February 2010

Evaluation of Medicare Care Management for High Cost Beneficiaries (CMHCB) Demonstration: Texas Senior Trails (TST)

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: TEXAS SENIOR TRAILS (TST)

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

February 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 Texas Tech University Health Sciences Center and its Texas Senior Trails (TST) 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 TST 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), TST 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 TST able to implement its program?
- *Reach.* How well did TST engage its intended audiences?
- *Effectiveness.* To what degree did TST 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

Texas Senior Trails launched its CMHCB demonstration program April 1, 2006. During the first year of operations, CMS offered TST a refresh population at twelve months post launch that would partially offset attrition due primarily to death. TST declined to accept a refresh

population. During the second year of operations, TST requested early termination of its CMHCB program. The TST program ended July 31, 2007, or 16 months after initiation.

Texas Senior Trails received monthly management fees for its full intervention population for the first six months of engagement. After the initial 6-month outreach period, TST accrued management fees 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 TST. 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 TST after a sentinel event, such as a hospitalization or an emergency room visit.

Texas Senior Trails negotiated a per beneficiary per month payment of $117 for the duration of the demonstration. At the end of the 16-month period, TST was contractually obligated to achieve a 5% savings in Medicare payments among the intervention group (regardless of participation in the TST program) compared with the comparison group and to cover the TST program fees collected. In addition, TST had the opportunity to share a portion of any savings beyond 5% that were achieved. Payments associated with the Medicare Part D benefit were not included in these calculations.

E.2 Overview of the TST CMHCB Demonstration Program

The overarching goal of the TST program was to help participants take an active role in their health and receive timely access to appropriate health and social services. TST used a holistic approach to deliver its multidisciplinary care management intervention to help coordinate health and social services for participants with multiple comorbid chronic conditions. Participants also had access to a variety of other services provided by nurse care managers and staff from TTUHSC's schools of medicine, nursing, and pharmacy.

Risk assessments of the TST participant population were conducted to inform the development of individualized care plans and assign participants to one of three risk categories to determine the level of service to be provided to each participant. TST reported that beneficiaries assigned to the low-risk intervention received support, composed primarily of educational information, from nurse care managers during monthly or quarterly telephone interactions. Beneficiaries at risk for repeated hospitalizations or nursing home placement received the medium-risk intervention (i.e., telephonic care management support conducted every 2 weeks to every 2 months). TST reported that the high-risk intervention was provided to approximately 5% to 7% of the TST participant population that had depression and/or potentially critical health problems that required immediate attention. These individuals were reported to have received support during telephone calls and in-person visits with nurse care managers.

Each participant was assigned to a care team composed of nurses employed by TrailBlazer, who provided care management support via telephone from Dallas, Texas, and nurse care managers employed by TTPA and located in Lubbock and Amarillo, who conducted on-site visits with beneficiaries at their homes, physician offices, or in the hospital, as needed. TST participants also received support from one of three social workers who addressed issues such as

transportation, housing, mental health problems, and preparing advance directives to prepare for end-of-life care needs.

One of the primary goals of the TST CMHCB program was to facilitate patient relationships with primary care physicians and specialists so that patients could receive appropriate preventive care to manage their illnesses and become informed consumers of health care. TST care managers educated beneficiaries about physician resources in their areas to help those individuals who had become habituated to receiving care at the emergency room. To further support patient relationships with physicians, care managers helped beneficiaries prepare for office visits by discussing issues that had arisen and helping patients identify questions for the doctor. TTPA care managers who provided onsite support were available to attend physician visits with patients who needed hands-on assistance. TST also encouraged participants, particularly those in nursing homes, to contact their TST care manager prior to seeking care at the emergency room.

E.3 Key Findings

In this section, we present key findings based upon the 16 months of TST operations. Our findings are based on the experience of approximately 10,000 ill Medicare beneficiaries assigned to an intervention or a comparison group. Five key findings on participation, beneficiary satisfaction, clinical quality and 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 TST demonstration program.

Of all TST intervention beneficiaries, 91% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. In spite of this high participation rate, we found that the participant population was healthier and younger than beneficiaries who never participated. 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. Given that TST was both provider- and hospital-based, it is surprising that they were unable to get the sickest eligible beneficiaries to participate. These findings suggest alternative recruiting and outreach strategies are needed to reach the sicker beneficiaries as well as dual Medicare/Medicaid enrollees and beneficiaries who have disabilities or are institutionalized.

Key Finding #2: The intensity of intervention with the participating beneficiaries is unlikely to produce significant behavioral change and savings.

Although there was no pre-determined expected number of contacts, the TST beneficiaries were a sick and costly group of FFS beneficiaries, and they reported significant unmet clinical and psychosocial needs. Across a 12-month intervention period that reflects maximum intervention months for fully eligible and participating beneficiaries, the median number of contacts was four. We found some evidence that TST targeted their intervention contacts to beneficiaries who were at risk of hospitalization or who had been hospitalized.

However, we found limited intervention services for all beneficiaries other than developing a care plan or a discharge plan. Given the low level of interaction with many of their participants, it is unlikely that they could be successful at changing beneficiary behavior with respect to self-management of their chronic illness leading to a reduction in use of acute care settings and savings.

Key Finding #3: There was no improvement in beneficiary satisfaction, experience with care, self-management, or physical and mental health functioning.

The TST CMHCB demonstration program employed strategies to improve quality of care while reducing costs by empowering Medicare beneficiaries to better manage their care. They did so 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 and should encourage more effective interactions between beneficiaries and their primary health care providers. TST hypothesized that lifestyle changes and better communication with providers would mitigate acute flare-ups in 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.

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. Among the 19 outcomes covered by the survey, only one statistically significant, counterintuitive, group difference was found: members of TST's intervention group were *less* (not more) certain that they could take their medications as prescribed. There was no improvement in beneficiary satisfaction, care experience, self-management, or physical and mental health functioning.

Key Finding #4: TST had a positive intervention effect on one of five quality of care process measures but no positive intervention effect on reduction in rates of hospitalization, emergency room visits, 90-day readmissions, or mortality.

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 TST CMHCB demonstration program. Out of five measures, there was only one observed increase in rate of receipt of evidence-based care (influenza vaccination) while at the same time there was a decrease in the rate of receipt of another process-of-care measure (oxygen saturation assessment).

Over the course of the demonstration, TST had expected to increase beneficiaries' rate of adherence to evidence-based care. Yet, between one-quarter to one-half of beneficiaries were not compliant with evidence-based guidelines during the last year of the CMHCB demonstration despite focused efforts by TST to encourage beneficiaries to become compliant. 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.

We did not find evidence of improvement in health outcomes as measured by a reduction in acute hospitalizations, 90-day readmissions, ER visits, or mortality. In fact, there was a statistically significant increase in the rate of all cause and ambulatory care sensitive condition (ACSC) hospitalizations among the intervention group relative to the comparison group. Although we observed a statistically higher mortality rate in the intervention group during the demonstration, multivariate analysis that controlled for imbalances in baseline health status between the two groups equalized the raw mortality rates.

Key Finding #5: Medicare cost growth was greater in the intervention group than the comparison group.

No Medicare savings were found for the intervention. In fact, cost growth was statistically greater ($123 per beneficiary per month), not less, in the intervention group compared with the comparison group. This increase was not due to nonparticipants in the intervention as participants had even slightly higher cost growth, on average ($126).

Not only did TST fail to achieve budget neutrality using RTI's evaluation method, it also doubled the $117 cost of Medicare's monthly per beneficiary management fee. If one believed that, in fact, TST's intervention actually increased beneficiary costs by $123, then the net effect on Medicare costs would be $123 plus $117, or $240 per beneficiary per month. Instead of reducing the 8.5% fee outlay on per beneficiary per month (PBPM) costs of the comparison group, the intervention would appear to have increased Medicare total costs from 8.5% to 17.4% of the comparison group's costs.

Multivariate regression methods were used to statistically control for both regression-to-the-mean and any biases against the intervention caused by initial imbalances in several beneficiary characteristics. The control variables had no material effect on the difference in growth rates in the two groups. Intervention cost growth remained statistically greater, not less, than in the comparison group. PBPM costs were nearly 14% greater in the intervention group than in the comparison group during the 16-month intervention period, compared with being only 5% higher in the base year. Because demonstration beneficiaries were selected, in part, because of their high pre-demonstration costs, both intervention and comparison groups experienced significant regression-to-the mean effects. Beneficiaries with initially high costs experienced large declines during the demonstration period while initially lower cost beneficiaries saw even larger increases, percentagewise. Yet, only 30% of the higher demonstration period costs could be explained by regression-to-the-mean or imbalances in beneficiary characteristics. Fully 70% would appear to be attributable to the more costly intervention. Where imbalances were large in relative terms, they were small percentagewise (e.g., 1% versus 3% of the population on Medicaid), had minor effects on cost *growth* (as distinct from initial cost *levels*), or were in favor of the intervention.

