison population; HCC = Hierarchical Condition Category; HF = heart failure; CFS = chronic fatigue syndrome.

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

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

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

N/S means not statistically significant.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: H:/project/07964/025 hiccup/pgm/larsen/programs/clm/9mo/tables/tableCLM-4x.sas 28AUG2009.

Table 4-9
Participation rates during the first 6 months of the CLM CMHCB demonstration by beneficiary characteristics, baseline characteristics, and intervention period health status: Original and refresh populations

Characteristics	Original (%)	Refresh (%)
Overall participation rate[1,2]	40	64
Beneficiary characteristics		
Male	41	64
Female	40	64
White	42	67
African American/other/unknown	36	57
Age < 65 years	38	66
Age 65-74	43	66
Age 75-84	41	64
Age 85 + years	36	60
Medicaid	35	52
Non-Medicaid	41	65
Urban	41	64
Rural	38	56
Baseline characteristics		
Baseline HCC score low	42	60
Baseline HCC score high	38	65
Low baseline PBPM	43	60
High baseline PBPM	38	63
Baseline Charlson score low	42	61
Baseline Charlson score medium	40	66
Baseline Charlson score high	39	64
Demonstration period health status		
Died	34	51
Alive	43	68
Institutionalized	14	30
Not institutionalized	44	66
Concurrent HCC score low	39	63
Concurrent HCC score high	39	60
Number of participants	4,406	7,828
Number of total beneficiaries	10,883	12,273

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

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

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

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: /vol1/project/07964/025 hiccup/pgm/larsen/programs/clm/9mo/partab2.sas 17AUG2009.

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

Model 1: Beneficiaries who participated at least 1 month in the first 6 months of the intervention period compared with all other beneficiaries (nonparticipants);

Model 2: Beneficiaries who participated at least 1 month during the full intervention period compared with all other beneficiaries (nonparticipants); and

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 and those who participated in the CMHCB demonstration more than 75% of their eligible months. Model 1 reflects the initial recruitment emphasis by the CMO, or characteristics of beneficiaries with whom the CMO 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 CMHCB demonstration. For Models 1 and 3, two versions were estimated; a version with just exogenous beneficiary characteristics and a full model. 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 4-10 and 4-11 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. An odds ratio less than 1 means that beneficiaries with a particular characteristic were less likely to participate; an odds ratio greater than 1 means that beneficiaries with the particular characteristic were more likely to participate. In general, the reference group comprises characteristics associated with younger and healthier beneficiaries. Across all three models, the explanatory power of the studied beneficiary characteristics was extremely low. This suggests that CLM had cast a fairly wide net when engaging their intervention population. Pseudo R-squares for most of the models were between 0.04 and 0.06, with the full Model 3 exhibiting pseudo R-squares of 0.04 for the original population and 0.03 for the refresh population. ***Supplement 4A*** contains tables that present the odds ratios and level of significance for Models 1 and 2.

Table 4-10
Logistic regression modeling results comparing beneficiaries that participated at least 75% of eligible months during the CLM CMHCB intervention period to all other intervention beneficiaries: Original population[1,2]

Characteristics	Model 1 OR	p^3	Model 2 OR	p^3
Intercept	0.58	**	0.57	**
Beneficiary characteristics				
Male	0.98	N/S	1.02	N/S
African American/other/unknown	0.77	**	0.79	**
Age < 65 years	0.79	**	0.83	**
Age 75-84	0.91	N/S	0.95	N/S
Age 85 + years	0.72	**	0.92	N/S
Medicaid	0.72	**	0.72	**
Urban	1.36	*	1.41	*
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.03	N/S
Baseline HCC score high	N/I	N/I	1.05	N/S
Medium baseline PBPM	N/I	N/I	0.96	N/S
High baseline PBPM	N/I	N/I	0.96	N/S
Baseline Charlson score medium	N/I	N/I	1.00	N/S
Baseline Charlson score high	N/I	N/I	0.87	*
Demonstration period health status				
Died	N/I	N/I	0.53	**
Institutionalized	N/I	N/I	0.22	**
Concurrent HCC score medium	N/I	N/I	1.24	**
Concurrent HCC score high	N/I	N/I	1.24	**
Number of cases	11,516	N/A	11,516	N/S
Chi-square (p<)	89.17	**	503.16	**
Pseudo R-square	0.01	N/A	0.04	N/S

NOTES: CLM = Care Level Management; 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 LT $1,671. The concurrent HCC score reference group is .616 or less.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: partab3 17AUG2009, partab4b 29JUL2009

Table 4-11
Logistic regression modeling results comparing beneficiaries that participated at least 75% of eligible months during the CLM CMHCB intervention period to all other intervention beneficiaries: Refresh population[1,2]

Characteristics	Model 1 OR	p^3	Model 2 OR	p^3
Intercept	0.24	**	0.19	**
Beneficiary characteristics				
Male	0.90	*	0.87	**
African American/other/unknown	0.81	**	0.85	**
Age < 65 years	0.97	N/S	0.92	N/S
Age 75-84	0.86	**	0.91	N/S
Age 85 + years	0.79	**	0.91	N/S
Medicaid	0.74	**	0.80	**
Urban	2.77	**	2.76	**
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.10	N/S
Baseline HCC score high	N/I	N/I	1.19	*
Medium baseline PBPM	N/I	N/I	1.11	N/S
High baseline PBPM	N/I	N/I	1.22	*
Baseline Charlson score medium	N/I	N/I	1.03	N/S
Baseline Charlson score high	N/I	N/I	0.97	N/S
Demonstration period health status				
Died	N/I	N/I	0.61	**
Institutionalized	N/I	N/I	0.13	**
Concurrent HCC score medium	N/I	N/I	1.18	**
Concurrent HCC score high	N/I	N/I	1.14	*
Number of cases	13,102	N/A	13,102	N/S
Chi-square (p<)	108.89	**	463.31	**
Pseudo R-square	0.01	N/A	0.03	N/S

NOTES: CLM = Care Level Management; 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 LT $780.40. The concurrent HCC score reference group is .886 or less.

Data Sources: RTI analysis of 2005-2008 Medicare enrollment, eligibility, claims and encounter data.

Program: partab3 17AUG2009, partab4b 29JUL2009

Based upon descriptive statistics of participation status, it is not surprising to see that having minority status, being under age 65, being 85 years of age and older, and having Medicaid status were negative predictors of participation. Not surprisingly, participants were more likely to live in an urban area (***Table 4-10***). Examining the full Model 3 for the original population ***(Table 4-10)***, we continue to observe the same pattern of influence of beneficiary characteristics on the likelihood of participation with one exception: the introduction of baseline and demonstration period health status measures negates the influence of age 85 and over on participation status. This implies correlation among these variables. Beneficiaries who were institutionalized during the first 6-month period of the demonstration were 80% less likely to participate than those not institutionalized, holding other factors constant. CLM had reported challenges engaging this population. Beneficiaries who died during the intervention period were less likely to be participants. However, beneficiaries with medium and high concurrent HCC scores were more likely to be participants. Most baseline health status characteristics (e.g., HCC risk score, PBPM costs, and comorbidity indices) had no impact on the likelihood of participation when controlling for baseline demographics and demonstration period health status.

There are a few noted differences in the results for the refresh population ***(Table 4-11)*** such as age less than 65 having no impact on the likelihood of participation and males being less likely to participate. Further, a high baseline HCC score and a high PBPM cost increased the likelihood of participation among the refresh population. These differences suggest that CLM was more successful enrolling the sicker and more costly beneficiaries into their CMHCB program with the refresh population.

4.3.4 Level of Intervention

In this section, we report the frequency of interaction between CLM and intervention beneficiaries for a subset of original intervention population beneficiaries who were fully eligible and participating in the CLM CMHCB demonstration program from month 7 through month 18. 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. During the site visit, CLM stated that it targeted beneficiaries based upon perceived need for services due to clinical deterioration or risk of hospitalization. Thus, we expect to see a pattern of higher levels of intervention contacts for beneficiaries in poorer health status or higher users of hospitalization services.

Descriptive statistics were performed on beneficiaries participating in the CLM demonstration program to determine the breadth and depth of contacts related to care management. The data represent beneficiaries who were fully eligible and participating (unless they died) for months 6 through 18 of the demonstration. A total of 4,041 unique beneficiaries met these criteria. Observations were weighted by the fraction of eligible days, accounting for fewer contacts due to attrition because of death, which resulted in 3,817 full-time equivalent beneficiaries. ***Table 4-12*** displays the overall distribution of care management related contacts. The mean number of contacts for each beneficiary was 31 and the median was 21. One-third of beneficiaries had less than 12 contacts and one-third of beneficiaries had 33 or more contacts.

Table 4-12
Distribution of number of contacts with participants in the CLM CMHCB demonstration: Original intervention population

Statistic	Number	Percent
Number of beneficiaries[1]	4,041	—
FTE beneficiaries[2]	3,817	—
Mean number of contacts	31	—
Median number of contacts	21	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-11 contacts	1,293	33.9%
12-32 contacts	1,240	32.5%
33+ contacts	1,284	33.6%
Total	3,817	100.0%

NOTES: CLM = Care Level Management; FTE = full time equivalent.

[1] Beneficiaries had to be fully eligible and full participants in months 7-18.

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

SOURCE: \\rtidcs01\hserprojdc\07964 HICUP\deliverables\CLM and TT final report\TTEncounter

Table 4-13 displays the percent of participants with completed Physician Home Intervention (PHI) visits, total visits that include both PHI and non-PHI visits made by physicians, telephone contacts with physicians or nurses, and total contacts by frequency of contact over 18 months. Two-thirds of beneficiaries had no PHI visits. Fourteen percent of beneficiaries had one PHI visit and another 12% had 2 to 4 PHI visits during the 12-month period.

Physicians also provided services to participating beneficiaries in their homes ranging from a physician visit in response to a call from a patient with a highly complex medical problem to a regularly scheduled visit for routine follow-up with no recommended change in treatment. When these types of visits are combined with PHI visits, 75% of beneficiaries had one or more physician visits. Twenty-two percent of beneficiaries had 10 or more visits, and 14% had 20 or more visits. Eighty-eight percent of beneficiaries received a telephone call from a nurse or physician, while 24% received 10 or more calls, and 39% of beneficiaries received 20 or more calls. Combining telephone and visit contacts, we observe that 11% of fully eligible and participating beneficiaries had none for the 12-month period and another 19% had fewer than 10 contacts. Yet at the same time, we observe 52% of beneficiaries had 20 or more contacts with the majority being telephone contacts.

Table 4-13
Percent distribution of participants with CLM physician home interventions, total visits, and telephone calls: Original intervention population

Type and frequency of contact	Number of FTE beneficiaries[1,2]	Percent
Physician home interventions		
0	2,655	69.5%
1	543	14.2%
2-4	471	12.4%
5-9	129	3.4%
≥10	17	0.4%
20+	3	0.1%
Total visits (visit service type + physician home visits)		
0	971	25.4%
1	347	9.1%
2-4	474	12.4%
5-9	646	16.9%
≥10	846	22.0%
20+	534	14.0%
Telephone contacts		
0	459	12.0%
1	181	4.8%
2-4	321	8.4%
5-9	470	12.3%
≥10	903	23.7%
20+	1,484	38.9%
Total all contacts		
0	457	11.9%
1	176	4.6%
2-4	262	6.9%
5-9	277	7.3%
≥10	669	17.5%
20+	1,977	51.8%

NOTES: CLM = Care Level Management; FTE = full time equivalent.

[1] Beneficiaries had to be fully eligible and full participants in months 4-15

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

SOURCE: S:\07964 HICUP\deliverables\CLM and TT final report\CLM\Encounter\

Table 4-14 displays the frequency of care management contacts by baseline HCC score and type of contact. Contact by mode was not mutually exclusive in that a beneficiary could have a combination of telephone and visit contacts any time during the demonstration. Beneficiaries were stratified into three HCC categories ranging from an HCC score greater than 3.0 to less than 2.0.

Physician home intervention—Beneficiaries in the high HCC risk group (34%) were 13 percentage points more likely to have had 1 or more PHI visits as beneficiaries in the low HCC risk group. High risk beneficiaries were 2.3 times more likely to have had 5 or more PHI visits compared with the low risk group (5.4% versus 2.4%). When PHI visits and other types of visits are combined, 79% of high risk beneficiaries had 1 or more visits compared to 70% of low risk beneficiaries. High risk beneficiaries were 1.6 times more likely than low risk beneficiaries to have had 10 or more physician visits (44% versus 28%). These findings suggest that CLM made a focused effort to visit their higher acuity beneficiaries.

Telephone contacts—Beneficiaries in both the high risk group (90%) and low risk group (87%) were almost equally likely to have had 1 or more telephone contacts with a physician or nurse care coordinator. However, beneficiaries in the high risk group were 1.5 times more likely than low risk beneficiaries to have 20 or more contacts (47% versus 31%). Conversely, beneficiaries in the low risk group were more likely to have 9 or fewer telephone contacts than beneficiaries in the high risk group (31% versus 21%).

When all modes of contact are combined, 90% of beneficiaries in the high risk group had one or more contacts as compared to 87% for beneficiaries in the low risk group. However, 60% of high risk beneficiaries had twenty or more contacts as compared to 43% for low risk category beneficiaries.

To more directly examine the targeting strategy of CLM, a multivariate logistic regression model was estimated with the number of total contacts as the dependent variable. The model estimates the likelihood of a participant receiving a high number of contacts. The medium contact group was omitted, thus comparing the high contact group to the low contact group. ***Table 4-15*** displays 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 months 7 through 18 of the demonstration, and their number of contacts categorized either as low (0-2) or high (5+). 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 2.6 times greater than those for a beneficiary age 65 and older, adjusting for any baseline difference in HCC score and characteristics. Conversely, male beneficiaries were 24% less likely to be in the high contact category. Being age 85 or older also increased the likelihood of being in the high contact category by 44%.