E.4 Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that

TST had limited success in improving key processes of care and no success in improving beneficiary satisfaction, self-management, or functional status and reducing hospital admissions, readmissions, emergency room visits, or mortality. The one process improvement success was done so at substantial cost to the Medicare program in the form of monthly management fees ($7.9 million) with no demonstrable savings in program outlays on health services. With only one statistical success, in the rate of influenza vaccination, the cost per successful improvement is $7.9 million or $2 million per significant percentage point improvement in a population of 5,000 FFS beneficiaries, or $400 per intervention beneficiary.

Despite a limited gain in one process-of-care measure, the lack of program savings to offset monthly management fees cannot justify the TST model for chronically ill Medicare fee-for-service beneficiaries on cost effectiveness grounds.

What might explain the lack of success in TST's demonstration?

Ineffective Targeting. 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). TST selected two large geographic areas in Texas that were medically underserved and where Texas Tech University Health Sciences Center (TTUHSC) and Texas Tech Physician Associates (TTPA) provided health care to a large proportion of the population in this area that had limited primary care access and high emergency room usage. Emergency departments are a significant source for hospital admissions in the target areas and therefore make important contributions to hospital profit margins. Thus, the TST program focused on (1) reducing emergency room utilization and (2) providing hospital discharge planning support to reduce readmissions.

To implement their care management strategy, each participant program was assigned to a care team composed of nurses who provided care management support via telephone from Dallas, Texas, and nurse care managers located in Lubbock and Amarillo, who conducted on-site visits with beneficiaries at their homes, physician offices, or in the hospital, as needed. When TST learned that one of its participants was admitted to the hospital, it reassigned this individual to its high-risk intervention, and when appropriate, a care manager visited the beneficiary in the hospital to determine the cause of the hospitalization and identify any new health or social issues to be addressed. Not surprisingly, TST adopted a strategy of targeting beneficiaries at greatest risk of a hospitalization and higher costs. Their targeting strategy was unsuccessful—and costly.

Using the total number of contacts with their intervention population reported by TST and $7.9 million in management fees, we calculate that each contact cost was roughly $332, or over 5 times the national average payment for a face-to-face office visit with an established patient with moderate complexity under the Medicare Fee Schedule.

Texas Senior Trails's lack of success is not surprising in light of the substantial regression-to-the-mean behavior that naturally occurs with the elderly chronically ill. Armed

with data on beneficiary disease, utilization, and cost profiles in the base period, health coaches were instructed to focus 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. Targeting this group 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 program was unable to predict future complications with any precision for those with initially stable, less costly, conditions. Lacking direct access to patients' medical records, the health coaches often began working with beneficiaries with incomplete information. Further, the health coaches 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 TST was unable to successfully improve patient self-management.

Because real time information on health status was not available on beneficiaries as their health declined, TST care managers focused on identifying beneficiaries at the time of discharge from the hospital. At that point, major inpatient costs had already been incurred along with a pre-ordained stream of post-acute care institutional and physician services. Unfortunately, even then, we find no evidence of success in reducing readmissions. It is not clear that any *ex post* success after discharge would have been sufficient, by itself, to save Medicare enough to justify the management fees negotiated by TST.

"Remote" Nurse-Patient Communication. 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 TST was ineffectual has to do with the limitations of remote nurse call centers in managing beneficiary utilization of health services. By complementing, not substituting, for the primary care physician, health coaches worked to change beneficiaries' lifestyles and encouraging them to take their medications, to track key health indicators such as blood sugar levels, and to follow physician orders. While all clinicians agree these are valuable adjuncts to successful self-care management, having nurses act as health coaches, listen on telephones, and make suggestions is not the same as directly determining whether a patient is admitted to a hospital and having a personal relationship cultivated over several years. Nor can call-center nurses have much effect on service intensity once patients are admitted for care. Nurses on telephones cannot admit patients, tell them to fill a prescription right away, or instruct them to "start exercising now or risk a heart attack in the next few months." Moreover, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments.

Latent Health Needs. A final possible reason why TST may have not been successful in reducing acute care utilization and Medicare program costs is the selection of medically underserved geographic areas. It may be reasonable to assume that once the health coaches conducted health assessments, they uncovered substantial unmet health care needs. By facilitating the patient-physician relationship, the health coaches may have actually increased health care utilization.

In light of the results of other CMS disease management demonstrations that relied to varying degrees on this model of care management, it is unlikely that simply managing elderly patients “at a distance” via the telephone or an occasional in-person visit will produce the kinds of savings for which policymakers had hoped. To the extent that detailed health assessments identify significant unmet health care needs, costs are likely to increase not decrease. It also suggests that to be effective at all an intervention may require intensive personal clinical attention, which costs money. A win-win scenario, in which better quality is achieved at lower overall cost, is not likely.

CHAPTER 1
INTRODUCTION TO THE MEDICARE CARE MANAGEMENT FOR HIGH COST BENEFICIARIES (CMHCB) DEMONSTRATION AND THE TEXAS SENIOR TRAILS (TST) 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 Texas Tech University Health Sciences Center and its Texas Senior Trails (TST) 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 group's 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 population 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 startup, 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. Because TST requested early termination from the CMHCB program after 18 months of operations, RTI did not conduct a follow-up site visit. RTI requested but TST did not make their senior leadership team available for a final telephone interview. The site visit findings were reported to CMS in a stand alone report (Brody and Bernard, 2007).

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. TST'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 18 months of the TST CMHCB program operations with their original population. TST did not accept a refresh population. We start by reporting on the degree to which TST was able to engage their intervention population. We measure degree of engagement in two ways: (1) participation rates and characteristics of participants; and (2) number and nature of contacts between TST and participating beneficiaries from encounter data provided to RTI from TST. We then report findings related to the effectiveness of TST to improve beneficiary and provider satisfaction, improve functioning and health behaviors, improve clinical quality and health outcomes, and achieve targeted cost savings.

1.2 TST's CMHCB Program Design Features

1.2.1 TST Organizational Characteristics

Texas Senior Trails was a CMHCB demonstration program run by a consortium of three organizations: Texas Tech University Health Sciences Center (TTUHSC), TrailBlazer Health Enterprises, LLC, and Texas Tech Physician Associates (TTPA). The consortium was formed for the sole purpose of developing and implementing a CMHCB program to deliver care management services to high-cost Medicare beneficiaries in 48 counties in the Texas Panhandle. The organization's culture embraced examining issues from a variety of perspectives in order to find creative solutions, as illustrated by its motto: "We can't solve problems by using the same kind of thinking we used when we created them." A description of each participating organization and its role is provided subsequently.

Texas Tech University Health Sciences Center—Created by the Texas legislature in 1969, TTUHSC has trained more than 7,000 health care professionals. TTUHSC, a university located on four campuses (Amarillo, El Paso, Lubbock, and Permian Basin), comprises the following five schools: School of Medicine, School of Nursing, School of Pharmacy, School of Allied Health Sciences, and the Graduate School of Biomedical Sciences. TTUHSC served as TrailBlazer's clinical partner in the TST consortium. TTUHSC management staff provided administrative oversight for the TST program, and physician leadership from TTUHSC served on utilization management committees in Lubbock and Amarillo that provided clinical oversight for the program. TTUHSC physicians, as well as other physicians who cared for TST participants, received information about the health status of their patients from TST care managers and feedback about performance from reports compiled by TrailBlazer.

TrailBlazer Health Enterprises, LLC—A wholly owned subsidiary of BlueCross/BlueShield of South Carolina, TrailBlazer Health Enterprises (referred to as "TrailBlazer") administers the Medicare program under contracting arrangements with CMS. At the outset of the demonstration, TrailBlazer served as both a Part A intermediary and a Part B carrier. TrailBlazer administers some aspect of the Medicare program for beneficiaries and providers in nearly every state in the nation. In Texas, TrailBlazer was the fiscal intermediary for Medicare Part A, the regional fiscal intermediary for Part A end-stage renal disease (ESRD) claims, and a carrier for Medicare Part B.

For the TST CMHCB program, TrailBlazer shared financial risk of the program with TTPA and employed care managers who provided telephonic care management support to program participants. TrailBlazer led data management and reporting efforts for TST because of its extensive data analysis capabilities as the State of Texas fiscal intermediary.

Texas Tech Physician Associates—TTPA is a physician group practice owned by TTUHSC that operates as a private, nonprofit entity. In 1995, TTPA was formed to allow TTUHSC to enter into risk-bearing agreements for physician and other medical services. The health maintenance organization (HMO) risk program co-owned by TTPA ended in 2002, at which time TTPA refocused its role as a health care provider and financial supporter of TTUHSC. All physicians at the TTUHSC School of Medicine who have a permanent Texas license to practice medicine are members of TTPA. For the TST CMHCB program, TTPA served as the risk contractor with primary responsibility for the program's financial goals. In addition, TTPA employed care managers who provided on-the-ground support for high-risk CMHCB participants in Lubbock, Amarillo, and surrounding areas.

1.2.2 Market Characteristics

Texas Senior Trails selected 48 counties in Western Texas, known as the Texas Panhandle and South Plains, as its target region for the CMHCB program for the following reasons:

- TTUHSC and TTPA provide health care to a large proportion of the population in this area,
- the area is medically underserved, and
- no other Medicare fee-for-service demonstrations or chronic care improvement programs was operating in the region.

The 46,355 square mile area that comprises the TST target region has a population of 837,898, including 122,000 Medicare beneficiaries. The Texas Panhandle is rural and remote, with a population density of 18.1 people per square mile. Average per capita income is $24,495, 16.4% of the population lives below the poverty level, and 1 of every 4 people does not have health insurance. The area is racially diverse, with a population that is 28.4% Hispanic, 5.6% African American, 64.5% Caucasian, and 1.5% of another race.

TTUHSC is the largest multispecialty practice in the area, competing primarily with one other large health system and a significant number of independent practitioners in the area. There are two hospital systems in Lubbock and two in Amarillo, which are supplemented by smaller rural hospitals. In Lubbock, TTUHSC has a close relationship with the University Medical Center (UMC), located adjacent to the TTUHSC campus. The UMC functions as a university hospital in that most of the TTUHSC faculty members serve as clinical department heads at UMC and most of TTUHSC's residents work at UMC. UMC is also a county hospital that serves a high proportion of indigent patients. In Lubbock, TTUHSC also works with Covenant Health System on a limited basis. In contrast, TTUHSC works with all the other hospitals in the Amarillo area, notably Northwest Texas Health Care System and Baptist St. Anthony, and

TTUHSC physicians in this area have privileges at multiple hospitals. In sum, 85% of UMC patients, 55% of Northwest patients, and 25% of Baptist St. Anthony patients receive care from TTUHSC.

TST chose to offer its CMHCB program to Medicare beneficiaries in the Texas Panhandle and South Plains because the area offered significant opportunities to improve the quality of health care delivered. These opportunities are illustrated by the following factors that contribute to the high rates of emergency room utilization in the area:

- Of the 48 counties served by TST, 40 are designated by the U.S. Health Services and Resources Administration (HRSA) as primary care shortage areas. As a result, a significant proportion of the population does not have an established relationship with a primary care provider and, therefore, typically seeks care at the emergency room.

- Transportation is a significant issue for many Medicare beneficiaries in the area. Medicare does not provide financial support for transportation to physician offices, so some individuals may opt to obtain care from an emergency department because Medicare will pay for ambulance transportation to the emergency room.

- The large hospitals in the area actively market their emergency room services to community residents and providers. Emergency departments are a significant source for hospital admissions and therefore make important contributions to hospital profit margins. Direct-to-consumer marketing has successfully maintained high levels of emergency department usage, even among patients who have primary care providers, especially in the Amarillo area.

- Finally, several physician practices in Amarillo that specialize in providing care to nursing home residents typically order residents to be transferred to emergency departments when health issues arise rather than deploying nurse practitioners or physician assistants to the facility to evaluate the patient.

1.2.3 TST Intervention and Comparison Populations

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

- was Medicare fee-for-service beneficiary with a primary residence in one of the designated counties of Texas and

- had high costs during the 12-month period between September 1, 2004, and August 31, 2005 (i.e., Medicare costs greater than or equal to $6,000 in 2004) **or** had high disease severity as indicated by Hierarchical Condition Categories (HCC) risk scores greater than or equal to 1.7 and

- had a majority[1] of visits to TTUHSC or UMC's primary care practice, Physician Network Services (PNS) as evidenced by claims for health care services obtained between September 1, 2004, and August 31, 2005.

The population was further restricted using the following exclusion criteria:

- receipt of hospice care,
- enrollment in the ESRD benefit or receipt of dialysis,
- enrollment in a Medicare Advantage (MA) plan,
- use of Medicare as a secondary payer, or
- lack of Medicare Part A and Part B coverage as of January 2, 2006.

Using these inclusion and exclusion criteria, a total of 5,117 Medicare beneficiaries were assigned to the TST intervention group.

The CMHCB demonstration program was designed using an intent-to-treat (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 are assigned to either the intervention or comparison group, they remain in their assigned group for all days in which they are eligible. Eligibility for the TST program and hence membership in either the intervention or comparison group is lost for the period(s) if the beneficiary

- enrolls in an MA plan,
- loses eligibility for Medicare Part A or B,
- gets a new primary payer (i.e., Medicare becomes the secondary payer),
- develops ESRD, or
- dies.

Comparison group—Following the development of the intervention group criteria, TST worked with CMS and RTI to develop specifications to select a comparison group of beneficiaries to be used in conducting the financial reconciliation and evaluation of this CMHCB program. A physician group practice-based (PGP) approach was used to select the comparison group from non-TTUHSC providers, an approach that parallels the procedures used to identify beneficiaries in the intervention group. The PGP-based approach requires a systematic method

[1] "Majority" here applies to those individuals with two or more visits to TTUHSC, unless they had only two or three visits to TTUHSC and nine or more to non-TTUHSC providers.

for identifying appropriate comparison PGPs. A claims-based approach was developed that identified high volume PGPs through their Tax Identification Numbers (TINs) and the selection of four TINs that had similar distribution of physician payments across major types of services as TTUHSC. The physician group practices had to have at least 20% of their payments for at least one of the following categories of services, which were the major types of services offered by TTUHSC: office visits, hospital visits, or routine diagnostic laboratory testing. Each selected practice also had to provide a broad range of other services to reflect those provided by TTUHSC

The comparison group beneficiaries were selected using the following three eligibility criteria:

- were Medicare fee-for-service beneficiaries with a primary residence in the 48-county area served by the TST program and
- had high costs during the 12-month period between September 1, 2004, and August 31, 2005 (i.e., Medicare costs greater than or equal to $6,000 in 2004) **or** had high disease severity as indicated by Hierarchical Condition Categories risk scores greater than or equal to 1.7 and
- had a majority of visits to four physician group practices similar to TTUHSC as evidenced by claims for health care services obtained between September 1, 2004, and August 31, 2005.

We evaluated three methods for matching comparison beneficiaries to the intervention group: (1) three total Medicare payment tertiles, (2) nine strata cross-classifying three payment tertiles by three HCC risk score tertiles, and (3) five payment quintiles. Of these methods, the third produced a comparison group that was the most similar to the intervention group and was the method selected by TST.

In order to ensure that the comparison group had Medicare costs similar to the intervention group, the comparison group members were randomly selected from each of five cost strata representing the cost quintiles observed in the intervention population. The number of comparison beneficiaries selected from each stratum was determined by the number of intervention beneficiaries in each stratum. The final comparison population of 5,117 beneficiaries had average monthly Medicare costs of $1,733 compared with $1,814 for the intervention group. After capping costs at or above the 99th percentile of costs in each group, the difference in per member per month Medicare spending was reduced to 3.3%.

1.2.4 TST Operations

Texas Senior Trails launched its CMHCB demonstration program April 1, 2006. During the first year of operations, CMS offered TST a refresh population at twelve months post launch that would partially offset attrition due primarily to death. TST declined to accept a refresh population. During the second year of operations, TST requested early termination of its CMHCB program. The Texas Senior Trails program ended July 31, 2007, or 16 months after initiation.

TST received monthly management fees for its full intervention population for the first six months of engagement. After the initial 6-month outreach period, TST accrued management fees 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 TST. 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 TST after a sentinel event, such as a hospitalization or an emergency room visit.

TST negotiated a per beneficiary per month payment of $117 for the duration of the demonstration. At the end of the 16-month period, TST was contractually obligated to achieve a 5% savings in Medicare payments among the intervention group (regardless of participation in the TST program) compared with the comparison group and to cover the TST program fees collected. In addition, TST had the opportunity to share a portion of any savings beyond 5% that were achieved. Payments associated with the Medicare Part D benefit were not included in these calculations.

1.2.5 Overview of the TST CMHCB Demonstration Program

RTI conducted an initial site visit to TST in Lubbock, Texas 7 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 TST and reporting early experiences with program implementation and engagement of eligible beneficiaries, providers, and CMS. Prior to the conduct of a second site visit, TST requested early termination from the CMHCB demonstration program. RTI attempted to conduct a telephone interview with key program staff to learn about changes in the TST program since the last contact with RTI and about the process of closeout of its program and contract with CMS, but TST was not responsive. The description of TST's CMHCB demonstration program and its activities in this report reflects TST's impressions and interpretation of its 16-month experience and does not necessarily reflect RTI's or CMS's perspective on these issues.