Baseline health status measured by the HCC risk score and the Charlson index score were created by RTI using the baseline claims for beneficiaries 1 year prior to the go-live date. We also estimated average PBPM Medicare payments during the baseline year. The presence of medium and high levels of baseline severity as measured by the HCC score and a high baseline level of costliness as measured by PBPM costs were positive predictors of being in the high

Table 4-14
Frequency of CLM contacts by HCC score for full participants during demonstration months 7-18: Original intervention population

Contact mode	HCC Score High (>3) N = 1294 Frequency	%	HCC Score Medium (2-3) N = 1033 Frequency	%	HCC Score Low (<2) N = 1490 Frequency	%
Visits (PHI)						
0	812	62.7%	704	68.2%	1,139	76.4%
1	214	16.5%	150	14.5%	179	12.0%
2-4	197	15.2%	137	13.3%	137	9.2%
5-9	62	4.8%	36	3.5%	31	2.1%
10-19	10	0.7%	4	0.4%	3	0.2%
20+	1	0.08%	1	0.10%	1	0.1%
Total visits (service type 1-7 and PHI type 1-4)						
0	271	21%	248	24%	451	30.3%
1	96	7.4%	86	8.3%	165	11.1%
2-4	128	9.9%	135	13.1%	210	14.1%
5-9	232	17.9%	169	16.4%	245	16.4%
10-19	320	24.7%	249	24.0%	278	18.7%
20+	248	19.1%	145	14.1%	140	9.4%
Telephone contacts (service type 10-12)						
0	127	9.8%	133	12.9%	199	13.3%
1	51	3.9%	42	4.1%	89	6.0%
2-4	92	7.1%	72	7.0%	156	10.5%
5-9	129	9.9%	119	11.5%	222	14.9%
10-19	284	21.9%	260	25.1%	360	24.2%
20+	613	47.4%	407	39.4%	464	31.1%
Total all contacts (total visits + telephone)						
0	126	9.7%	132	12.8%	199	13.3%
1	49	3.8%	42	4.0%	86	5.7%
2-4	79	6.1%	58	5.6%	125	8.4%
5-9	74	5.7%	66	6.4%	137	9.2%
10-19	186	14.4%	175	16.9%	307	20.6%
20+	780	60.3%	560	54.3%	636	42.8%

NOTES: CLM = Care Level Management; HCC =Hierarchical Condition Category; N = number of beneficiaries; PHI = physician home intervention.

SOURCE: S:\07964 HICUP\deliverables\CLM and TT final report\CLM\Encounter\

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

Characteristics	Odds ratio[1,2]	p<
Demographics		
Male	0.76	**
African American/other/unknown	0.98	N/S
Age <65	2.59	**
Age 75-84	1.00	N/S
Age 85+ years	1.44	**
Urban	1.57	N/S
Medicaid beneficiary	0.99	N/S
Baseline health status & utilization		
Baseline HCC score medium (2.00 - 3.1)	1.39	**
Baseline HCC score high (>3.1)	1.52	**
Baseline Charlson score medium	1.14	N/S
Baseline Charlson score low	1.17	N/S
One hospitalization	1.36	**
Multiple hospitalizations	2.07	**
Baseline Medicare payments		
Medium base PBPM	1.09	N/S
High base PBPM	1.31	*
Demonstration health status		
Died	1.43	**
NH/LTC/SNF resident	0.41	**
Concurrent HCC score medium (0.315-1.121)	1.22	N/S
Concurrent HCC score high (>1.121)	1.45	**
Number of cases	2,726	N/S
Chi-square (p<)	299.30	**
Pseudo R2	0.10	N/S

NOTES: CLM = Care Level Management; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

NOTES: CLM = Care Level Management; FTE = full time equivalent.

[1] Beneficiaries had to be fully eligible and full participants in months 4-15

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

* denotes statistical significance at the 5% level.

** denotes statistical significance at the 1% level.

N/S means not statistically significant.

SOURCE: RTI analysis of Medicare enrollment and claims and encounter data.

S:\07964 HICUP\deliverables\CLM and TT final report\CLM\Encounter\interrun02

contact group. Demonstration period acute care utilization was also a powerful explainer of more contacts. Beneficiaries who had multiple hospitalizations were 2 times more likely to be in the high contact group (2.07), and those who had no hospitalizations. A high concurrent HCC score, or health status measured during the first 6 months of the demonstration period, was also a positive predictor of being in the high contact group. Beneficiaries who died during the demonstration or who were institutionalized at the start of the demonstration were less likely to be in the high contact category. These findings suggest that CLM made a focused effort to contact beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program.

4.4 Summary

For CLM, we find that participants from the original population were healthier and younger than beneficiaries who never participated. The disabled (under age 65), the very old (85 years of age and older), Medicaid enrollees, institutionalized beneficiaries, and those who died during the demonstration were less likely to be participants. In the multivariate regression analysis, most baseline health status characteristics (e.g., prospective HCC risk score, PBPM costs, and Charlson comorbidity indices) had no impact on the likelihood of participation after controlling for baseline demographics and demonstration period health status. However, beneficiaries with medium and high concurrent HCC scores were more likely to be participants. This suggests that CLM was unable to engage the historically sicker Medicare beneficiaries but did make some inroads into engaging those with acute clinical deterioration as measured by the concurrent HCC score. Further, a high baseline HCC score and a high PBPM cost increased the likelihood of participation among the refresh population. These differences suggest that CLM was more successful engaging the sicker and more costly beneficiaries in their program as it matured.

A cornerstone of CLM's program was physician home intervention visits including home hospitalization services. Yet, only 30% of their fully participating beneficiaries for a 12-month period received a PHI visit. When we add in other types of visits—including routine follow-up care—we do observe a higher level of physician interaction. However, 25% of beneficiaries received no physician visits and one-half of all beneficiaries received less than five visits during a 12-month period. Telephone contact was the most dominant "frequent" form of contact. In our multivariate regression modeling of likelihood of being in a high contact versus low contact group, we found some evidence that CLM made a focused effort to contact beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program. Given CLM's high monthly management fee (almost $300 per month) and the population-based financial risk feature of this demonstration, the concentration of physician visits, in general, and PHI visits, in particular, suggests that CLM would have to be extremely successful in reducing costs associated with the beneficiaries they were targeting. Prior evaluations of Medicare care management programs that were primarily telephonic have not demonstrated savings sufficient to cover fees one-third the size of CLM's fee.

CHAPTER 5
CLINICAL QUALITY PERFORMANCE

5.1 Introduction

RTI's analysis of quality of care focuses on measuring effectiveness of the Care Management for High Cost Beneficiaries (CMHCB) demonstration interventions by answering the following evaluation question:

- *Clinical Quality of Care:* Did Care Level Management's (CLM's) CMHCB demonstration 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 CLM CMHCB demonstration by examining changes in the rate of receipt of four 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 four measures appropriate for different populations of elderly beneficiaries: influenza vaccine for all beneficiaries; low-density lipoprotein cholesterol (LDL-C) testing for beneficiaries with diabetes or ischemic vascular disease (IVD); annual oxygen saturation assessment for beneficiaries with chronic obstructive pulmonary disease (COPD); 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. Therefore, in our evaluation, we selected measures that could be reliably calculated using Medicare administrative data to assess improvements in quality of care and health outcomes. Further, these data are available for both the intervention and comparison populations and do not require medical record abstraction or beneficiary self-report. Medical record data are not available to us for either the intervention or comparison populations, and beneficiary self-report data would only available for the intervention beneficiaries who participated during the demonstration. Further, beneficiary self-report is subject to recall error and to the willingness of beneficiaries to provide the information.

5.2 Methodology

We created process-of-care measures for the 12-month period immediately prior to CLM's go-live date and for two intervention periods (months 6-17 and months 18-29) for its original population and for one intervention period (months 7-18, 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 this analysis. ***Table 5-1*** provides the number of beneficiaries who were included in the analyses of the quality of care measures, in total, and by three disease cohorts: COPD, diabetes, and IVD.

For the financial and acute care utilization analyses, claims data were only included during periods that the beneficiary met the eligibility criteria for the CMHCB demonstration in both the baseline and intervention periods. Therefore, costs incurred while beneficiaries were not eligible for the demonstration, for example, after electing the hospice benefit, were excluded. As

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

Statistics	All	COPD	Diabetes	Ischemic vascular disease
Original beneficiaries				
Months 6-17				
Intervention				
Total number of beneficiaries	10,480	3,053	3,897	6,173
Full time equivalents[1]	10,446	3,041	3,888	6,156
Comparison				
Total number of beneficiaries	4,163	1,237	1,551	2,549
Full time equivalents[1]	4,147	1,232	1,544	2,539
Months 18-29				
Intervention				
Total number of beneficiaries	8,762	2,384	3,223	5,223
Full time equivalents[1]	8,740	2,377	3,215	5,210
Comparison				
Total number of beneficiaries	3,458	960	1,279	2,133
Full time equivalents[1]	3,446	957	1,273	2,125
Refresh beneficiaries				
Months 7-18				
Intervention				
Total number of beneficiaries	11,431	3,393	4,199	5,384
Full time equivalents[1]	11,357	3,373	4,171	5,249
Comparison				
Total number of beneficiaries	4,587	1,342	1,751	2,170
Full time equivalents[1]	4,558	1,336	1,737	2,157

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; COPD = chronic obstructive pulmonary disease.

[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.

described in Chapter 2, we elected a somewhat different approach for the quality of care measures. Medicare claims for the full baseline and intervention period were included regardless of beneficiary eligibility for the CMHCB 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 care management organization (CMO) in case the services occurred after exposure to the CMHCB 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 CMHCB demonstration—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, as described in Equation 2.1. 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 CMHCB demonstration was based on two nested cohort samples of Medicare beneficiaries who were assigned to intervention and comparison groups within five strata defined by baseline costs. In addition, an eligibility fraction ranging from 0 to 1 was assigned to the pre- and post- time periods for each sample member. STATA SVY was used to fit the model with robust variance estimation. Operationally, the five strata and a beneficiary identifier were included in the SVYSET statement to reflect the stratified sampling design. The eligibility fraction was included as the weight to reflect the period of time during which the beneficiary met the CLM CMHCB demonstration eligibility criteria in the baseline and demonstration periods.

As described in Equation 2.1, the β_3 interaction coefficient tests whether the intervention group's performance profile differs over time from the comparison group's performance profile. 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 β_3 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 CLM CMHCB demonstration original and refresh beneficiaries into four categories based upon whether or not they received each of the four quality of care measures during the pre-demonstration baseline period and the last 12 months of the demonstration: compliant in both baseline and demonstration; compliant in baseline but not in demonstration; not compliant in baseline but compliant in demonstration; and not compliant in both periods. We report on the natural trends observed in the comparison and intervention populations over the 3-year period.[10] 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 CLM CMHCB demonstration 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 rate (baseline period intervention versus comparison rate difference minus intervention period intervention versus comparison rate difference). Positive difference-in-differences rates per 100 beneficiaries indicate that the intervention group's mean rate improved more than the comparison group's mean rate or the intervention group's mean rate declined at a lower rate than the comparison group's mean rate. Negative difference-in-differences rates per 100 beneficiaries indicate that comparison group exhibited higher rates of growth or less of a decline than the intervention group.

At baseline, rates for the five measures in the original comparison group ranged from a low of 25% for oxygen saturation for beneficiaries with COPD to a high of 74% for HbA1c testing for beneficiaries with diabetes. With the exception of influenza vaccine, rates in the comparison group either remained the same or declined over the course of the 29-month demonstration period. We observe a more than 10 percentage point increase in rate of influenza vaccine and a nearly 10 percentage point decrease in oxygen saturation assessment for beneficiaries with COPD. As time elapsed during the demonstration, the rate of compliance with HbA1c testing among the original comparison beneficiaries with diabetes declined.

Between baseline and months 6-17 of CLM's original demonstration period, there was no statistically significant quality improvement across the five measures. However, during the last 12 months of the demonstration period, the rate of influenza vaccine increased in the original intervention population by 3 percentage points more than the comparison group. And, we observe a similar increase in the refresh intervention population. There were no improvements in rates of oxygen saturation assessment among the original and refresh beneficiaries with COPD, rates of HbA1c or LDL-C testing among the original and refresh intervention beneficiaries with diabetes, or rates of LDL-C testing among the original and refresh intervention beneficiaries with ischemic vascular disease.

---

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

Table 5-2
Comparison of rates of guideline concordant care for the first and last 12 months of the CLM demonstration period with rates for a 1-year period prior to the start of the CLM 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-inD Rate per 100	D-in-D OR	D-in-D p	D-in-D CI Low	D-in-D CI High
ORIGINAL POPULATION									
Months 6-17									
All beneficiaries									
Influenza vaccine	28	28	42	41	0.92	1.04	0.45	0.93	1.17
Beneficiaries with COPD									
Oxygen saturation test	25	25	18	18	-0.00	1.00	0.99	0.79	1.27
Beneficiaries with diabetes									
HbA1c test	75	73	73	72	-0.30	0.98	0.86	0.81	1.19
LDL-C test	67	68	67	67	1.42	1.07	0.49	0.89	1.28
Beneficiaries with IVD[2]									
LDL-C test	71	68	69	67	-0.67	0.97	0.65	0.84	1.12
Months 18- 29									
All beneficiaries									
Influenza vaccine	28	28	42	40	3.30	1.15	0.02	1.02	1.30
Beneficiaries with COPD									
Oxygen saturation test	25	25	17	16	0.08	1.02	0.90	0.77	1.34
Beneficiaries with diabetes									
HbA1c test	75	74	68	69	-1.83	0.91	0.40	0.74	1.13
LDL-C test	70	70	68	71	-1.89	0.92	0.40	0.75	1.12
Beneficiaries with IVD[2]									
LDL-C test	73	71	70	70	-2.34	0.89	0.16	0.76	1.05
REFRESH POPULATION									
Months 7-18									
All beneficiaries									
Influenza vaccine	33	33	48	45	3.36	1.15	0.01	1.03	1.27
Beneficiaries with COPD									
Oxygen saturation test	24	24	18	19	-0.40	0.97	0.77	0.77	1.22
Beneficiaries with diabetes									
HbA1c test	76	77	70	70	-0.39	0.98	.096	0.82	1.18
LDL-C test	72	71	68	70	-2.86	0.87	0.14	0.73	1.04
Beneficiaries with IVD[2]									
LDL-C test	70	69	66	66	-1.18	0.95	0.49	0.81	1.11

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries; I = intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odds ratio; COPD = chronic obstructive pulmonary disease; 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 care management organization (CMO) was active in the program. Only beneficiaries who had at least one day of eligibility in both the baseline and demonstration periods are included in this analysis.

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

SOURCE: RTI analysis of 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: qcvars; gcc_tab1.

Table 5-3 displays the percentages of CLM's CMHCB demonstration original and refresh beneficiaries who did or did not receive one of the four process-of-care measures during the baseline and last 12 months of its respective demonstration period. We display the distribution of its 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 either the baseline or intervention period; and
- became compliant, meaning noncompliant in the baseline period but compliant in the intervention period.

The first column for each quality of care measure contains the percentage distributions for the comparison populations and the second column displays the percentage distributions for the intervention populations. Only about one-half of original and refresh intervention and comparison beneficiaries with diabetes or IVD were compliant in both baseline and demonstration periods for HbA1c and LDL-C testing. Less than 10% of beneficiaries with COPD were compliant in both time periods for oxygen assessment, and 20% or fewer beneficiaries were compliant in both time periods for influenza vaccination. Although we observe a statistically significant increase in rate of compliance in influenza vaccine in both the original and refresh populations, over one-half of all intervention beneficiaries were not compliant during the demonstration period. This is surprising given that CLM's demonstration program focused on some of the most critically ill Medicare fee-for-service (FFS) beneficiaries. Across the five measures original and refresh beneficiaries in both groups were more likely to become noncompliant rather than compliant during the course of the demonstration with the exception of influenza vaccination.

Of particular note is the low rate of compliance with oxygen saturation assessment; more than 80% of intervention beneficiaries were not compliant during the intervention period. Annual oxygen saturation is a National Quality Forum–endorsed quality of care measure that is owned by the American Medical Association's Physician Consortium of Performance Improvement (AMA-PCPI). In its specifications for this measure, the relevant population is restricted to patients aged 18 years of age and older with a diagnosis of COPD. The specifications further restrict the population to those patients with a functional expiratory volume of less than 40% of predicted value. However, we were unable to impose that eligibility restriction with only claims data available to us. On one hand, the low rates of adherence we observed could, in part, reflect a denominator population that is larger or that includes beneficiaries with a higher level of pulmonary function than the population envisioned by the AMA-PCPI. On the other hand, it is difficult to imagine that nearly 80% of an ill FFS population with COPD would not qualify for this measure. The low rates of adherence could also be driven by the use of only Current Procedural Terminology (CPT) procedure codes for identification of the test being conducted. The specifications also allow for identification through laboratory data, It may be that this

measure is not well suited for measurement with Medicare claims data alone and use of CPT procedure codes only.

Table 5-3
Percentage of comparison and intervention beneficiaries meeting process of care standards in the baseline year and last 12 months of the CLM 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	Oxygen assessment	Oxygen assessment	Influenza vaccine	Influenza vaccine
Original population	C	I	C	I	C	I	C	I	C	I
Always compliant	57%	57%	55%	54%	57%	58%	8%	8%	16	17
Became noncompliant	18	18	16	17	15	16	17	17	12	11
Never compliant	13	14	13	15	15	14	67	65	48	46
Became compliant	12	11	15	14	13	12	9	9	23	26
Refresh population	C	I	C	I	C	I	C	I	C	I
Always compliant	61	61	56	56	53	54	8	8	21	22
Became noncompliant	16	16	15	16	17	17	16	16	12	11
Never compliant	13	15	15	16	17	17	65	66	42	41
Became compliant	9	9	13	12	13	12	11	10	25	26

NOTES: CMHCB = Medicare Care Management for High Cost Beneficiaries; CLM = Care Level Management; 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 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: qcvars, gcctab3.sas.

5.4 Summary of Findings and Conclusion

In this chapter, we report on RTI's assessment of the effect of the CMHCB program on quality of care. Specifically, we report findings for the key research question: did CLM improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care? We find no evidence of systematic improvement in quality of care in the CLM CMHCB demonstration program. Out of five measures, there was only one observed increase in rate of receipt of evidence-based care (influenza vaccination). We observe this increase in both the original and refresh population and during the last 12 months of CLM's operations.

Over the course of the demonstration, CLM had expected to increase rates of adherence to evidence-based care. However, during the last year of its demonstration program, we observe lower rates of adherence to the selected measures among its intervention beneficiaries than we do

among the comparison group beneficiaries for all measures with the exception of influenza vaccination. We also observe between roughly one-third to one-half of intervention beneficiaries in both the original and refresh populations were not compliant during the last year of the CMHCB demonstration despite focused efforts by CLM to encourage beneficiaries to become compliant with evidence-based care. As noted above, more than 80% of intervention beneficiaries with COPD were not compliant with annual oxygen saturation assessment. These findings suggest that improving or sustaining adherence to guideline concordant care in a cohort of ill Medicare FFS beneficiaries was more challenging than originally envisioned.

CHAPTER 6
HEALTH OUTCOMES

6.1 Introduction

RTI's analysis of health outcomes focuses on measuring effectiveness of the Care Level Management (CLM) Medicare Care Management for High Cost Beneficiaries (CMHCB) demonstration program by answering the following two evaluation questions:

- Did the CLM program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and emergency room (ER) utilization?
- Did the CLM 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 6-17 and the last 12 months of the CLM 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 CLM demonstration program date, for months 6-17 for the original population, and the last 12 months of the intervention period for both the original and refresh populations. We constructed rates of all-cause hospitalization and ER visits and a combined utilization measure for 10 ambulatory care sensitive condition (ACSC) reasons for admission—heart failure, diabetes, asthma, cellulitis, chronic obstructive pulmonary disease (COPD) and chronic bronchitis, dehydration, bacterial pneumonia, septicemia, ischemic stroke, and urinary tract infection (UTI)—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 CLM 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. In Equation 2.1, we report the IRR associated with the β_3 interaction term, or 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 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 CLM's program. For the original population's first demonstration period, we included index admissions for months 3 through 14, and for the second demonstration period, we included index admissions for months 15-26. For the refresh population's demonstration period, we included index admissions for months 4-15. As described in Chapter 2, we counted readmissions that occurred within 90 days after an index hospitalization discharge date. Therefore, readmissions for baseline period admissions were counted through the first 3 months of the demonstration period. Demonstration period readmissions were counted through the end of the demonstration period.

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 was 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 between the groups. If CLM reduced admissions and readmissions, we expect to observe negative D-in-D, 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. The test of the significance of the D-in-D estimate is based on the β_3 interaction term in Equation 2.1. We report the odds ratio (OR) from the logistic regressions and the IRR from the negative binomial regressions, 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, indicating that CLM 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.

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 CLM 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 periods. 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. The last four columns contain the IRR and its statistical level of significance (*p*) value as well as the 95% confidence interval for the IRR.

Table 6-1
Comparison of rates of utilization for months 6-17 and the last 12 months of the CLM CMHCB demonstration with rates of utilization for a 1-year period prior to the start of the CLM 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 6-17									
Hospitalizations									
All cause	1,847	1,913	1,155	1,229	-7	0.97	0.43	0.91	1.04
10 ACSCs[5]	507	516	403	437	-25	0.94	0.27	0.84	1.05
ED/Obs visits									
All cause	905	929	801	882	-56	0.93	0.50	0.76	1.14
10 ACSCs	151	152	128	143	-13	0.91	0.41	0.71	1.15
Months 18-29									
Hospitalizations									
All cause	1,781	1,828	1,099	1,199	-53	0.94	0.11	0.87	1.01
10 ACSCs	460	451	398	456	-67	0.86	0.01	0.76	0.97
ED/Obs visits									
All cause	890	898	780	897	-110	0.88	0.29	0.69	1.12
10 ACSCs	141	145	124	144	-16	0.89	0.41	0.66	1.18

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ED/Obs = emergency room visits, including observation bed stays; CMO = care management organization; COPD = chronic obstructive pulmonary disease; UTI = urinary tract infection.

[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 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 acsctab acsc acsctab1 26JUN2009.

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.

Not unexpectedly, the baseline rates of hospitalization and ER visits were very high in the CLM intervention and comparison populations. The baseline rate of all-cause hospitalization was 1,847 per 1,000 original intervention group beneficiaries. And, the baseline rate of all-cause ER visits was 905 per 1,000 original intervention beneficiaries. Original population beneficiaries eligible for the later months of the demonstration had modestly lower baseline utilization rates. The 10 ACSC reasons for hospitalization combined accounted for roughly one-quarter of all-cause hospitalizations and all-cause ER visits. Thus, Medicare fee-for-service (FFS) beneficiaries in the CLM demonstration program were being treated in acute care settings for many reasons other than prevalent chronic medical conditions such as heart failure, diabetes, and COPD as well as prevalent acute medical conditions such as pneumonia.

Between baseline and both intervention periods, we observe substantial regression-to-the mean in the hospitalization rates. In the original comparison population, the rate of all-cause hospitalization declined roughly 35% between baseline and both intervention periods. ER visits at baseline were less than half of the hospitalization rates, and we observe virtually no change in rates between the baseline and intervention periods.

The rates of all-cause and ACSC hospitalization and ER visits declined similarly in the original intervention and the comparison groups between the baseline and the initial demonstration period. However, there was a statistically significant greater decline in the 10 ACSC hospitalization rate for the original intervention group than in the intervention group during the last 12 months of the demonstration. The D-in-D rate is -67 per 1,000 beneficiaries lower in the intervention group. There are no other positive intervention effects in acute care utilization during the last 12 months of the demonstration for the original population.

Rates of hospitalization and ER visits per 1,000 refresh population beneficiaries for the year prior to go-live and months 7-18 of the CLM refresh demonstration period are presented in ***Table 6-2***. We observe no pattern of regression-to-the mean in the refresh populations' acute care utilization measures. In fact, we observe an increase in the 10 ACSC hospitalization rate and all-cause and 10 ACSC ER visit rates in the comparison group during the pilot period.

The rate of all-cause hospitalizations declined more in the refresh intervention group than in the comparison group between baseline and the last 12 months of the CLM demonstration period by -90 per 1,000 beneficiaries. And, the rate of 10 ACSC causes of hospitalization increased less in the refresh intervention group than in the comparison group by -58 per 1,000 beneficiaries. Both differences are statistically significant. We observe no differential rates of ER usage—either all-cause or for ambulatory care sensitive conditions—during the demonstration period relative to the baseline period.

Table 6-2
Comparison of rates of utilization for months 6-17 and the last 12 months of the CLM CMHCB demonstration with rates of utilization for a 1-year period prior to the start of the CLM CMHCB demonstration: Refresh population

Utilization	Baseline rate per 1,000 I[1,2,3]	Baseline rate per 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	1,547	1,531	1,338	1,412	-90	0.94	0.05	0.88	1.00
10 ACSCs[5]	518	495	516	551	-58	0.89	0.03	0.81	0.99
ED/Obs visits									
All cause	711	733	759	820	-39	0.95	0.42	0.85	1.07
10 ACSCs	109	121	120	125	7	1.06	0.55	0.87	1.30

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; ED/Obs = emergency room visits, including observation bed stays; CMO = care management organization; COPD = chronic obstructive pulmonary disease; UTI = urinary tract infection.

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

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

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

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

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

SOURCE: RTI analysis of 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 acsctab acsc acsctab1 26JUN2009.

6.3.2 Rates of 90-Day Readmissions

Table 6-3 displays the number of original and refresh population 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. There are no statistically significant reductions in admissions or readmissions among the original intervention beneficiaries during the early stage of the demonstration (months 3-14). However, during the last 12-month demonstration period, there is a statistically significant reduction in the percent of original intervention beneficiaries with an ACSC admission, a modest D-in-D rate of decline of -2 per 1,000 beneficiaries[11]. Yet, there are no observed reductions in percent of beneficiaries with a 90-day ACSC readmission or the ACSC readmission rate per 1,000 beneficiaries. Also during the last 12 months of the demonstration period, we observe no decline in percent of beneficiaries with an all-cause admission[12] but we do observe a sizeable reduction in the all-cause readmission rate of -225 per 1,000 beneficiaries. Given that we observe no decline in percent of beneficiaries with all-cause readmissions, the decline in the readmission rate implies that CLM was more successful at reducing readmissions for beneficiaries with frequent readmissions than for beneficiaries with less frequent readmissions.

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

Counts of beneficiaries	Intervention	Comparison
Original beneficiaries		
Months 3-17		
Total number of beneficiaries	11,150	4,446
Full time equivalents[1]	11,114	4,430
Months 15-29		
Total number of beneficiaries	9,209	3,634
Full time equivalents[1]	9,185	3,621
Refresh beneficiaries		
Months 4-18		
Total number of beneficiaries	12,308	4,928
Full time equivalents[1]	12,224	4,897

NOTES: CLM = Care Level Management.

[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.

[11] As seen in ***Table 6-1***, there is a reduction in the ACSC admission rate.

[12] We do not observe a decline in rate of all-cause hospitalizations.

Table 6-4
Change in 90-day readmission[1] rates between the year prior to the CLM CMHCB demonstration and months 3-14 and months 14-24 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 3-14									
Hospitalizations									
Percent with an admission	76	77	49	50	1	1.06	0.28	0.95	1.19
Percent with ACSC[5] admission	32	32	23	24	-0	0.98	0.68	0.87	1.09
All-cause 90-day readmission									
Percent with readmission	56	57	46	49	-3	0.90	0.11	0.79	1.02
Readmission rate / 1,000	1,189	1,214	1,099	1,163	-40	0.96	0.49	0.87	1.07
ACSC same-cause 90-day readmission									
Percent with readmission	15	15	16	15	0	1.02	0.86	0.78	1.34
Readmission rate / 1,000	230	220	250	261	-20	0.92	0.57	0.68	1.24
Months 15-26									
Hospitalizations									
Percent with an admission	75	76	46	46	1	1.05	0.47	0.93	1.18
Percent with ACSC admission	29	29	20	23	-2	0.87	0.04	0.77	0.99
All-cause 90-day readmission									
Percent with readmission	55	55	44	47	-3	0.90	0.14	0.77	1.04
Readmission rate / 1,000	1,117	1,106	1,053	1,268	-225	0.82	0.00	0.73	0.93
ACSC same-cause 90-day readmission									
Percent with readmission	14	12	15	17	-3	0.80	0.15	0.58	1.09
Readmission rate / 1,000	217	177	255	281	-66	0.74	0.10	0.52	1.06

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odd ratio; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; COPD = chronic obstructive pulmonary disease; UTI = urinary tract infection.

[1] 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.

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

[3] 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.

[4] 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: readm04 readm05 tab6_3_1x 26JUN2009

Table 6-5 displays the percent of refresh population beneficiaries with an admission and readmission and rate of readmission per 1,000 beneficiaries. We observe a statistically significant decline in percent of beneficiaries with an ACSC hospitalization among intervention beneficiaries relative to the comparison beneficiaries during the last 12 months of the demonstration, but it is a very modest decline of -2 per 1,000 beneficiaries[13]. Yet, we do not observe a decline in the percent of beneficiaries with an ACSC readmission or in the rate of ACSC readmissions. Although we do not observe a reduction in the percent of beneficiaries with an all-cause hospitalization, we observe an unexceptional reduction in the percent of refresh intervention beneficiaries with an all cause 90-day readmission, 42 per 1,000, but do not observe a statistically significant decline in rate of all-cause readmission.