The overarching goal of the TST program was to help participants take an active role in their health and receive timely access to appropriate health and social services. TST used a holistic approach to deliver its multidisciplinary care management intervention to help coordinate health and social services for participants with multiple comorbid chronic conditions. Participants also had access to a variety of other services provided by nurse care managers and staff from TTUHSC's schools of medicine, nursing, and pharmacy. Using a combination of telephonic and in-person support services, TST reported that they:

- facilitated patient relationships with physicians and helped patients to comply with physician care plans, including receipt of preventive and routine care,
- provided guidance and support to reduce emergency room utilization, as appropriate,
- supported hospital discharge planning,
- helped patients with adhering to medication regimens,

- educated patients about self-management activities to decrease risk for acute exacerbations of chronic diseases,
- provided resources to address mental health issues, such as depression, and social issues such as transportation needs,
- supplied access to a multidisciplinary clinic, and
- targeted care management support to nursing home residents.

Participants could receive support from any or all of these service areas during the demonstration program, depending on their needs. Next, we describe the process that TST used to assess participant risk for acute events, hospitalizations, and emergency room visits followed by a brief discussion of the key services that were available to participants of different risk levels. Our understanding of the TST program is drawn from our site visit made 7 months after TST launched its CMHCB demonstration program.

Risk assessment—TST enlisted the services of Geriatric Health Systems, LLC (GHS) to conduct risk assessments of the TST participant population to inform the development of individualized care plans and assign participants to one of three risk categories to determine the level of service to be provided to each participant. TST used GHS's Senior Health Profiles® that produced three risk scores: (1) Probability of Repeated Hospital Admission, (2) Frailty Score, and (3) Depression Scale Score.

Levels of care—TST used the information from Senior Health Profiles® to determine the level of care provided to each beneficiary. TST reported that beneficiaries assigned to the low-risk intervention received support, composed primarily of educational information, from nurse care managers during monthly or quarterly telephone interactions. Beneficiaries at risk for repeated hospitalizations or nursing home placement received the medium-risk intervention (i.e., telephonic care management support conducted every 2 weeks to every 2 months). TST reported that the high-risk intervention was provided to approximately 5% to 7% of the TST participant population that had depression and/or potentially critical health problems that required immediate attention. These individuals were reported to have received support during telephone calls and in-person visits with nurse care managers. In addition, beneficiaries who were hospitalized received the high-risk intervention, which included discharge planning support. TST participants who resided in nursing homes also received the high-risk intervention and received on-site visits from care managers to supplement the monthly physician visits typically conducted with this population. In addition, care managers contacted nursing home residents and their families by phone to address health care issues.

Participants transitioned between the different levels of care management throughout the demonstration period based on recommendations of care managers and informed by program guidelines. Typically, participants moved into the high-risk intervention if there were issues that could not be addressed telephonically or if the beneficiary was hospitalized. Once a participant stabilized, he or she would be transferred into the low- or moderate-risk interventions.

Care management services—Each participant was assigned to a care team composed of nurses employed by TrailBlazer, who provided care management support via telephone from Dallas, Texas, and nurse care managers employed by TTPA and located in Lubbock and Amarillo, who conducted on-site visits with beneficiaries at their homes, physician offices, or in the hospital, as needed. TST participants also received support from one of three social workers who addressed issues such as transportation, housing, mental health problems, and preparing advance directives to prepare for end-of-life care needs.

Nurses and social workers drew on community resources such as social services and area agencies on aging to assist participants with social issues. In addition, TST participants also benefited from access to various services provided by TTUHSC's schools of medicine, nursing, and pharmacy. For example, participants residing in Lubbock who were unable to travel to a physician's office could receive home-based care from a program called "Senior House Calls," which is run by the TTUHSC School of Nursing. Participants who had a complex medication regimen of at least eight medications were referred to a School of Pharmacy intervention, which entailed a pharmacist review of the appropriateness of a patient's medication regimen.

One of the primary goals of the TST CMHCB program was to facilitate patient relationships with primary care physicians and specialists so that patients could receive appropriate preventive care to manage their illnesses and become informed consumers of health care. TST care managers educated beneficiaries about physician resources in their areas to help those individuals who had become habituated to receiving care at the emergency room. To further support patient relationships with physicians, care managers helped beneficiaries prepare for office visits by discussing issues that had arisen and helping patients identify questions for the doctor. TTPA care managers who provided onsite support were available to attend physician visits with patients who needed hands-on assistance. TST also encouraged participants, particularly those in nursing homes, to contact their TST care manager prior to seeking care at the emergency room.

Despite the fact that most hospitals have discharge planning coordinators, the services provided to support patients in their transfer home varied widely across facilities. Therefore, TST reported that they provided discharge planning support to its participants. TTUHSC was able to access information about hospitalizations at UMC in Lubbock through its information systems. In Amarillo, TST had an arrangement with one of the two main hospitals to receive census data routinely, and learned about admissions through the TTUHSC family practice residency at the other Amarillo hospital. When TST learned that one of its participants was admitted to the hospital, they reassigned this individual to its high-risk intervention and, when appropriate, a care manager visited the beneficiary in the hospital to determine the cause of the hospitalization and identify any new health or social issues to be addressed.

Sharing of information with physicians—The primary service that TST felt they provided to physicians involved sharing information to facilitate efforts to care for their patients. In particular, care managers informed physicians about care received from other physicians in the community and alerted physicians to health issues that required attention between routine care visits. Care managers typically contacted physicians by either mail or telephone, often interacting with the physician's office staff or nurse, who conveyed messages to the provider. In Lubbock, TTUHSC used an electronic information system to manage appointments, so TST

nurses could find out when participants were coming to the clinic for a visit and could provide the physician with current information about the patients' statuses.

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 CMHCB demonstration period and provider satisfaction with the 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 TST 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.

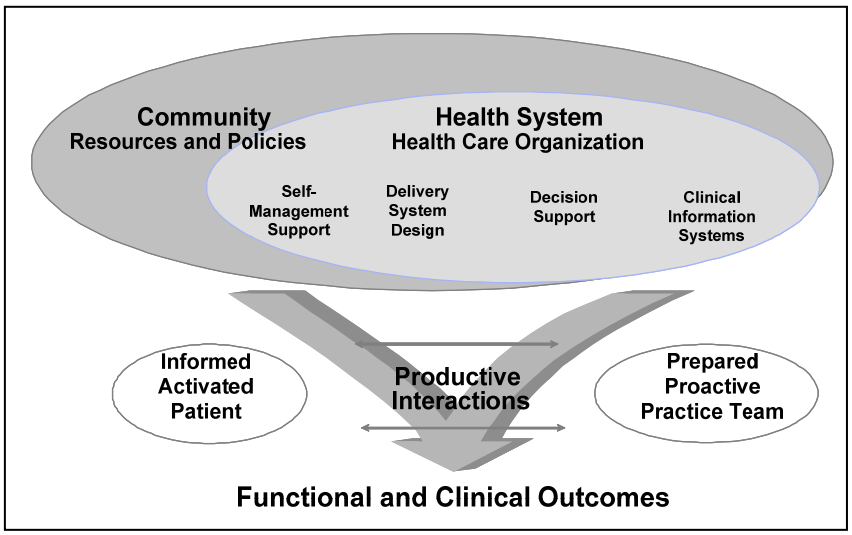
2.1.2 Emerging Approaches to Chronic Care

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 control group, and two had higher costs (Peikes et al., 2009).[2] 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. We briefly describe features of the CMHCB programs below.

- *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.

---

[2] 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.

- *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 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.