6.3.3 Mortality

Table 6-6 displays mortality rates during the CLM CMHCB demonstration for both the original and refresh intervention and comparison populations. Over the 29-month demonstration period for the original population, 30% of both the intervention and comparison group beneficiaries died. And, during the 18-month demonstration period for the refresh population, 26% of the intervention group beneficiaries and 27% of the comparison group beneficiaries died. Thus, we observe no statistically significant differences in mortality rates for either population. However, the percentage point difference in mortality rates is quite small despite an 11 month longer demonstration period for the original population. As noted in Chapter 4, the refresh population was substantially sicker than the original population as measured during the year prior to each population entering the demonstration. The original population had a mean Hierarchical Condition Category (HCC) score of 2.8 in contract to a mean HCC score of 3.8 for the refresh population.

6.4 Conclusions

RTI's analysis of quality of care focuses on measuring effectiveness of the CLM CMHCB demonstration intervention by answering the following evaluation questions:

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

During the first 14 months of CLM program operations, CLM was not successful in reducing acute hospitalizations, ER visits, or 90-day readmissions in their original intervention population. However, during the last 12 months of the demonstration, CLM was somewhat successful in reducing acute care utilization. Within its original population, the percent of intervention beneficiaries admitted for 1 of 10 ambulatory care sensitive conditions and the admission rate per 1,000 intervention beneficiaries declined modestly. There was no decline in all-cause admissions or the percent with an all-cause admission, however we do observe a sizeable reduction in the all-cause readmission rate of -225 per 1,000 beneficiaries. Given that we observe no decline in percent of beneficiaries with all-cause readmissions, the decline in the

[13] As seen in Table 6-2, there is a lower rate of growth in ACSC admissions within the intervention group.

Table 6-5
Change in 90-day readmission[1] rates between the year prior to the CLM CMHCB demonstration and months 4-15 of the demonstration: Refresh population

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo period rate per 1,000[1,2,3] I	Demo period rate per 1,000[1,2,3] C	D-in-D	OR/IRR[4]	p	Low CI	High CI
Months 4-15									
Hospitalizations									
Percent with an admission	62	62	52	54	-2	0.93	0.17	0.85	1.03
Percent with ACSC[5] admission	31	30	26	27	-2	0.90	0.05	0.81	1.00
All-cause 90-day readmission									
Percent with readmission	57	55	49	52	-4	0.85	0.01	0.74	0.96
Readmission rate / 1,000	1,246	1,235	1,222	1,276	-66	0.95	0.27	0.86	1.04
ACSC same-cause 90-day readmission									
Percent with readmission	17	17	17	18	-0	0.97	0.83	0.77	1.23
Readmission rate / 1,000	270	299	282	318	-7	0.98	0.89	0.76	1.27

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odd ratio; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition; COPD = chronic obstructive pulmonary disease; UTI = urinary tract infection.

[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: readm04 readm05 tab6_3_1x 26JUN2009

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

Description	Intervention number of deaths	Percent	Comparison number of deaths	Percent	Difference	p value
Original population (29 months)	3,392	29.5	1,361	29.8	0.4	0.63
Refresh population (18 months)	3,462	26.4	1,390	27.0	0.1	0.88

NOTES: CLM = Care Level Management; CMHCB = Medicare Care Management for High Cost Beneficiaries.

SOURCE: RTI analysis of 2004-2008 Medicare enrollment data base.

Computer runs: mortality.sas 27JUN2009

readmission rate implies that CLM was more successful at reducing readmissions for beneficiaries with frequent readmissions than for beneficiaries with less frequent readmissions.

During that last 12 months of their demonstration program, CLM also had some success in reducing the acute care utilization in its refresh population. Rates of all-cause and ACSC hospitalizations declined, as did the percent with a readmission for all causes or an ACSC. CLM was not successful reducing rates of ER visits in either time period for the original population or in the last 12 months of the demonstration for the refresh population. CLM had no success reducing mortality in either the original or refresh population.

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 29 months that the Care Level Management (CLM) Care Management for High Cost Beneficiaries (CMHCB) program was in operation (or 16 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 29 or 16 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, alone, materially reduce the intervention's overall cost savings?
- 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?
- How did Medicare savings for the 29- or 16-month period compare with the fees that were paid out? How close was CLM in meeting budget neutrality?
- 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?
- Did the intervention have a differential effect on high cost and high risk beneficiaries?
- What evidence exists for regression-to-the-mean (RtoM) in Medicare costs for beneficiaries in the intervention and comparison groups?

The cost analyses presented in this section differ from those that will be conducted for financial reconciliation by Actuarial Research Corporation (ARC) under contract to the Centers for Medicare & Medicaid Services (CMS). ARC will determine savings based on the demonstration's terms and conditions negotiated between CMS and CLM. RTI's estimation of savings, detailed subsequently, differs in that

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

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

The rest of this chapter has five sections. The next two sections describe our data sources, variable construction, and analytic methods. ***Section 7.4*** presents our primary findings on trends in PBPM costs between base and demonstration periods. ***Section 7.5*** shows PBPM savings in relation to average monthly fees and whether CLM achieved budget neutrality using RTI's costing methods. ***Section 7.6*** stratifies 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 samples that might affect t-tests of mean differences in PBPM growth rates. The chapter concludes in ***Section 7.9*** with a summary of key findings.

7.2 Data and Key Variables

7.2.1 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. Eligibility in the original and refresh samples was based on the following criteria.

Original Sample:

- Medicare fee-for-service beneficiaries with a primary residence in designated counties of California, Florida, or Texas, with high costs in 2004 (i.e., top 5% of annual Medicare costs),
- Two or more hospitalizations in 2004, and
- At least one diagnosis from a list provided by CLM, such as heart failure (HF).

CLM then requested that the original sample be reconfigured by carving out a portion of high cost beneficiaries not meeting the following criteria. Beneficiaries from the original starting population would be retained if they:

Carve-out (Modified Original Sample):

- had a Hierarchical Condition Category (HCC) score that was 2.75 or greater, or
- had an HCC score less than 2.75 but had a diagnosis of selected clinical conditions such as peripheral vascular disease, ischemic heart disease, hypertensive heart and/or kidney disease, heart failure, chronic obstructive pulmonary disease (COPD), and asthma.

Refresh Sample:

- were Medicare fee-for-service beneficiaries with a primary residence in designated counties of California, Florida, or Texas that had claims in calendar year 2005,
- had a 2005 HCC risk score >2.75, and
- had two or more hospitalizations in 2005.

Of these beneficiaries, the following were excluded:

- beneficiaries in HCC groups 51 (drug/alcohol psychosis), 52 (drug/alcohol dependence), 54 (schizophrenia), and 55 (major depressive, bipolar, and paranoid disorders), or
- hospitalizations with specific diagnosis-related group (DRG) codes were not to be counted toward the admission requirement (e.g., acute major eye infections, kidney and urinary tract infections). In addition to the DRG exclusion, inpatient claims that had discharge dates equal to admission date were excluded, as well as inpatient claims where the admission date matched the discharge date of a prior claim, or
- beneficiaries who were institutionalized during the last three months of 2005.

The original sample focused on beneficiaries with very high annual costs, and it would be quite likely to expect lower costs during the demonstration period based on RtoM. By still requiring very high HCC scores and multiple hospitalizations, RtoM is still likely to be profound in the comparison group, implying major reductions in intervention costs required to achieve budget neutrality.

Because of more than a year's gap between selection for and the start of the demonstration, a new base year of claims data were extracted for the intervention and comparison populations. Consequently, it is likely that some beneficiaries who originally qualified during the randomization process would no longer qualify for the demonstration during the base period just 1 year before CLM's start date. They still remain in the intervention and comparison groups, however, for our analysis.

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

7.2.2 Constructing PBPM Costs

All financial analyses were conducted on a PBPM basis, or the ratio of eligible Medicare costs to eligible months. The baseline period is defined as 365 days (or 1 year) prior to CLM's

start date. The 29-month demonstration period for the original population includes 882 days (29 months × 30.42 days/month) after the start date. The refresh population covers 16 months, or 487 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[14].

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 29 or 16 months of demonstration period eligibility. However, some beneficiaries had fewer than the maximum number of eligible months (or days), usually due to death. At the extreme, a beneficiary could have a 10-day hospital admission at the beginning of the intervention period with a combined Part A and B payment of $30,000 before dying. This $30,000 outlay is divided by approximately 1/3 (10 days / 30.42 days), resulting in an adjusted PBPM 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 mean intervention group PBPM costs in CLM's original population (11,516) stratified by beneficiaries' number of eligible days in the demonstration period (487 maximum). Those with 10 or fewer eligible days had overall PBPM costs averaging $10,191. Beneficiaries eligible for a year or more averaged $2,435. Maximum PBPM costs were over $200,000. Thus, beneficiaries with very truncated eligibility averaged monthly costs 4.2 times greater than those with much longer eligibility. Although beneficiaries with a month or less of eligibility were only about three-tenths of 1% of the entire intervention group, their PBPM costs add disproportionately both to the mean and variation in PBPM costs. (See ***Section 7.3.2*** for statistics on PBPM variation.)

Table 7-2 shows the effects of short term eligible beneficiaries in the refresh population (13,104). Again, short-eligibility beneficiaries were over 4 times as costly per month as those with more than 1 year's eligibility. Maximum PBPM costs were over $95,000.

Variation can be reduced by trimming high PBPM outliers, as done by CMS for financial reconciliation at the 99th percentile. In addition, no maximum spending threshold was applied to any beneficiary's average PBPM cost. While the 1% trim reduces the CMO's financial risk, we wanted to avoid biasing comparisons against interventions that constrained spending among the most expensive beneficiaries.

---

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

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

Eligible days[1]	N (%)	PBPM	Range
< 10	37 (0.3%)	$10,191	$0–116,933
11–30	108 (0.9)	15,049	0–208,446
31–60	188 (1.6)	12,951	0–150,025
61–90	206 (1.8)	8,719	0–64,897
91–365	1,577 (13.7)	6,356	0–68,742
366+	9,400 (81.6)	2,435	0–56,183
Mean	11,516	3,399	0–208,446

NOTES: Observations unweighted. CLM = Care Level Management; 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 200x-200Y Part A & B claims; clm/9mo/COSTRUN2(8/17/09).

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

Eligible days[1]	N (%)	PBPM	Range
< 10	73 (0.6%)	$11,203	$0–71,016
11–30	196 (1.5)	11,887	0–95,712
31–60	223 (1.7)	12,404	0–91,142
61–90	266 (2.0)	9,624	0–60,376
91–365	2,319 (17.7)	6,715	0–82,391
366+	10,027 (76.5)	2,441	0–51,257
Mean	13,104	3,703	0–95,712

NOTES: Observations unweighted. CLM = Care Level Management; 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 200x-200Y Part A & B claims; clm/9mo/COSTRUN2(8/17/09).

Alternatively, the method RTI adopted was to weight PBPM mean costs and standard errors by each beneficiary's eligible fraction of days for the intervention period. In the previous example, the beneficiary's adjusted $90,000 PBPM cost is weighted by 10/882 = 0.0113 in the original population, or about 88 times less than weight given to 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 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 CLM, its negotiated management fee was $295.

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 CMO was neither advantaged nor disadvantaged by the costliness of their sample 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 CLM begin 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 CLM'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.[15] 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

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

zero to assure the government that the measured savings rate was not due to chance.[16] 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 CLM.

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[17] 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 $DT/PBPM_C$.

Tables 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 CLM's intervention and comparison, original (carved out) and refresh samples. Mean PBPM costs in the base period ranged from a low of $0 to a high of $208,446 in the 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 equal to mean costs). CVs in the

---

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

17 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.

original and refresh samples increased substantially during the demonstration period, implying growing variation in monthly costs across beneficiaries.

Table 7-3
CLM CMHCB 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)	(4,561)	(11,516)	(4,561)	(11,516)
Minimum	$0	$0	$0	$0
<10%	349	338	231	240
<25%	1,141	1,105	640	625
Median	2,930	2,886	1,758	1,729
>25%	4,924	4,843	4,126	3,927
>10%	7,152	7,346	7,770	7,725
Maximum	41,354	57,227	171,244	208,446
Mean	3,527	3,522	3,420	3,399
CV	.94	.96	1.80	1.75

NOTES: Observations unweighted. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; N = number of beneficiaries; CV = coefficient of variation.

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

SOURCE: Medicare 200x-200Y Part A & B claims; COSTRUN2(7/9/09).

Table 7-4
CLM CMHCB 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)	(5,240)	(13,104)	(5,240)	(13,104)
Minimum	$0	$0	$0	$0
<10%	0	0	215	215
<25%	311	318	604	605
Median	1,812	1,833	1,882	1,867
>25%	4,298	4,249	4,676	4,487
>10%	7,571	7,355	9,043	8,937
Maximum	62,539	74,491	155,152	95,712
Mean	3,020	2,982	3,798	3,703
CV	1.33	1.26	1.69	1.55

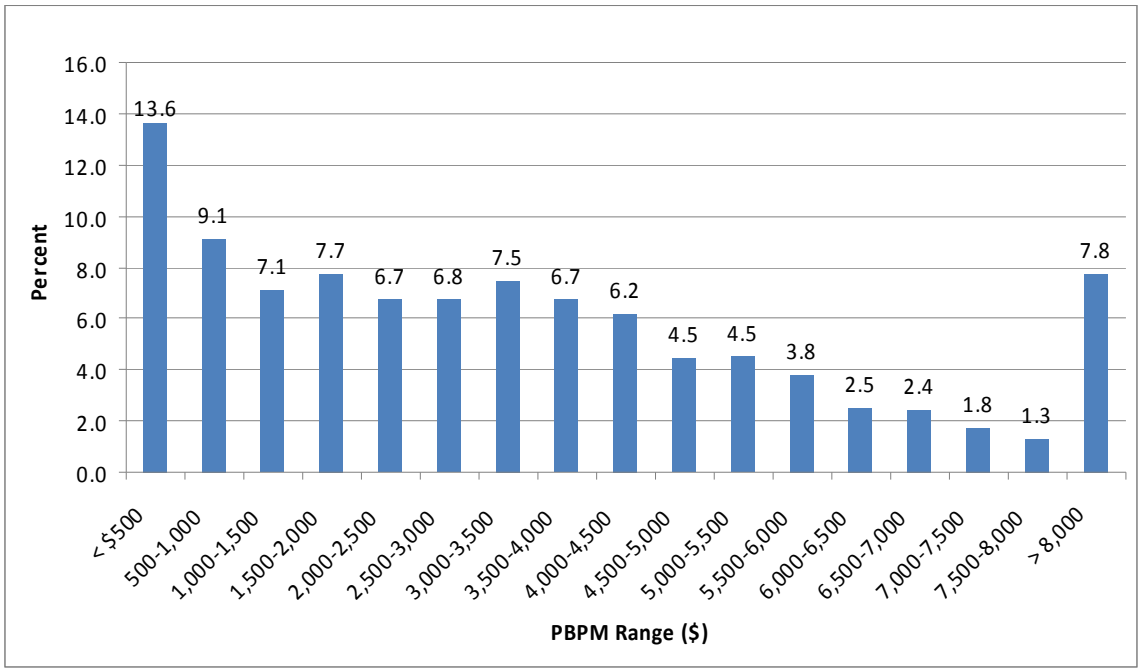
NOTES: Observations unweighted. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; N = number of beneficiaries; CV = coefficient of variation.

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

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/COSTRUN2(8/17/09).

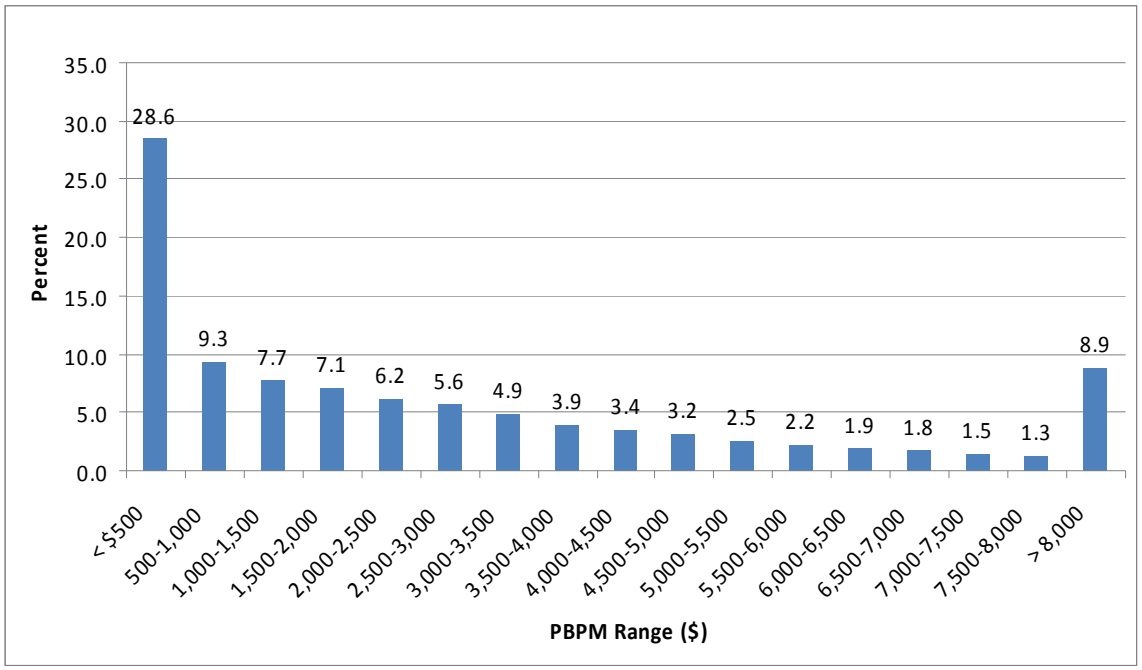
The difference between median and mean PBPM costs indicates how skewed costs actually are. (Also, see ***Figures 7-1*** and ***7-2***.) Mean costs are about 20% greater than median costs in the original sample's base year, but the gap increases to a factor of two in the demonstration period, indicating a strong right tail of high costs. Costs were more skewed in the refresh group. Note the high percentage with less than $500 in base year costs. This group experienced large increases in costs during the demonstration, as shown later. Maximum values show how high PBPM costs can be before weighting—greater than $200,000 per month in the demonstration period. 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.

Figure 7-1
Frequency distribution of PBPM costs, comparison group, original sample, base year: CLM



NOTES: PBPM = per beneficiary per month; CLM = Care Level Management.

Figure 7-2
Frequency distribution of PBPM costs, comparison group, refresh sample, base year: CLM



NOTES: PBPM = per beneficiary per month; CLM = Care Level Management.

Despite the large variances in costs in the intervention and comparison groups during the base year and demonstration period, the power of the sample sizes to detect as small a difference as 5% between the entire intervention and comparison group's PBPM cost trend was still 85%.[18]

7.3.3 Budget Neutrality

Each CMO is obligated to produce net savings for the Medicare program. The net savings requirements for those CMOs that complete a 36-month demonstration period are 5% for the original cohort and 2.5% for the refresh cohort. However, since CLM's demonstration was terminated on February 29, 2008, the net savings requirement was pro-rated (as stated on pages 9 and 10 of the financial protocol). The pro-rated net savings requirement is determined by the number of months after the effective start date for each cohort that the demonstration is terminated. In CLM's case, the termination date is 29 months after the start date of the original cohort and 18 months after the start date of the refresh cohort. Therefore, the net savings requirements for CLM are 4% for the original cohort and 2% for the refresh cohort.

Unless the CMO achieves 4% net savings for the original cohort and 2.5% net savings for the refresh cohort, it must return some or all of its fees. In the original population, the intervention PBPM cost must be equal or less than 96% of the comparison group PBPM cost minus the average monthly fee (MF); that is,

$$PBPM_I <= 0.96PBPM_C - MF, \qquad (7.6a)$$

or as a fraction of the comparison PBPM cost,

$$PBPM_I / PBPM_C <= 0.96 - (MF / PBPM_C), \qquad (7.6b)$$

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 91% or less of monthly comparison costs to avoid paying back fees. Debt obligation per intervention beneficiary month is the positive difference: $PBPM_I - [0.96PBPM_C + MF]$.

RTI's conclusion regarding budget neutrality will differ from those of the 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.

---

18 Power = $\Phi[-z_{.975} + \Delta / SE_{[\Delta I - \Delta C]}]$. For CLM, Power = Φ [-1.96 + .05*$2,501/$133] = Φ [-1.04] = 0.85. See Table 7-5 for comparison group PBPM mean costs and standard errors of differences in growth rates. Power figures would be higher when adjusting the standard errors for the positive correlation of PBPM costs by beneficiary between the base and demonstration periods. (See Rosner, 2006, p. 333)

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 actually 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 (φ 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 comparison group's average PBPM cost in the base year, while γ = each beneficiary's dollar increase in PBPM costs over 14 months (i.e., the sixth month of the base year to the eighth midperiod month of the demonstration). γ provides a test of RtoM effects (see ***Supplement 7-1***). If γ is less than 1.0, the beneficiary's PBPM during the demonstration period increased by a lower dollar amount (and percentage) the greater the beneficiary's base period PBPM cost. 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 Sample

Table 7-5 displays PBPM cost levels and rates of growth in average PBPM costs between the 12-month base year and the 29-month demonstration period for the original (carve-out) sample. 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
CLM CMHCB PBPM cost growth rates between base year and demonstration period, intervention and comparison groups: Original population

Study group	Beneficiaries	Base year PBPM Mean[1]	Base year PBPM SE	Demo PBPM Mean[1]	Demo PBPM SE	Differences in means	SE
Intervention	11,516	$3,296	28.9	$2,478	36.9	-$818**	36.1
Participants	7,279	3,207	35.3	2,329	32.7	-879**	40.7
Nonparticipants	4,227	3,509	50.8	2,836	66.1	-673**	72.2
Comparison	4,561	3,360	46.7	2,501	47.8	-859**	57.5
Differences							
I – C	—	-64	54.6	-23	58	41	67.9
Participants - C	—	-153**	58.7	-173**	56.5	-20	69.5
Nonparticipants - C	—	149*	69.4	335**	80.3	186*	91.6
Participants - Nonparticipants	—	-302**	63.4	-507**	67.6	-206**	79.2

NOTE: CLM = Care Level Management; 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 clm/final/9mo/costrun1(8/17/09).

The weighted base year average PBPM cost was $64 less ($p$ = insig) in the intervention group versus the comparison group ($3,296 versus $3,360), or 1.9%. The intervention-comparison gap in PBPM Medicare costs shrank to -$23 ($p$ = insig) in the demonstration period ($2,478 versus $2,501). Between the base year and the 29-month demonstration period, the average comparison group PBPM declined significantly by $859 ($p < .01$), while the intervention group's PBPM average Medicare costs declined by $818 ($p < .01$). Thus, the intervention group's PBPM cost declined $41 slower ($p$ = insig) than the comparison group's PBPM cost. Intervention beneficiaries, who were 1.9% less costly at baseline, became only 0.9% less costly, on average, than the comparison group after 29 months.

The participation rate, based on beneficiaries used in this cost analysis, was 63% (7,279 / 11,516 - 1). Participants in the CLM intervention group were $153 less costly ($p < .01$) than the comparison group beneficiaries in the base period and nonparticipants were $149 more costly ($p < .05$). Nonparticipants became $673 less costly ($p < .01$) during the demonstration period while participants became $879 less costly ($p < .01$). The participant group's PBPM cost decreased $20 faster ($p$ = insig) than the comparison group's while the nonparticipant group's PBPM cost decreased $186 slower ($p < .05$) than the comparison group's PBPM cost.

7.4.2 Refresh Sample

Table 7-6 displays PBPM cost levels and rates of growth in average PBPM costs between the 12-month base year and the 16-month demonstration period for the refresh sample. The weighted base year average PBPM cost was $3 more ($p$ = insig) in the intervention group versus the comparison group ($2,763 versus $2,760), or 0.1%. The intervention-comparison gap in PBPM Medicare costs grew to -$26 ($p$ = insig) in the demonstration period ($2,780 versus $2,807). Between the base year and the 16-month demonstration period, the average comparison group PBPM increased $46 ($p$ = insig), while the intervention group's PBPM average Medicare costs increased $18 ($p$ = insig). Thus, the intervention group's PBPM cost increased $29 slower ($p$ = insig) than the comparison group's PBPM cost. Intervention beneficiaries, who were 0.1% more costly at baseline, became only 1.0% less costly, on average, than the comparison group after 16 months.

The participation rate, based on beneficiaries used in the refresh cost analysis, was 66% (8,670/13,104 – 1). Participants in the CLM intervention group were $3 more costly ($p < .01$) than the comparison group beneficiaries in the base period and nonparticipants were $1 more costly ($p$ = insig). Nonparticipants became $380 more costly ($p < .01$) during the demonstration period while participants became $124 less costly ($p < .01$). The participant group's PBPM cost decreased $170 faster ($p < .05$) than the comparison group's while the nonparticipant group's PBPM cost increased $334 faster ($p < .01$) than the comparison group's PBPM cost.

Table 7-6
CLM CMHCB PBPM cost growth rates between base year and demonstration period, intervention and comparison groups: Refresh population

Study group	Beneficiaries	Base year PBPM Mean[1]	Base year SE	Demo PBPM Mean[1]	Demo PBP M SE	Differences in means	SE
Intervention	13,104	$2,763	30.1	$2,780	33.4	18	38.8
Participants	8,670	2,763	35.7	2,639	36.9	-124**	44.5
Nonparticipants	4,434	2,761	56.4	3,142	70.0	380**	76.6
Comparison	5,240	2,760	49.5	2,807	53.6	46	73.1
Differences							
I - C	—	3	57.0	-26	62.7	-29	73.1
Participants - C	—	3	60.4	-167**	63.7	-170*	76.0
Nonparticipants - C	—	1	75.5	335**	87.1	334**	99.0
Participants - Nonparticipants	—	2	67.0	-502**	74.2	-504**	86.2

NOTE: CLM = Care Level Management; 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 clm/final/9mo/costrun1(8/17/09).

7.5 Savings and Budget Neutrality

7.5.1 Original Sample

Table 7-7 presents summary statistics on savings from the CLM intervention with the original (carve-out) sample. 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. CLM's monthly fee is reported also as a percentage of the comparison group's PBPM cost.

Over the course of the 29-month intervention, average monthly costs decreased $818 in the intervention group and $859 in the comparison group. The result was a $41 greater relative increase in PBPM cost growth in the intervention group. This positive difference implies *dissavings* rather than savings, at a rate of 1.6% of the comparison group's PBPM cost.

With roughly 11,500 beneficiaries in the intervention group and 4,500 in the comparison group, we had the power to detect a savings of $133 or more (in negative terms) at the 95% confidence level. This rate is 5.3% of the comparison group's PBPM cost, implying substantial power to detect relatively modest savings, if they existed.

CLM's monthly fee was $295, which amounted to 11.8% of the comparison group's PBPM during the demonstration period. Thus, CLM would have had to achieve 15.8% (11.8% + 4%) savings in order to retain all of its fees—at least according to RTI's calculations, which are not official under financial reconciliation.

If one believed that, in fact, CLM's intervention actually increased beneficiary costs by $41, then the net effect on Medicare costs would be $295 plus $41, or $336 per beneficiary per month. Instead of reducing the 11.8% fee outlay on PBPM costs, the intervention would appear to have increased Medicare total costs from 11.8% to 13.4% of the comparison group's costs.

7.5.2 Refresh Sample

Table 7-8 presents summary statistics on savings from the CLM intervention with the refresh sample. Over the course of the 16-month intervention, average monthly costs increased $18 in the intervention group and $46 in the comparison group. The result was a $29 smaller relative increase in PBPM cost growth in the intervention group. This positive difference implies *savings* at a rate of 1.0% of the comparison group's PBPM cost.

With roughly 13,400 beneficiaries in the intervention group and 5,200 in the comparison group, we had the power to detect a savings of $143 or more (in negative terms) at the 95% confidence level. This rate is 5.1% of the comparison group's PBPM cost, implying substantial power to detect relatively modest savings, if they existed.

Ignoring the fact that the $29 in intervention savings was not statistically different from 0, the net fee to Medicare was reduced from $295 per beneficiary per month to $266, resulting in a net cost of 9.5% of the comparison group's average monthly outlay on claims.

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

Description	PBPM cost change
Intervention group	-$818
Comparison group	-$859
Difference	$41
Gross (dis)saving %[1]	1.6%
Minimal Detectable Savings[2]	
Absolute	-$133
% of comparison PBPM[3]	-5.3%
Monthly Fee	
Absolute[4]	$295
% of comparison PBPM[3]	11.8%
Net Fee	
Absolute	$336
% of comparison PBPM[3]	13.4%

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison PBPM (= $2,501). 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 in demonstration period.

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

SOURCE: Medicare 200X-200Y Part A&B claims; PBPM cost changes and detectable savings: Table 7-5; monthly fees: CMS Project Officer.

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

Description	PBPM cost change
Intervention group	$18
Comparison group	$46
Difference	-$29
Gross (dis)saving %[1]	-1.0%
Minimal Detectable Savings[2]	
Absolute	-$143
% of comparison PBPM[3]	-5.1%
Monthly Fee	
Absolute[4]	$295
% of comparison PBPM[3]	11.8%
Net Fee	
Absolute	$266
% of comparison PBPM[3]	9.5%

NOTES: CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month.