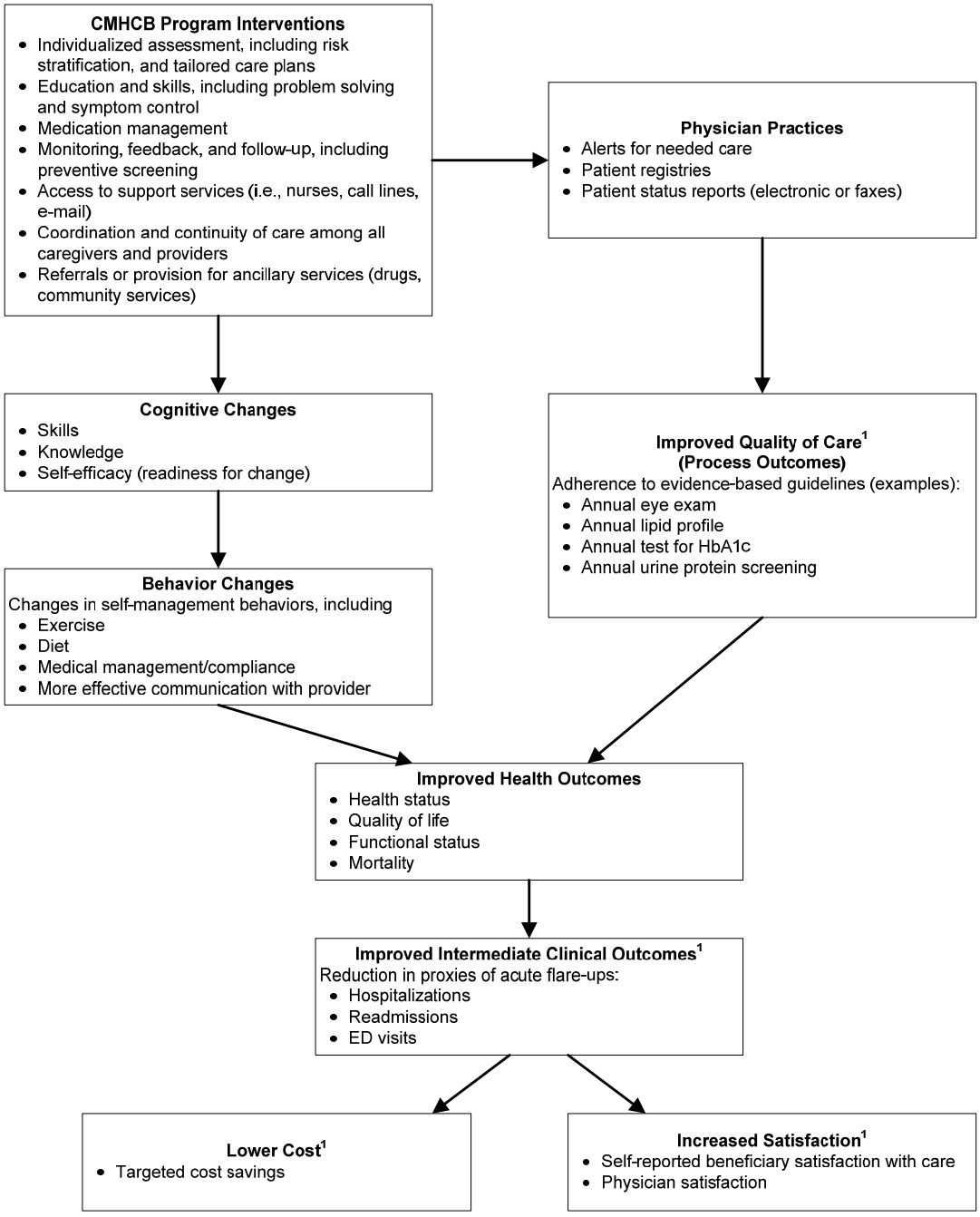
- *Monitoring, feedback, and follow-up*. Activities in this domain include ongoing biomonitoring of beneficiaries by placing scales or other equipment in their homes or by having the beneficiaries self-report their weights, blood sugars, or other measures. When data on preventive services, screenings, 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 Texas Senior Trails program contained elements of the common features. 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

Figure 2-2
Conceptual framework for the CMHCB programs



NOTE: CMHCB = Care Management for High Cost Beneficiaries; CMO = Care Management Organization; ACE = angiotensin-converting enzyme; LVF = left ventricular function; 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 Congressional Budget Office (CBO) (2004).

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 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 Texas Senior Trails (TST) was able to engage its assigned 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 TST implementation experience is reported in Chapter 1.

Table 2-1
Evaluation research questions and data sources

Research questions	Site visits	CMO data	Claims	Survey
IMPLEMENTATION: To what extent was TST 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 TST or partners?	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 TST engage physicians and physician practices in their programs?	Yes	No	No	No
4. How might CMS change the structure of the TST program, if it were to be expanded?	Yes	No	No	No
REACH: How well did TST 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 TST 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 TST program?	No	Yes	Yes	No
4. To what extent were the intended audiences exposed to the TST 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 TST intervention versus a low level of intervention?	No	Yes	Yes	No

(continued)

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

Research questions	Site visits	CMO data	Claims	Survey
EFFECTIVENESS: To what degree was TST 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 TST 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 TST intervention?	Yes	No	No	No
Functioning and health behaviors				
1. Did the program improve knowledge and self-management skills?	No	No	No	Yes
2. Did the TST program result in greater engagement in health behaviors?	No	No	No	Yes
3. Did the TST 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 TST 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 TST program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and ER utilization?	No	No	Yes	No
3. Did the TST 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 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 16-month period compare with the fees that were paid out? How close was TST 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; TST = Texas Senior Trails; CMS = Centers for Medicare & Medicaid Services; CMHCB = Care Management for High Cost Beneficiaries; 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 "prerandomized" 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

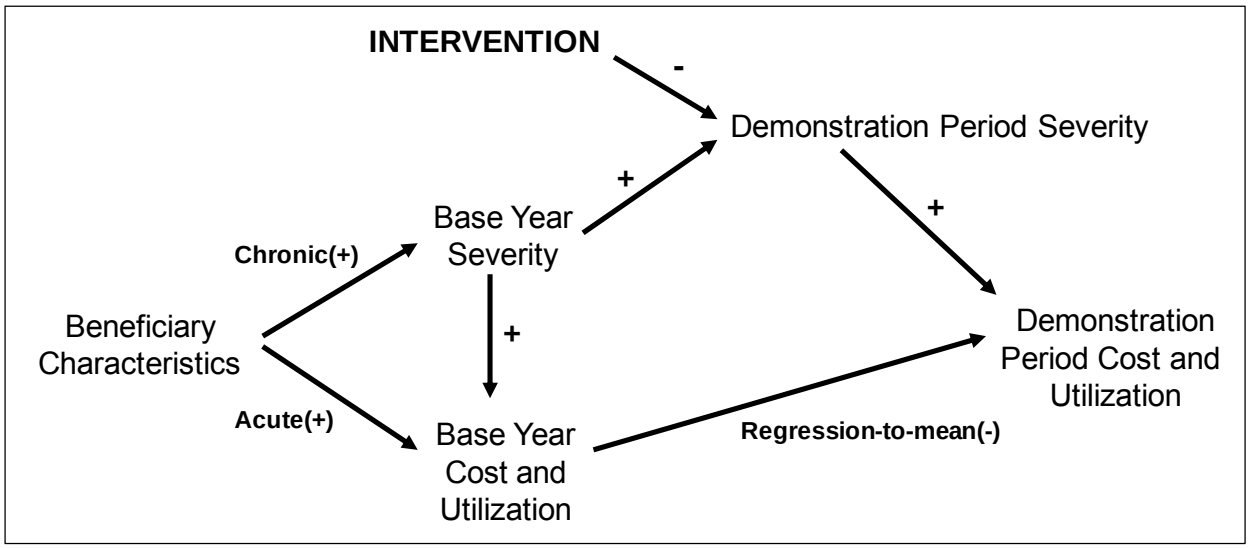
ε = a regression error term.

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 pre-demonstration 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 zero if pre-demonstration assignment 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 that there is 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.

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



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

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 TST CMHCB demonstration program. As noted in Chapter 1, we also conducted a survey of TST 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 seven 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:

- *Eligible intervention beneficiaries data file.* Actuarial Research Corporation (ARC) provided RTI with a data file for TST that contained a listing of beneficiary identification numbers for the selected intervention beneficiaries, their Hierarchical Condition Category (HCC) scores, and their annual Medicare costs. This file was used to do proportional sampling for the comparison group. These beneficiaries were selected from the 48 counties TST was operating in and met the criteria of having annual costs greater than or equal to $6,000 and HCC scores greater than or equal to 1.7. Both the cost and risk score calculations were based on claims incurred for the 12-month period from September 1, 2004 through August 31, 2005.

- *Eligible comparison beneficiaries data file*. ARC provided RTI with a data file for TST that contained a listing of beneficiary identification numbers for those beneficiaries in the 48 counties that were not selected for the intervention group, their Hierarchical Condition Category (HCC) scores, and their annual Medicare costs. RTI subsetted this file to those beneficiaries that met the same cost and HCC criteria as the intervention group. The final comparison sample was selected to match the cost quintiles of the intervention group.

- *Participant status files*. We received participant status files from ARC. The participant status information originates from TST and was submitted to ARC. This file was updated quarterly and logged status changes among the intervention groups by TST. 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 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 TST beneficiary was assigned—intervention or comparison.

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

- ARC provided RTI with an EDB file for TST comprised of all intervention and comparison beneficiaries. RTI used these files to determine daily eligibility based on TST 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 TST'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 assignment (March 2, 2006) for TST.

Table 2-2
Criteria used for determining daily eligibility for TST

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: TST = Texas Senior Trails; ESRD = end-stage renal disease; MA = Medicare Advantage; GHO = Group Health Organization.

- *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 intervention and comparison beneficiaries during the 12 months prior to TST'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, TST 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 intervention period for the participation analysis.

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

Table 2-3
Analysis periods used in the TST 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
4/1/06	7/31/07	16	4/1/05	3/31/06

NOTES: CMHCB = Care Management for High Cost Beneficiaries; TST = Texas Senior Trails.

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 comparison group selection, March 2, 2006. Medicaid enrollment was determined at any time during the baseline period and was also determined using the EDB.

Daily eligibility variables were used to create analytic variables representing the fraction of the baseline and demonstration period that the intervention and comparison beneficiaries were CMHCB program eligible. These eligibility fractions were created based on the time period of the analysis. For example, 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 TST was active in the demonstration. The numerator is the number of days the beneficiary is eligible during that time period. TST participated in the demonstration for 16 months, so the number of days in the denominator for each beneficiary in TST is 487 (TST end date minus TST 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 487, or 0.862.

2) Institutionalized Status. Four binary indicators of institutionalization were created:

- 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 SNF costs.

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

3) 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.

- A *concurrent HCC score* calculated by RTI for the first 6 months of the intervention period. 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 assignment.

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

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

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

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

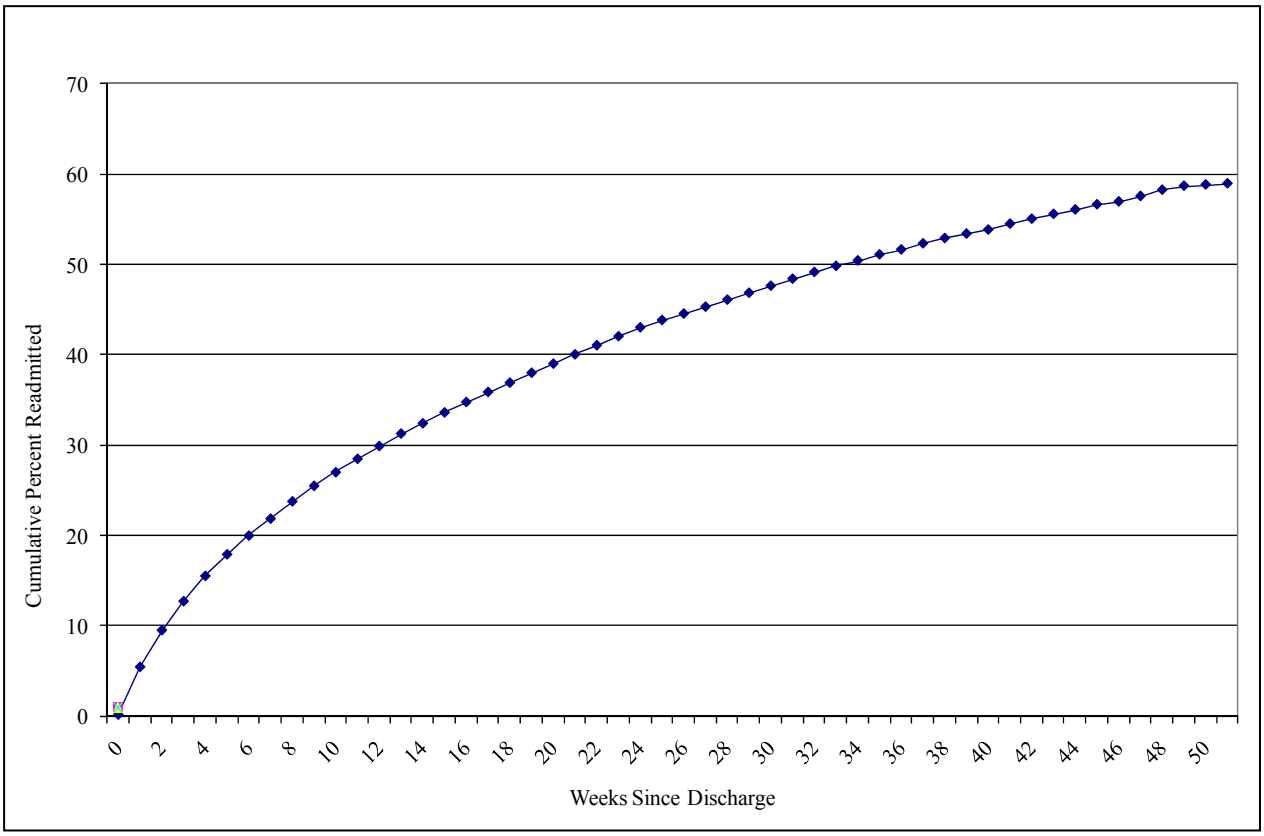
Only claims that occurred during periods of eligibility were included in the utilization measures. For both the demonstration and baseline periods, claims were included if services were started during days that the beneficiary met TST's CMHCB program eligibility criteria, as determined from the ARC daily 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 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[3].

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



We examined readmissions following admissions that occurred during a 12-month period. 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

3 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.

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 period examined admissions during the period from 13 months through 2 months prior to the end of the demonstration and included readmissions through the end of the demonstration period. 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 ten ambulatory care sensitive conditions.

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 16-month 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 TST. 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 behavior of intervention and comparison populations that was not subject to Medicare eligibility or payment rules and to provide credit to TST in case the services occurred after exposure to the CMHCB intervention and during the intervention period. One could envision that TST encouraged the receipt of the process-of-care measures; however, the actual service was provided during a brief period of ineligibility, for example, 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 TST, 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 a count of the types of services received during the demonstration.

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[4].

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 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. The findings from all six CMHCB beneficiary surveys were reported to the Centers for Medicare & Medicaid Services (CMS) in RTI's Third Annual Report (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 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 CMHCB interventions: 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

---

[4] 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 TST 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 TST intervention on the beneficiaries' assessment of satisfaction that their full health care team is helping them to cope with their chronic conditions.

communication with health care providers. Last, the survey instrument included four physical and mental health functioning measures.

3.1.1.1 Measures of Experience and Satisfaction with Care

The impact of the 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 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 prior 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 assessing how quickly the health care team gets back to beneficiaries with answers to their medical questions. The questions ask how often beneficiaries get answers the same day during office hours or if they call after regular office hours, how often their questions were answered.

Six frequency options (always, almost always, usually, sometimes, almost never, and never) are converted into composite scores ranging from 0 (never to all items) to a maximum of 100 (always to all items).

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

3.1.1.2 Self-Management Measures

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

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

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

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

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

2000). Respondents indicate the number of days (0-7) in the past week that they performed each self-care activity.

3.1.1.3 Physical and Mental Health Function

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

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

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

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

$$Y = a + b_1X_1 + b_kX_k + e, \qquad (3\text{-}1)$$

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 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 data base (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. TST'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 TST was 74%. 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 TST

This section presents the results of the Medicare Health Services Survey data analysis for TST. 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. We wrap up this section with a summary of results for TST and our conclusions.

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 TST.

Table 3-1
Medicare Health Services Survey: Estimated intervention effects for experience and satisfaction with care,
TST
(N = 618)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Helping to cope with a chronic condition (1 to 5)	3.51	3.71	-0.07	N/S
Number of helpful discussion topics (0 to 5)	1.85	1.68	0.14	N/S
Discussing treatment choices (1 to 4)	3.00	3.16	-0.07	N/S
Communicating with providers (0 to 100)	70.9	75.7	-4.0	N/S
Getting answers to questions quickly (0 to 100)	57.0	63.3	-5.4	N/S
Multimorbidity Hassles score (0 to 24)	3.53	3.20	0.02	N/S

NOTES: TST = Texas Senior Trails; ANCOVA = Analysis of Covariance.

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

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

Overall experience: 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 3.5 for the intervention group, or about midway between "very good" and "good" ratings. The average score for the comparison group was about 3.7. More than 57% of TST beneficiaries rated their experience as "excellent" or "very good," and about 26% selected "good." It is not uncommon among the elderly to report high satisfaction

ratings. For that reason, the mean scale score was used in the analyses to capture transitions between all response categories.

For this overall satisfaction measure, we observe no statistically significant intervention effect for TST. TST's intervention was not effective in improving beneficiary overall satisfaction in helping them cope with their chronic illness. Among other covariates in the ANCOVA model, gender, education, and survey mode of administration were significant predictors of overall satisfaction. Older and better educated TST beneficiaries were more likely to rate this outcome higher. Mail survey respondents gave lower ratings how well the TST program helped them to cope with their chronic condition.

Number of helpful discussion topics—The mean number of services for which beneficiaries had helpful discussions with their health care team was comparable between the intervention and the comparison groups. For this measure, we observe no statistically significant intervention effect for TST. Female TST beneficiaries were more likely to have had more helpful discussions than males.