[1] Gross (Dis)Savings % = Difference in PBPM outlay changes as % of comparison PBPM (= $2,807). 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 in demonstration period.

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

SOURCE: Medicare 200X-200Y Part A&B claims; PBPM cost changes and detectable savings: Table 7-6; monthly fees: CMS Project Officer.

7.6 Imbalances between Intervention and Comparison Samples

In the CLM demonstration before the carve out, beneficiaries were eligible if they incurred annual costs that put them in the top 5% of Medicare beneficiaries in the year prior to selection. After the carve out, eligibility was based on an HCC score greater than 2.75 or if they had a particular chronic disease. We were still interested in how balanced the intervention samples were on both cost and severity, as proxied by the HCC score. Besides the HCC > 2.75 criterion, we also identified another group that would have qualified based only on high cost. We chose an $18,000 annual threshold to qualify that reflected the very high costs in general of CLM beneficiaries.[19] These two criteria produced four cost-risk groups: (1) high cost only > $18,000; (2) high risk only > 2.75; (3) both high cost and high risk; and (4) neither. The neither group will include beneficiaries who were included because of their particular chronic disease yet less than $18,000 in costs or with an HCC < 2.75. Another source of intervention-comparison group differences in beneficiary characteristics is that a more current period than the sampling period was used to estimate the base year PBPM costs. Differential attrition between intervention and comparison groups could produce further deviations in underlying beneficiary characteristics, including their costliness and HCC scores just prior to the start date.

Initial random sampling should have balanced the intervention and comparison groups. Yet, it is still possible that small, but possibly important, imbalances remained simply by chance. It is possible that high cost and high risk beneficiaries exhibit opposing RtoM cost trends between the base and demonstration periods. High cost beneficiaries should have declining costs, while high risk but lower cost beneficiaries might have increasing costs. If the distribution of high cost and high risk beneficiaries differs between CLM'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.

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.

Because beneficiaries were randomly assigned to the intervention and comparison groups, differences in cost-risk and patient characteristics across the two groups should be minimal even with some attrition. Nevertheless, we test for the cost impacts of any imbalances as shown below.

---

[19] According to the financial reconciliator, ARC, the mean PBPM cost of the intervention group in California was $4,506, or $54,072 annually. We chose $18,000 as a minimum cost threshold to qualify on cost to approximate the top 5% threshold that would be consistent with mean costs of $54,000.

7.6.1 Frequencies of Beneficiary Characteristics

Table 7-9 and ***7-10*** show that the intervention and comparison groups were nearly identically distributed by cost and risk during the randomization period. Nearly 40% of original (carve-out) beneficiaries in the official base year prior to the demonstration had *both* an HCC score greater than 2.75 and annual costs above $18,000. Only about 4% would have qualified based only on their high risk score if that had been a second criterion. In the period that base costs were determined, over one-quarter of beneficiaries did not meet either selection criterion and likely qualified because of their chronic diseases, but this shift did not unbalance the two groups, at least regarding costs and HCC severity risk. These similarities would indicate that the lack of intervention savings cannot be explained by intervention-comparison group differences in cost and risk group status.

7.6.2 PBPM Cost Levels and Trends by Cost and Risk Group

7.6.2.1 Original Sample

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 intervention beneficiaries qualifying only with an HCC score greater than 2.75 averaged PBPM costs of just $919 compared with $3,510 for high cost–only beneficiaries (3.8 times greater) and both high cost and high risk beneficiaries ($5,417; 5.9 times greater). The gap narrowed considerably in the demonstration period as a result of opposing cost trends. Both high cost groups experienced large declines in their PBPM costs while the high risk–only group experienced an average increase of $1,432 ($p < .01$). The comparison group shows almost identical patterns of cost levels and trends.

Focusing on the difference in trends at the bottom of Table 7-11, we find no statistically significant differences between intervention and comparison group growth rates in the four cost-risk groups. The $170 faster cost declines in the high cost intervention group was not quite significant at the 10% confidence level.

7.6.2.2 Refresh Sample

Table 7-12 presents similar results on PBPM cost trends by the four cost-risk groups for the refresh sample. None of the differences-in-differences in growth rates are statistically significant across the four groups. The high cost-only refresh group had faster cost reductions in the intervention group ($236) but this difference was not statistically significant.

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. Using comparison group data avoids any effects the intervention might have on the underlying RtoM phenomenon. Unweighted mean costs were $3,020 in the comparison group's base period in the original (carve-out) sample, with an overall increase of $778. Cost increases are inversely correlated with a beneficiary's base period PBPM costs. At the extremes, beneficiaries with less than $500 in base period PBPM costs saw their average costs increase by $2,931 while those with initial costs greater than $8,000 experienced average decreases of

$4,252. Mean costs in both periods are well above median costs and indicate a strong skewness in PBPM costs.

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

Characteristics	Intervention (%)	Comparison (%)
COST-RISK Group		
High cost > =$ 18,000	28.5%	28.8%
Both	38.5	40.1
High risk: HCC > 2.75	4.5	4.3
Neither	28.5	26.8
Age Group		
<65	13.5	12.2
65-69	16.2	15.7
70-74	18.3	18.5
75-79	19.0	20.7
80-84	16.2	17.2
85+	16.5	15.8
Gender		
Female	53.7	50.6
Male	46.3	49.5
Race		
Minority	26.8	26.0
White	73.2	74.0
MEDICAID Eligible		
No	91.1	90.6
Yes	8.9	9.4
DISABLED		
No	87.5	89.1
Yes	12.5	10.9
Urban residence		
No	2.4	2.1
Yes	97.6	97.9
Long-term care		
No	97.9	98.3
Yes	2.1	1.7
SNF		
No	73.8	73.6
Yes	26.2	26.4

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/Cost4b1(9/10/09).

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

Characteristics	Intervention (%)	Comparison (%)
COST-RISK Group		
High cost > =$ 18,000	8.3%	7.8%
Both	44.4	44.5
High risk: HCC > 2.75	10.1	10.6
Neither	37.2	37.1
Age Group		
<65	12.0	12.6
65-69	11.6	11.8
70-74	17.2	16.6
75-79	19.5	19.5
80-84	18.8	19.5
85+	21.0	20.0
Gender		
Female	52.9	52.6
Male	47.1	47.4
Race		
Minority	29.2	29.1
White	70.9	70.9
MEDICAID Eligible		
No	91.8	91.1
Yes	8.2	8.9
DISABLED		
No	89.1	88.3
Yes	10.9	11.7
Urban residence		
No	2.5	2.7
Yes	97.5	97.3
Long-term care		
No	98.1	98.1
Yes	1.9	1.9
SNF		
No	82.9	84.3
Yes	17.1	15.7

NOTE: Beneficiaries weighted by fraction of eligible days in demonstration period. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; HCC = Hierarchical Condition Category; SNF = skilled nursing facility.

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/Cost4b1(9/10/09).

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

Description	High-cost and high-risk PBPM	High-cost and high-risk SE	High-cost only PBPM	High-cost only SE	High-risk only PBPM	High-risk only SE	Neither PBPM	Neither SE
Intervention (N)	(4,845; 42%)	—	(3,105: 27%)	—	(523: 5%)	—	(3,043; 26%)	—
Base Year	$5,417	50.1	$3,510	30.5	$919	17.4	592	7.7
Demonstration	3,503	59.2	2,054	48.4	2,352	109.0	1,537	39.7
Difference	-1,913	69.6	-1,456	55.0	1,432	110.2	945	39.4
% Change	-35%	—	-42%	—	56%	—	60%	—
Comparison (N)	(1,957; 43%)	—	(1,239; 27%)	—	(207; 5%)	—	(1,158; 25%)	—
Base Year	5,383	80.6	3,510	47.7	896	25.2	578	12.3
Demonstration	3,363	83.5	2,225	83.5	2,254	173.0	1,533	72.6
Difference	-2,020	104.6	-1,286	90.1	1,357	174.4	975	72.5
% Change	-38%	—	-37%	—	51%	—	69%	—
Difference-in-Differences	107	127.6	-170	104.0	75	208.3	-30	78.4

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; 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 $18,000 in base period.

High-Risk: HCC > 2.75 in base period.

% Change: Difference/Base Year.

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/Cost4b1 (9/10/09).

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

Description	High-cost and high-risk PBPM	High-cost and high-risk SE	High-cost only PBPM	High-cost only SE	High-risk only PBPM	High-risk only SE	Neither PBPM	Neither SE
Intervention (N)	(6,142; 47%)	—	(1,027: 8%)	—	(1,270: 10%)	—	(4,665; 36%)	—
Base Year	$5,181	48.3	$3,118	54.6	$929	10.9	301	6.3
Demonstration	3,573	56.0	2,203	90.6	1,937	69.9	2,195	49.6
Difference	-1,608	64.0	-916	98.0	1,008	69.9	1,894	50.6
% Change	-31%	—	-29%	—	109%	—	629%	—
Comparison (N)	(2,456; 47%)	—	(387; 7%)	—	(538; 10%)	—	(1,859; 35%)	—
Base Year	5,195	81.2	3,111	100.9	914	17.0	297	9.9
Demonstration	3,609	90.3	2,431	169.7	2,034	129.1	2,146	74.1
Difference	-1,587	106.1	-680	189.4	1,120	129.7	1,849	75.9
% Change	-31%	—	-22%	—	123%	—	623%	—
Difference-in-Differences	-21	121.3	-236	197.5	-112	136.8	45	93.3

NOTE: Beneficiary PBPM weighted by fraction of eligible days in demonstration period. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; 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 $18,000 in base period.

High-Risk: HCC > 2.75 in base period.

% Change: Difference/Base Year.

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/Cost4b1 (9/10/09).

Table 7-13
CLM CMHCB Regression to the Mean in comparison group PBPM costs: Original population

Base year PBPM level	N	Base year PBPM	Demonstration period PBPM	Change
< $500	621	$245	$1,802	$1,557
500-1,000	416	732	2,544	1,812
1,000-1,500	324	1,239	2,769	1,530
1,500-2,000	353	1,752	2,859	1,108
2,000-2,500	306	2,255	3,340	1,085
2,500-3,000	308	2,750	3,802	1,053
3,000-3,500	341	3,239	3,066	-172
3,500-4,000	306	3,749	2,716	-1,033
4,000-4,500	283	4,240	3,453	-787
4,500-5,000	204	4,758	3,722	-1,036
5,000-5,500	207	5,236	3,691	-1,545
5,500-6,000	172	5,749	3,507	-2,242
6,000-6,500	116	6,243	4,681	-1,562
6,500-7,000	110	6,714	4,707	-2,007
7,000-7,500	80	7,181	8,415	1,234
7,500-8,000	60	7,731	4,110	-3,621
> 8,000	354	11,518	6,664	-4,854
Mean	4,561	3,527	3,420	-107
Median	—	2,930	1,758	-1,172

NOTES: Observations unweighted. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 200x-200Y Part A & B claims; COSTRUN2(7/9/09).

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

Base year PBPM level	N	Base year PBPM	Demonstration period PBPM	Change
< $500	1,498	$86	$3,017	$2,931
500-1,000	485	743	2,150	1,406
1,000-1,500	404	1,254	2,476	1,223
1,500-2,000	372	1,748	3,410	1,662
2,000-2,500	325	2,237	3,333	1,096
2,500-3,000	294	2,741	4,053	1,313
3,000-3,500	257	3,215	3,544	329
3,500-4,000	206	3,763	3,879	116
4,000-4,500	178	4,250	3,722	-527
4,500-5,000	169	4,733	4,467	-266
5,000-5,500	129	5,235	4,066	-1,169
5,500-6,000	116	5,734	4,805	-929
6,000-6,500	100	6,221	5,192	-1,029
6,500-7,000	95	6,752	4,753	-1,999
7,000-7,500	78	7,280	5,871	-1,409
7,500-8,000	70	7,772	5,169	-2,603
> 8,000	464	12,437	8,185	-4,252
Mean	5,240	3,020	3,798	778
Median	—	1,812	1,882	70

NOTES: Observations unweighted. CLM = Care Level Management; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; N = number of beneficiaries.

SOURCE: Medicare 200x-200Y Part A & B claims; COSTRUN2(7/9/09).

Regression-to-the-mean is also quite strong in the refresh sample (***Table 7-14***). However, mean (unweighted) costs fell slightly in the comparison group and median costs fell $1,172, or 40%. Thus, one-half of the refresh group had PBPM costs less than $2,930 in the base period, while fully one-half had costs less than $1,758 in the demonstration period. For mean PBPM costs to be little changed between the two periods ($107 less in the demonstration period), the size of the increases must have greatly exceeded the size of the decreases, albeit from a smaller group with increases.

7.8 Multivariate Regression Tests of Intervention Savings

7.8.1 Original Sample

Three 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 differences between each beneficiary's demonstration and base year PBPM costs. In the first column of results controlling only for each beneficiary's base period PBPM cost, the intervention coefficient of -5.1 is insignificant implying no statistical difference between intervention and comparison groups in terms of average cost changes. This intervention effect is somewhat different than shown in Table 7-5 (plus $41) because of the way beneficiaries' PBPM costs have been accounted for. In this regression format, changes in beneficiary costs are benchmarked against an average base year effect on future costs for each beneficiary instead of their actual change. Using either method, no significant savings are found.

The base period PBPM cost coefficient (0.283; $p < .01$), when combined with the intercept coefficient, implies substantial RtoM of costs (= 0.336 - 1 = -0.664, the RtoM effect). Imagine two comparison group beneficiaries, one with a relative low ($1,000) and another with a relatively high ($4,000) PBPM cost in the base period. The predicted PBPM cost of the initially "low cost" comparison beneficiary would increase 83% during the intervention period, while that of the "high cost" beneficiary would decline by one-third.[20] Whereas cost differences were 4:1 in the base period, they would now be compressed to 1.5:1.

RtoM effects are quite substantial but clearly not in one direction. Including only high cost beneficiaries in the original sample 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. Also note that the standard error of comparison group costs was slightly higher in the demonstration period, not lower (Table 7-5), as might be expected with compression of costs.

[20] The calculation is as follows based on Table 7-5, column 1:

PBPM[base]	PBPM[demo]	PBPM Change	%Change
$1,000	$1,829	$829	+83%
$4,000	$2,678	-$1,322	-33%

Table 7-15
Regression results: CLM intervention gross savings controlling for base period PBPM and beneficiary characteristics: Original population

Independent Variable	PBPM_ Demo Coefficient	PBPM_ Demo t	PBPM_ Demo Coefficient	PBPM_ Demo t	PBPM_ Demo Coefficient	PBPM_ Demo t
Intercept	1,546	39.0	1,414	28.5	1,475	5.2
Intervention	-5.1	0.1	2.5	0.1	-17.5	0.3
PBPM_Base	0.283	35.3	0.218	21.0	0.194	17.3
High cost–high risk	N/I	N/I	868.8	11.0	868.2	10.7
High cost	N/I	N/I	-76.4	1.1	-23.1	0.3
High risk	N/I	N/I	712.1	5.6	657.6	5.2
Male	N/I	N/I	N/I	N/I	2.6	0.1
Minority	N/I	N/I	N/I	N/I	336.5	5.8
Age 65-69	N/I	N/I	N/I	N/I	-709.6	3.0
70-74	N/I	N/I	N/I	N/I	-670.5	2.8
75-79	N/I	N/I	N/I	N/I	-729.2	3.1
80-84	N/I	N/I	N/I	N/I	-653.9	2.7
85+	N/I	N/I	N/I	N/I	-449.1	1.9
Medicaid	N/I	N/I	N/I	N/I	138.8	1.6
Disabled	N/I	N/I	N/I	N/I	120.6	0.5
Urban	N/I	N/I	N/I	N/I	453.1	2.8
LTCB	N/I	N/I	N/I	N/I	1,025.8	5.5
SNFB	N/I	N/I	N/I	N/I	-66.9	1.0
R^2	.072	N/A	.086	N/A	.098	N/A
N	16,076	N/A	16,076	N/A	16,076	N/A

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. CLM = Care Level Management; PBPM = per beneficiary per month; LTCB = long-term care beneficiaries; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

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

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

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

High Cost-High Risk: PBPM > $18,000 and HCC > 2.75 in base year.

High Cost: PBPM > $18,000 and HCC< 2.75.

High Risk: PBPM < $18,000 and HCC > 2.75.

LTCB, SNFB = 1 if beneficiary had long-term care hospital or SNF payments in base year.

N/I = not included; N/A means not applicable.

SOURCE: Medicare 200x-200Y Part A & B claims ; clm/9mo/Cost4b1 (9/10/09).

The second regression model controls for which cost-risk group the beneficiary was in during the base period. The key intervention coefficient is still insignificant. This is true even though two of the three cost-risk groups are much more costly than the neither group. The lack of effect is due to the initial balance of the intervention and comparison groups. The PBPM_base coefficient is even smaller, implying more RtoM. The large coefficient for the high cost-high risk group implies that the RtoM coefficient overstates the compression for this group and their costs do not fall quite as much as predicted. The large positive coefficient for the high risk group implies that the average RtoM effect would understate the cost increase for this initially lower cost group.

In the third model controlling for beneficiary characteristics, the intervention coefficient remains highly insignificant (-$17.5; t = 0.3). After controlling for the beneficiary's base year PBPM cost, the cost-risk group, and many other sociodemographic and utilization characteristics, we still find no cost-saving intervention effect on the trend in Medicare PBPM claims costs. Controlling for other beneficiary characteristics, the growth effects of the cost-risk groups decline somewhat but remain statistically significant. All age coefficients for the over-65 elderly are negative and significant, implying higher costs, on average, among the under-65 disabled population. Minorities in this very sick, costly, group were $337 more costly even controlling for their high cost, high risk status. Beneficiaries living in urban areas were more costly as well, ceteris paribus. Being admitted to a long-term care facility in the base period also added over $1,000 to monthly costs in the demonstration period.

Table 7-16 presents marginal effects on PBPM cost growth caused by average differences in beneficiary characteristics between the intervention and comparison groups. Intervention minus comparison (I - C) impacts are derived by multiplying I and C differences during the demonstration period by their corresponding regression coefficient (e.g., PBPM base cost impact = 0.194*[$3,296 - $3,360] = -$12.40; high cost/risk = $868*[0.385 - 0.401] = -$13.90). These coefficient-weighted effects approximate the amount of the overall cost difference in intervention and comparison group costs during the demonstration that is attributable to the difference in a particular beneficiary characteristic.

The difference in demonstration period PBPM costs was -$23 (see ***Table 7-5***). The "pure" intervention effect explains 75% (-$17.50 / -$23.00) of the -$23 difference, leaving 25% to be explained by unbalanced characteristics. Alternatively, unbalanced characteristics, including the difference in base period PBPM costs, explain only 25% ($5.50) of the demonstration period cost difference. The difference between base period mean PBPM costs can account for roughly 54% (-$12.40) of the -$23 difference in demonstration period PBPM costs between the intervention and comparison groups. The slight imbalance in the high cost–high risk group added another 60% to the overall greater costs in the intervention group. Together, the two effects explain 110% of the difference, but were offset primarily by lower frequencies of intervention beneficiaries aged 70 and 84. The relatively high number of Medicare disabled in the intervention group actually raised intervention costs somewhat (8.4% of the -$23 difference). The disabled are largely in the group of beneficiaries younger than 65 years that also constitutes a substantial number of the high cost and high risk groups whose effects are accounted for elsewhere in the model. The higher frequency of long-term care hospital beneficiaries in the intervention group as in the comparison group added nearly 18% to intervention costs.

Table 7-16
Marginal effects of beneficiary characteristics on CLM PBPM cost growth: Original population

Regression variable	Coefficient	I-value	C-value	I - C impact	% PBPM change
PBPMb	0.194	3296	3360	-12.4	54.0
High cost–high risk	868.2	0.385	0.401	-13.9	60.4
High cost	-23.1	0.285	0.288	0.1	-0.3
High risk	657.6	0.045	0.043	1.3	-5.7
Male	2.6	0.463	0.495	-0.1	0.4
African American	336.5	0.268	0.26	2.7	-11.7
Age 65-69	-709.6	0.162	0.157	-3.5	15.4
Age 70-74	-670.5	0.183	0.185	1.3	-5.8
Age75-79	-729.2	0.190	0.207	12.4	-53.9
Age80-84	-653.9	0.162	0.172	6.5	-28.4
Age 85+	-449.1	0.165	0.158	-3.1	13.7
Medicaid	138.8	0.089	0.094	-0.7	3.0
Disabled	120.6	0.125	0.109	1.9	-8.4
Urban	453.1	0.976	0.979	-1.4	5.9
LTCB	1,025.8	0.021	0.017	4.1	-17.8
SNFB	-66.9	0.262	0.264	0.1	-0.6

NOTES: PBPM = per beneficiary per month; CLM = Care Level Management; I = intervention; C = comparison; LTCB = long-term care beneficiaries; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries; I-C impact is difference between intervention and comparison group regression variables times regression coefficient; % change = I-C impact as share of total I-C difference in PBPM.

SOURCE: Medicare 200x-200Y Part A & B claims

7.8.2 Refresh Sample

In the first column of refresh results in ***Table 7-17,*** controlling only for each beneficiary's base period PBPM cost, the intervention coefficient of -27.0 is insignificant, implying no statistical difference between intervention and comparison groups in terms of average cost changes, ceteris paribus.

The base period PBPM cost coefficient (0.283; $p < .01$), when combined with the intercept coefficient, again implies substantial RtoM of costs in the refresh sample (= 0.336 - 1 = -0.664, the RtoM effect).

The second regression model controls for which cost-risk group the beneficiary was in during the base period. The key intervention coefficient switches sign (25.4) but is still quite insignificant. This is true even though two of the three cost-risk groups are much less costly than the neither group. The lack of effect of the high risk and cost groups is due to the initial balance of the intervention and comparison groups. The PBPM_base coefficient declines slightly, implying more RtoM within each of the cost-risk groups.

In the third model, controlling for beneficiary characteristics, the intervention coefficient changes sign once again but remains highly insignificant (-$18.60; t = 0.3). After controlling for the beneficiary's base year PBPM cost, the cost-risk group, and many other sociodemographic and utilization characteristics, we still find no cost-saving intervention effect on the trend in Medicare PBPM claims costs. Controlling for other beneficiary characteristics, the growth effects of the cost-risk groups decline somewhat but remain statistically significant except for the high cost–high risk group. All age coefficients for the over-65 elderly are negative but insignificant. Minorities in this very sick, costly, refresh group were $201 more costly even controlling for their high cost, high risk status and other factors. Beneficiaries living in urban areas and admitted to a long-term care facility in the base period also added substantially to monthly costs in the demonstration period over and above other factors.

Table 7-18 presents marginal effects on PBPM cost growth caused by average differences in beneficiary characteristics between the intervention and comparison groups. The difference in demonstration period refresh PBPM costs was -$26 (see ***Table 7-6***). The "pure" intervention effect explains 71% (-$18.60 / -$26.00) of the -$26 difference, leaving about 30% to be explained by unbalanced characteristics. Imbalances accounting for the $26 lower intervention cost are different than in the original sample. For example, the high cost–high risk group explained 60% of the difference in the original sample compared but only 1% in the refresh sample. The opposite is true for the high cost–only group that added less to costs in the refresh group due to a higher frequency than the comparison group (8.3% versus 7.8%). A higher frequency of high cost beneficiaries actually reduces cost increases due to severe RtoM in this group. Two other groups also contributing to smaller intervention cost increases include the Medicaid eligible and beneficiaries in a nursing home prior to the demonstration.

Table 7-17
Regression results: CLM intervention gross savings controlling for base period PBPM and beneficiary characteristics: Refresh population

Independent variable	PBPM_ Demo Coefficient	PBPM_ Demo t	PBPM_ Demo Coefficient	PBPM_ Demo t	PBPM_ Demo Coefficient	PBPM_ Demo t
Intercept	1,999	51.2	2,092	43.4	1,889	6.1
Intervention	-27.0	0.4	25.4	0.4	-18.6	0.3
PBPM_Base	0.283	36.0	0.272	26.2	0.268	23.7
High Cost-High Risk	N/I	N/I	75.3	1.0	244.0	3.0
High Cost	N/I	N/I	-680.4	6.2	-534.0	4.8
High Risk	N/I	N/I	-385.1	4.0	-264.2	2.7
Male	N/I	N/I	N/I	N/I	-162.7	2.9
Minority	N/I	N/I	N/I	N/I	200.7	3.2
Age 65-69	N/I	N/I	N/I	N/I	-404.5	1.5
70-74	N/I	N/I	N/I	N/I	-225.4	0.9
75-79	N/I	N/I	N/I	N/I	-192.0	0.7
80-84	N/I	N/I	N/I	N/I	-183.2	0.7
85+	N/I	N/I	N/I	N/I	-125.6	0.5
Medicaid	N/I	N/I	N/I	N/I	464.5	4.5
Disabled	N/I	N/I	N/I	N/I	-10.3	0.0
Urban	N/I	N/I	N/I	N/I	323.8	1.9
LTCB	N/I	N/I	N/I	N/I	759.7	3.6
SNFB	N/I	N/I	N/I	N/I	-302.0	3.6
R^2	.066	N/A	.069	N/A	.074	N/A
N	18,343	N/A	18,343	N/A	18,343	N/A

NOTES: Dependent Variable: Beneficiary's demonstration period PBPM cost. CLM = Care Level Management; PBPM = per beneficiary per month; LTCB = long-term care beneficiaries; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

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

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

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

High Cost-High Risk: PBPM > $18,000 and HCC > 2.75 in base year.

High Cost: PBPM > $18,000 and HCC< 2.75.

High Risk: PBPM < $18,000 and HCC > 2.75.

LTCB, SNFB = 1 if beneficiary had long-term care hospital or SNF payments in base year.

N/I = not included; N/A means not applicable.

SOURCE: Medicare 200x-200Y Part A & B claims; clm/9mo/Cost4b1 (9/10/09).

Table 7-18
Marginal effects of beneficiary characteristics on CLM PBPM cost growth: Refresh sample

Regression variable	Coefficient	I-value	C-value	I - C Impact	%PBPM Change
PMb	0.268	2763	2760	0.804	-3.1
High Cost/Risk	244	0.444	0.445	-0.244	0.9
High Cost	-534	0.083	0.078	-2.670	10.3
High Risk	-264.2	0.101	0.106	1.321	-5.1
Male	-162.7	0.471	0.474	0.488	-1.9
Black	200.7	0.292	0.291	0.201	-0.8
Age 65-69	-404.5	0.116	0.118	0.809	-3.1
Age 70-74	-225.4	0.172	0.166	-1.352	5.2
Age75-79	-192	0.195	0.195	0.000	0.0
Age80-84	-183.2	0.188	0.195	1.282	-4.9
Age 85+	-125.6	0.21	0.2	-1.256	4.8
Medicaid	464.5	0.082	0.089	-3.252	12.5
Disabled	-10.3	0.109	0.117	0.082	-0.3
Urban	323.8	0.975	0.973	0.648	-2.5
LTCB	759.7	0.019	0.019	0.000	0.0
SNFB	-302	0.171	0.157	-4.228	16.3

NOTES: CLM = Care Level Management; PBPM = per beneficiary per month; I = intervention; C = comparison; LTCB = long-term care beneficiaries; SNFB = skilled nursing facility beneficiaries; N = number of beneficiaries.

SOURCE: Medicare 200x-200Y Part A & B claims

7.9 Conclusion

PBPM costs showed considerable variability because of nature of the population selected for the demonstration, including a few very high cost beneficiaries with short spells of eligibility. Nevertheless, the over 11,000 original (and over 13,000 refresh) beneficiaries in the intervention group and 4,500 original (and 5,200 refresh) beneficiaries in the comparison groups are large enough to enable detection of an intervention savings rate of slightly over 5% of the comparison group's PBPM cost.

No statistically significant savings, however, were found for the intervention in either the original or refresh sample. Costs fell $41 slower (not faster, as required) in the original intervention group (1.6% of comparison costs), but savings needed to exceed $133 to be

considered statistically significant. Instead of offsetting its $295 monthly care management fee, CLM may have increased Medicare's costs to $336 per beneficiary per month.

CLM may have performed slightly better with its refresh sample because intervention costs increased $29 less than in the comparison group. This difference, however, was insignificantand savings needed to be $143 to be considered statistically significant.

Because CLM's intervention and comparison groups were randomly determined, no material imbalances were found across many cost, severity, and other patient characteristics in the base period. Consequently, any slight differences that did exist in the subsequent base year had little effect on our final conclusion of no significant savings.