Discussing treatment choices—For this item, beneficiaries were asked whether their health care teams talked about pros and cons of each treatment choice with answers ranging from 1 = "definitely no" to 4 = "definitely yes." The mean score for the intervention group was 3.0, compared with 3.2 for the control group. For TST, we observe no statistically significant intervention effect on this satisfaction item. Gender was a significant predictor for this satisfaction with care outcome: Female TST beneficiaries were more likely to rate this outcome higher than males.

Communication with health care team—The score for communication with health care team could range from 0 to 100, with 0 indicating never for all items in the composite and 100 indicating always for all items in the composite. TST showed somewhat lower average communication scores, with 71 for the intervention group and about 76 for the comparison group. For TST, we observe no statistically significant intervention effect on this satisfaction item, indicating that TST intervention did not affect communication with providers. Older TST beneficiaries and beneficiaries with health insurance coverage in addition to Medicare were more likely to rate this outcome higher.

Getting answers to questions quickly—The score for getting answers to questions quickly could range from 0 to 100, with 0 indicating never for both items in the composite and 100 indicating always for both items in the composite. TST showed moderate average scores for intervention and control groups. For this measure, we observe no statistically significant intervention effect for TST. Two other covariates in the model were significant predictors for this outcome. Older TST beneficiaries and females were more likely to report that they got the answers to their questions quickly than their counterparts.

Multimorbidity Hassles scale—Multimorbidity Hassles Scale, designed to measure frustrating problems that patients experience in getting comprehensive care for chronic illnesses, is measured on a scale from 0 to 24. High score indicates more problems. TST showed relatively low Multimorbidity Hassles scores for intervention and comparison groups (3.5 and 3.2, respectively). In the original development sample, the mean Hassles score for these six items was

5.86 (Parchman, Noel, and Lee, 2005). For this measure, we observe no statistically significant intervention effect for TST. Two other covariates in the model were significant predictors for the Multimorbidity Hassles score. Older TST beneficiaries experienced fewer frustrating problems in getting comprehensive care for their chronic illnesses than their counterparts. Mail survey responders reported more problems.

In summary, across the six measures of experience and satisfaction with care, we observe no significant intervention effects for TST. TST was unable to influence any aspect of its members' experience and satisfaction with care with their CMHCB intervention.

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 TST.

Table 3-2
Medicare Health Services Survey: Estimated intervention effects for self-management,
TST
(N = 618)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
Percent receiving help setting goals	45.8	43.7	1.5	N/S
Percent receiving help making a care plan	46.4	38.9	8.2	N/S
Self-efficacy ratings				
Take all medications (1 to 5)	4.12	4.46	-0.23	*
Plan meals and snacks (1 to 5)	3.83	4.12	-0.21	N/S
Exercise 2 or 3 times weekly (1 to 5)	3.36	3.67	-0.25	N/S
Self-care activities				
Prescribed medications taken (mean # of days)	6.53	6.75	-0.14	N/S
Followed healthy eating plan (mean # of days)	4.75	4.83	-0.07	N/S
30 minutes of continuous physical activity (mean # of days)	2.93	3.17	-0.28	N/S

NOTES: TST = Texas Senior Trails; ANCOVA = Analysis of Covariance.

Statistical significance (Stat sign): * Indicates significance at the 5% level; ** Indicates significance at the 1% level; otherwise no statistically significant.

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 a beneficiary's health care team helped set goals or make a plan to take care of the beneficiary's health problems. For TST beneficiaries, in the intervention group 46% received help setting goals and 46% received help making a care plan. In the comparison group, 44% and 39%, respectively, received assistance with these self-management activities. The ANCOVA results reveal that TST was not effective at helping beneficiaries set goals for self-care management, nor was it effective at providing help in making health care plans. Both coefficients, though positive in direction, are not statistically significant. There were a few other covariates that predicted receiving help on these two measures: females, proxy respondents, and those with higher baseline HCC scores were significantly more likely to get help setting goals, and females and proxy respondents were also more likely to receive help with making a care plan.

Self-efficacy ratings—To assess self-efficacy, respondents were asked how sure they were that they could perform each of three specific behaviors. Overall, TST beneficiaries typically reported relatively high levels of self-efficacy with mean ratings averaging around 4 (somewhat sure of their ability to perform self-care activities) out of a maximum of 5 (very sure). The highest self-efficacy scores were reported for taking medications as prescribed, and the lowest scores were for getting exercise 2 or 3 times per week. On average, TST beneficiaries in the intervention group were significantly less sure that they could take their medications as often as prescribed (4.1), compared with 4.5 in the comparison group. The intervention and comparison groups rated confidence in planning meals and snacks 3.8 and 4.1, respectively, and confidence in exercising as 3.4 and 3.7, respectively. Thus, of the three self-efficacy items, the only significant intervention effect was an unfavorable impact on taking medications as prescribed. Older beneficiaries were more likely to be confident about taking their medications, proxy respondents and those with higher baseline HCC score were significantly more confident about planning their meals and snacks, and Hispanic TST beneficiaries were more confident about getting needed exercise. Females and proxy respondents were less confident about their exercise regimen.

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 was quite high for taking medications and more modest for exercise. There were no significant group differences in the rates for any of these three self-care activities. In terms of other significant covariates for self-care outcomes, greater age for TST beneficiaries decreased their compliance with prescription medications but increased their compliance with diet and exercise. The medication and exercise compliance was lower for proxy respondents compared with self-respondents. Diet compliance improved with age and with Hispanic status but decreased for those who lived alone and for mail respondents. Compliance with exercise was also better for Hispanic and African American beneficiaries compared with Whites and for those with higher PBPM expenditures. Self-care exercise behavior was less consistent for females, those with higher baseline HCC score and for mail respondents.

3.1.3.3 Physical and Mental Health Functioning

Physical and mental function—***Table 3-3*** displays the mental and physical functioning outcomes for TST. On average, TST respondents had the mean PHC score for the intervention

group at 28.3, slightly lower when compared with 29.5 for the comparison group but not significantly so, according to the ANCOVA adjustment. The mean MHC score for the intervention group was 35.8 and the PHQ-2 score of 2.6, compared with 38.7 and 2.0 for the comparison group. The ANCOVA estimation revealed that for Texas Tech there are no statistically significant intervention effects for physical and mental functioning outcomes.

For TST members, self-reported physical function, as shown by PHC scores, increased significantly with age and years of education but was lower for females, those responding to the survey by proxy, and mail respondents. Similarly, mental function, as expressed by MHC scores, was significantly higher with increased age and every additional year of education but was lower for females than males and for proxies compared with self-respondents. PHQ scores were lower for those with more years of education and higher for proxy respondents.

Table 3-3
Medicare Health Services Survey: Estimated intervention effects for physical and mental health function,
TST
(N = 618)

Outcome	Intervention mean	Comparison group	ANCOVA-adjusted intervention effect	Stat. sig.
PHC score (physical health, mean =50, std = 10)	28.3	29.5	-0.4	N/S
MHC score (mental health, mean = 50, std = 10)	35.8	38.7	-1.3	N/S
PHQ-2 score (depression, 0 to 6)	2.64	2.03	0.27	N/S
Number of ADLs difficult to do (0 to 6)	2.75	2.32	0.04	N/S
Number ADLs receiving help (0 to 6)	1.30	0.82	0.27	N/S

NOTES: TST = Texas Senior Trails; ANCOVA = Analysis of Covariance; PHC = Physical Health Composite; MHC = Mental Health Composite; PHQ-2 = Patient Health Questionnaire 2; ADLs = activities of daily living.

Statistical significance (Stat sig.): * Indicates significance at the 5% level; ** Indicates significance at the 1% level; otherwise no statistically significant.

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, respondents in the TST intervention group reported being limited on 2.8 ADLs compared with 2.3 ADLs for the comparison group. TST beneficiaries also reported that they received help with an average of 1.3 to 0.8 ADLs. ANCOVA results indicate that there was no difference in functional status as a result of the intervention. For TST members, when other characteristics are held constant, females reported significantly more ADL limitations than males and proxy respondents reported more than self-respondents. As expected, those with higher baseline HCC scores also report significantly higher levels of functional impairment. Those who had higher baseline HCC scores, had higher PBPM expenditures, or responded by mail also reported significantly higher levels of ADL impairment. TST members who have additional health insurance coverage report fewer ADL limitations than those who

only have Medicare. In terms of needing help with ADLs, the patterns are similar: females, proxy respondents, and members with higher baseline HCC score report needing help on a significantly higher number of ADLs. Those who lived alone and those with additional health coverage 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. The CMHCB demonstration programs hypothesized that lifestyle changes and better communication with providers would mitigate acute flare-ups in 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, only one statistically significant group difference was found—members of TST's intervention group were less certain that they could take their medications as prescribed. This difference, however, was not reflected in actual weekly medication usage rates, which were only slightly lower for the intervention group.