Responding to CLM's request, CMS staff selected a very costly, complex set of Medicare beneficiaries for their intervention and comparison groups. Mean per beneficiary per month claims costs were approximately $3,300 in both groups, a figure several times higher than in the general Medicare population and at least twice as high as in other high cost beneficiary demonstration sites (e.g., Texas Senior Trails). As a result, the comparison group exhibited extreme RtoM effects: initially lower cost beneficiaries experienced large increases in their monthly costs during the demonstration and vice-versa for initially high cost beneficiaries. While the randomized experimental design should cancel out RtoM effects and isolate a pure intervention effect, the large churning of beneficiaries adds considerable statistical noise to the test of savings. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater is the potential for RtoM, the greater the effort required to identify those lower cost, lower utilizing, beneficiaries to avoid expensive hospitalizations in the near future. Part of the problem comes from using the prospective HCC score as a selection indicator. Although this score is based on cost weights that predict future costs, it may be biased in certain ways against identifying the chronically ill and favoring those with acute flare-ups. While HCC scores may correctly predict higher costs next period, on average, the higher the HCC score, the greater the reduction in a beneficiary's costs *even though costs still may be higher than average*. Moreover, the chronically ill have wide swings in monthly and even annual costs, albeit around a higher average than healthier Medicare beneficiaries. These swings in costs will be far greater for an elderly group than in a younger chronically ill population because of the much higher frequency of new, costly, diseases and problems (e.g., cancer, hip fractures). Successfully managing one chronic problem in a younger population can substantially reduce costs, but the same cannot be said with the same assurance in an elderly and disabled population.

CHAPTER 8
KEY FINDINGS FROM THE CARE LEVEL MANAGEMENT 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 Care Level Management (CLM) Medicare Care Management for High Cost Beneficiaries (CMHCB) demonstration program. Our evaluation focuses upon three broad domains of inquiry:

- **Implementation.** To what extent was CLM able to implement its program?
- **Reach.** How well did CLM engage its intended audience?
- **Effectiveness.** To what degree was CLM 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.

In this chapter, we present key findings based upon the 29 months of CLM operations with its original population and 18 months with its refresh population. Our findings are based on the experience of approximately 34,000 ill Medicare beneficiaries assigned to an intervention or a comparison group. Six key findings on participation, intensity of engagement in the CLM 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.

Key Finding #1: Several vulnerable subpopulations of Medicare FFS beneficiaries were less likely to agree to participate in the CLM demonstration program.

Of all CLM intervention beneficiaries, 65% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. We found that Medicaid enrollees and institutionalized beneficiaries were less likely to be participants; both groups are costly and high users of acute care services. In general, participants tended to be healthier than nonparticipants using baseline characteristics including the prospective HCC score. However, beneficiaries with higher *concurrent* HCC scores based on the first 6 months of the demonstration were more likely to participate than healthier beneficiaries. This suggests that CLM made some inroads into engaging those with acute clinical deterioration. Further, as CLM's program matured, they appeared to be more successful engaging sicker and more costly beneficiaries based on baseline health status; however, those with Medicare/Medicaid dual enrollment and the institutionalized were still less likely to become participants. These findings suggest alternative recruiting and outreach strategies are needed to reach dual Medicare/Medicaid enrollees and beneficiaries who are institutionalized.

Key Finding #2: CLM's physician home intervention (PHI) was limited to less than one-third of their participating beneficiaries.

A cornerstone of CLM's program was physician home intervention (PHI) visits including home hospitalization services. We found some evidence that CLM made a focused effort to contact beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program. Yet, only 30% of their fully participating beneficiaries received a PHI visit during a 12-month period. When we add in other types of visits – including routine follow-up care, we do observe a higher level of physician interaction. However, 25% of beneficiaries received no physician visits. Telephone contact was the most dominant "frequent" form of contact.

Key Finding #3: CLM improved some aspects of beneficiary reported experience with care, level of physical activity, and self-reported physical health.

The beneficiary survey was designed to obtain assessments directly from beneficiaries about key outcomes of beneficiary ***experience of care, self-management***, and ***physical and mental function.*** We asked beneficiaries about the extent to which their health care providers helped them to cope with their chronic condition. We supplemented this item with questions related to two key components of the CLM 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.

Among the 19 outcomes covered by the survey, four statistically significant positive intervention effects were found—discussing treatment choices, communication with their health care providers broadly defined, 30 minutes of continuous physical activity, and, most notably, physical health. However, there was no improvement in overall beneficiary satisfaction that their health care team helped them deal with their chronic condition in spite of the positive intervention effects of two related experience with care measures—helpfulness and the quality of discussions with their health care team and improvement in self-reported physical health.

A positive intervention effect of 30 minutes of continuous physical activity did not translate into greater *confidence* on the part of the CLM population in exercising 2 to 3 times a week. The mean number of days of 30 minutes of exercise was 2.8 in the comparison group and 3.2 in the intervention group.

We used the RAND-12 scoring algorithm 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 CLM intervention population had a mean PHC score of 30 versus 28 in its comparison group, a statistically significant difference. Further, a 2 percentage point increase is generally viewed as a clinically meaningful improvement. There was

no difference in mental health functioning as a result of the CLM intervention. Nor was there was any difference in abilities to do activities of daily living.

Key Finding #4: CLM had a positive intervention effect on one of five quality of care process measures.

We have defined quality improvement for this evaluation as an increase in the rate of receipt of claims-derived, evidence-based process-of-care measures (e.g., serum cholesterol testing) and improvement in health outcomes as a reduction in the rate of hospitalizations, readmissions, and emergency room (ER) visits, and a reduction in mortality rates. We find no evidence of systematic improvement in quality of care in the CLM CMHCB demonstration program. Out of five measures, there was only one observed increase in rate of receipt of evidence-based care (influenza vaccination). We observe this increase in both the original and refresh population and during the last 12 months of CLM's operations.

Over the course of the demonstration, CLM had expected to increase rate of adherence to evidence-based care. However, during the last year of their demonstration program, we observe lower rates of adherence to the selected measures among their intervention beneficiaries than we do among the comparison group beneficiaries for all measures with the exception of influenza vaccination. We also observe between roughly one-third to one-half of intervention beneficiaries in both the original and refresh populations were not compliant during the last year of the CMHCB demonstration despite focused efforts by CLM to encourage beneficiaries to become compliant with evidence-based care. As noted above, over 80% of intervention beneficiaries with COPD were not compliant with annual oxygen saturation assessment. These findings suggest that improving or sustaining adherence to guideline concordant care in a cohort of ill Medicare FFS beneficiaries was more challenging than originally envisioned.

Key Finding #5: During the last 12 months of the demonstration program, CLM had some success in reducing acute care utilization in both the original and refresh population. CLM had no success reducing mortality in either the original or refresh population.

During the first half of program operations, CLM was not successful in reducing acute hospitalizations, ER visits, or 90-day readmissions in their original intervention population. But during the last 12 months of the demonstration, CLM was modestly successful in reducing the percent and rate of admissions for ten ambulatory care sensitive conditions (ACSCs). Most notably, CLM achieved a sizeable reduction in the all-cause readmission rate of -225 per 1,000 beneficiaries. During that same time period, CLM also had some modest success in reducing the acute care utilization in its refresh population. Rates of all-cause and ACSC hospitalizations declined, as did the percent with a readmission for all causes or an ACSC.

CLM was not successful reducing rates of ER visits in either time period for the original population or in the last 12 months of the demonstration for the refresh population. CLM had no success reducing mortality in either the original or refresh population.

Over the course of the first year of operations, CLM reported that they modified their program in an effort to identify, in real time, participants whom they believed would most benefit from their interventions by changing how they stratified beneficiaries into levels of visit

urgency. CLM also reported that they reorganized their patient care teams to include more nursing support. CLM believed that this arrangement would allow patients to bond with the nurse care manager over time, whereas CLM had observed that the clinical specialists were not able to forge a sufficient bond as evidenced by the fact that some of their participants were going to the hospital rather than calling the clinical specialists when problems arose. CLM also reported that they felt the members of the refresh population had an illness profile that made them more appropriate to the CLM program. The data on health outcomes suggests for the original population that the program modifications had the desired effect of reducing some acute care utilization, but not ER visits; while the data on health outcomes for the refresh population shows that CLM was more successful reducing acute care utilization at an earlier stage in their demonstration period than for the original population.

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

No statistically significant savings were found for the intervention in either the original or refresh populations. Costs fell $41 slower (not faster as required) in the original intervention group (1.6% of comparison costs) but savings needed to exceed $133 (5%) to be considered statistically significant. Instead of offsetting its $295 monthly care management fee, Care Level Management may have increased Medicare's costs to $336 per beneficiary per month.

CLM performed slightly better with its refresh population as intervention costs increased $29 less than in the comparison group. This difference, however, was highly insignificant, as savings needed to be $143 to be considered statistically significant. Among refresh participants, alone, monthly Medicare costs did decline significantly (-$170) but were offset by large increases among non-participants who were one-third of the entire intervention population. It is not possible to rule out the selection of beneficiaries who are more responsive to the intervention in explaining success only with participants.

Multivariate regression was used to control for any imbalances in the intervention and comparison groups prior to the disease management intervention. No cost savings were found after adjusting for minor imbalances in the two groups.

Base year per beneficiary per month claims costs averaged $3,300 in both groups, a figure several times higher than in the general Medicare population. As a result, the comparison group exhibited extreme regression-to-the-mean effects with costs declining $859, on average, during the demonstration period. At the same time average group costs were falling, initially lower cost beneficiaries saw their costs rising by several thousand dollars during the intervention period. The greater is the potential for regression-to-the-mean in either direction, the greater is the challenge facing care management groups in identifying the appropriate beneficiaries to target for intervention.

8.1 Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that CLM had limited success in improving key processes of care, beneficiary experience with care, self-management, or functional status, and reducing hospital admissions. CLM was most successful at reducing 90-day all-cause readmissions by -225 per 1,000 among its original

beneficiaries. However, the overall set of modest improvements were achieved at substantial cost to the Medicare program in the form of monthly management fees ($58 million) with no demonstrable savings in program outlays on health services. Despite the limited gains, the lack of program savings to offset monthly management fees cannot justify the CLM model for chronically ill Medicare fee-for-service beneficiaries on cost effectiveness grounds.

What might explain the lack of success in CLM's demonstration program? One explanation may be the inability to accurately target beneficiaries at greatest risk of intensive, costly, service use (as distinct from the need for general care management). A cornerstone of CLM's program was physician home intervention (PHI) visits including home hospitalization services with a focus on ambulatory care sensitive conditions (ACSCs). Yet, only 30% of their fully participating beneficiaries for a 12-month period received a PHI visit. When we add in other types of visits—including routine follow-up care—we do observe a considerably higher level of physician interaction. However, 25% of beneficiaries received no physician visits and one-half of all beneficiaries received less than five visits during a 12-month period. Telephone contact was the most dominant "frequent" form of contact.

In our multivariate regression modeling of likelihood of being in a high contact versus low contact group, we found some evidence that CLM made a focused effort to interact with beneficiaries who were at high risk of hospitalization or who had been hospitalized, a key stated component of their program. Given CLM's high monthly management fee (almost $300 per month) and the population-based financial risk feature of this demonstration, the concentration of physician visits, in general, and PHI visits, in particular, suggests that CLM would have had to have been extremely successful in reducing costs associated with the beneficiaries they were targeting. Descriptive analyses showed that the rates of all-cause and ACSC hospitalizations during the demonstration were higher among beneficiaries who received PHI visits or had more than 20 contacts as compared to beneficiaries with no PHI visits or who received fewer than 20 contacts. The pattern was consistent across both the original and refresh populations.

Prior evaluations of Medicare care management programs that were primarily telephonic have not demonstrated savings sufficient to cover fees one-third the size of CLM's fee. CLM was successful in only modestly reducing hospitalizations during the last 12 months of demonstration operations, with no particularly greater success for ACSC hospitalizations. CLM was more successful at reducing readmissions but only among its original population. The lack of substantive improvements in acute care utilization broadly across their intervention population translated into limited financial savings. And, their targeting strategy was costly. Each contact cost was roughly $266, or over twice the national average payment for a face-to-face office visit with an established patient with the ***highest level of complexity*** under the Medicare Fee Schedule[21].

CLM's lack of success is not surprising in light of the extreme regression-to-the-mean (RtoM) behavior that we observed among their beneficiaries who had been selected based upon high prior costs or high prior rates of hospitalizations. Armed with data on beneficiary disease, utilization, and cost profiles in the base period, CLM focused first on those most likely to be

[21] National non-facility price of $124.79 for HCPCS code 99215 for 2009.

major users of acute care services. Yet, many of these beneficiaries experienced declines in use and costs regardless of the intervention, as evidenced in the control group. The large increases in demonstration period costs in otherwise less costly beneficiaries in the base period make it very difficult for intervention staff to target those at highest risk of increasing costs. In fact, the greater the potential for RtoM, the greater the effort required to identify those lower cost, lower utilization, beneficiaries to avoid expensive hospitalizations in the future. Targeting the group who had a high number of hospitalizations in the base period focused extensive management resources on many "false positive" beneficiaries, who ultimately did not need nearly as many costly services as they did in the year prior to the demonstration.

The quixotic, immediate, nature of many elderly diseases calls for real time information on health status. Unable to predict future health status with any precision for those with initially stable, less costly, conditions, and lacking direct access to patients' medical records, the CLM physicians and nurses often began working with beneficiaries with incomplete information. Further, the CLM physicians and nurse care managers were not part of the beneficiaries' primary health care teams, further hindering their ability to directly interact with the beneficiaries' primary care providers and effectively help facilitate changes in medical care plans to mitigate deterioration in health status. It is not surprising that CLM was unable to successfully improve patient self-management.

Because targeting care management resources is so difficult with the elderly, and errors so costly, the way in which the clinical team communicates and interacts with them is extremely important. Yet, another possible reason why CLM was ineffectual has to do with the limitations of CLM's personal visit physician (PVP) and nurse care manager model. The PVP served only as an adjunct to the patients' primary care physicians with a stated goal of facilitating the relationship between the patient and his or her community-based provider with a focus on chronic issues. The PVP consulted with community-based providers if significant changes in treatment regimens were indicated and to ensure that they both were implementing a common care plan. Nurse care managers worked with the PVPs to telephonically help coordinate care services with patients, family members, community-based primary care physicians and specialists, and home health nurses. By complementing, not substituting, for the primary care physician, the PVPs and 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. Moreover, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments, and the most dominant form of contact was telephonic.

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.

Barnett, A.G., van der Pols, J., and Dobson, A.: (2005). Regression to the mean: what it is and how to deal with it. Int. J. Epidemiol. 34:215-220, 2005.

Brody, E.R., and Bernard, S.L.: Evaluation of the Medicare Care Management For High Cost Beneficiaries demonstration: Final site visit report on Care Level Management. Prepared for the 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.

Gelman, A., and Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge and New York. Cambridge University Press, 2007.

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.: Economic Analysis. Upper Saddle River, N.J. Prentice-Hall, 2000.

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., & Bernard, S.: Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Second Annual Report. Prepared for the 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.: Fundamentals of Biostatistics. Pacific Grove, Calif. Duxbury Thomson Press, 2006.

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.

Stigler, S.: The History of Statistics: The Measurement of Uncertainty before 1900. Cambridge, Mass. Harvard University Press, 1986.

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.