3.2 Physician Satisfaction

RTI made one site visit to meet with the TST program staff during the demonstration period. The site visit was conducted in November 2006, 8 months after initiation of the TST demonstration program. A small number of physicians participated in a focus group to discuss their knowledge of and satisfaction with various aspects of the TST CMHCB demonstration program. In addition, an interview was held with the TST medical director.

RTI had planned to conduct a more comprehensive evaluation of physician satisfaction with the TST demonstration program during its second site visit that would be held approximately 18 months after program launch. Because of TST's request for early termination, a second site visit was not conducted. Therefore, physician assessment of the value of the TST demonstration program is based on a very small number of physicians whose responses were reported fairly early in the demonstration program. The three interviewed physicians had limited firsthand experience of interacting with the care managers. Their responses were primarily related to the potential value they saw of the program.

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

3.2.1 TST Outreach to Physicians

A major goal of the TST program was to use its care managers to facilitate physician-patient relationships so that patients could receive appropriate preventive care to manage their illnesses and become better informed consumers of health care. Thus, the TST model required active participation by all physicians providing care to their intervention beneficiaries.

TST used a multipronged approach to inform physicians about the TST program. All physicians who had provided care to beneficiaries eligible for the TST intervention (based on claims data) received an introductory letter from CMS explaining the program and emphasizing the fact that the program is not intended to interfere with current physician-patient relationships. TST followed up by sending marketing materials to providers. A provider outreach nurse conducted face-to-face meetings with all providers who cared for at least 100 members of the intervention group. In addition, TST staff conducted in-service presentations on various occasions (e.g., faculty meetings) to explain the elements of the program and how physicians could determine if one of their patients was eligible or participating in the program. In particular, TST focused on marketing the program to the University Medical Center (UMC) Physician Network Services (PNS) group because PNS is composed largely of primary care physicians responsible for making referrals to specialist care.

In addition to distributing marketing materials and conducting group presentations, a TST nurse with utilization management and provider relations experience visited the offices of the 250 doctors in the Lubbock and Amarillo areas with the highest numbers of CMHCB demonstration-eligible patients. This nurse was largely successful in gaining physician support for the program, often as a result of spending time with physician office staff and administrators who conveyed information about the program to the physicians. She found that physicians in the Amarillo area and the southern rural section of the target regions were the most challenging to engage in the TST program. Similarly, the TST medical director in Amarillo had so much difficulty obtaining physician support via phone calls to these individuals that he ceased conducting these outreach calls. At the time of our site visit, the TST medical director and managing director were continuing to look for ways to market the program to providers who were not supportive initially. In particular, they were developing messages that conveyed the fact that the program can serve as a resource for physicians, by providing support for patients who are hard to manage because of mental health and/or social issues.

3.2.2 Sharing of Information/Ongoing Relationship with Providers

The primary service that TST provided to physicians involved sharing information to facilitate efforts to care for their patients. In particular, care managers informed physicians about care received from other physicians in the community and alerted physicians to health issues that would require attention between routine care visits. Care managers typically contacted physicians by either mail or telephone, often interacting with physicians' office staff or nurses, who conveyed messages to the providers. In Lubbock, Texas Tech University Health Sciences Center (TTUHSC) used an electronic information system to manage appointments so TST nurses could find out when participants were coming to the clinic for a visit and provide the physicians with current information about patients' status.

At the time of the site visit, TST was still considering various options for presenting physicians with feedback about the quality of care provided to patients. For example, regional benchmarks were available for many of the quality measures that the program was tracking because they are part of the National Committee for Quality Assurance (NCQA) Health Employer Data and Information Set (HEDIS) and are available within Agency for Healthcare Research and Quality (AHRQ) datasets already used at the university. However, members of the TST leadership observed that physicians were often most receptive to information about specific

patients. Therefore, TST was interested in finding a way to present trend reports and benchmark reports to physicians in a way that could help them improve performance without feeling criticized. One of the physicians interviewed during the site visit noted the importance of receiving timely feedback rather than reports of obsolete data.

In addition to providing physicians with information, TST anticipated that physicians would benefit by being less burdened by difficult patients, who would receive care management support from TST nurses. TTUHSC physicians also could acquire new patients (and the associated income) as a result of TST's efforts to link participants with primary care providers, if they did not have one.

3.2.3 Physician Assessment of the TST Demonstration Program Value

RTI interviewed three TTUHSC physicians who specialized in internal medicine, family practice, and oncology, who were generally pleased with their limited experiences with the TST program. Because the CMHCB demonstration population received care from a variety of providers, the physicians interviewed felt that the coordination provided by TST care managers was valuable and had the potential to reduce health care costs by reducing emergency room visits and inappropriate use of subspecialty care. In particular, the interviewed physicians felt that the care managers were well suited to educate patients and their family members and caregivers about appropriate use of the emergency department.

The physicians also noted that 24-hour access to care managers, home visits, and home assessments were potentially valuable for providing care management support to the CMHCB demonstration population, and ideally, TST participants would learn to call their care managers whenever faced with health issues. However, one physician was concerned that, when the program ended, patients would revert to previous behaviors such as seeking nonurgent care at the emergency room.

The TST program staff noted that noninterviewed physicians had been supportive of the TST program, which was evidenced by the fact that some physicians were calling TST directly if they learned that a patient needed assistance. In addition, physicians had requested to enroll additional patients in the program.

TST staff reported that most participants had a primary provider that was associated with TTUHSC; however, many patients, particularly in Amarillo, received care from additional providers that were not associated with the university. These providers typically operated independent practices, so TST care managers had to establish relationships with a number of different practices. Generally, care managers interacted with physician office staff members and nurses, rather than directly with the physicians. None of these providers were made available for interviews, so we are unable to provide their assessment of the value of the TST demonstration program.

CHAPTER 4
PARTICIPATION RATES IN THE TST 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 TST 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 TST program?
- To what extent were the intended audiences exposed to the TST programmatic interventions? To what extent did participants engage in the various features of the program?
- What beneficiary characteristics predict a high level of TST 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 assigned to the intervention group and compared with all eligible beneficiaries selected for 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 November 2006 site visit, Texas Senior Trails (TST) reported that they reached its outreach goal by engaging 88% of the CMHCB intervention population as of October 20, 2006 (Brody and Bernard, 2007). 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 assigned beneficiaries. The TST intervention had a variety of telephonic and in-person elements (e.g., it facilitated patient relationships with physicians and helped patients 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 TST and their participants. For each participating beneficiary, TST 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. TST 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.[5] These latter two sets of numbers provide an estimate of the number of beneficiaries with whom the CMOs 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 one 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 X_i vector), and e = the base of natural logarithms. The probability of a beneficiary being in the participant group is thus explained by the X_i variables.

[5] 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 assignment:

- 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 between 2 and 3 (medium) and 4 or greater than (high); Charlson score less than 2 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 each CMHCB demonstration program was greater than or equal to $528

and less than $1,612 (medium) and $1,612 or greater (high); PBPM cost less than $528 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 one 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 intervention period was greater than 0.315 but equal to or less than 1.125 (medium) and greater than 1.125 (high); concurrent HCC score less than or equal to 0.313 is the reference group.

4.2.2 Level of Intervention Analysis Methods

On a quarterly basis, TST 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 TST participants. Contacts for care management services were made between participants and either care managers or physicians. Contacts between participants and physicians were exclusively by telephone; while all three modes of contact were used by care managers. In addition, TST also provided information on the nature of the contact. TST identified 15 different care coordination or management services such as care plan development, discharge planning, care management adherence, care coordination for home health, and other services. However, only about 50% of care manager contacts had associated content or nature of the contact, and about 80% of physician contacts had content. Further, there were records provided to us that contained content but no mode of contact. Therefore, we report these two elements of the TST intervention data separately.

Using the encounter data submitted by TST, we constructed counts of the number of contacts with participants—either telephonically, in person, or via written communication—as well as total contacts (all three) and total contacts by the care managers versus physicians. We report the distribution of beneficiaries across five categories of contacts (0, 1, 2-4, 5-9, and 10 or more) and the mean and median number of total contacts. We also estimated a multivariate logistic regression model of the likelihood of being in the high total contact category relative to the low total contact category. Three contact categories were constructed using the tertiles (low, medium, and high) of the distribution of total contacts. A dichotomous dependent variable was created and set at 1 for beneficiaries who had a high level of contact with TST and 0 for beneficiaries who had a low level of contact. Beneficiaries who had a medium level of contact with TST 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 TST 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 starting with month 4 and ending with month 16. We excluded the first 3 months of the demonstration period from this analysis to reduce confounding intervention contacts with engagement contacts. While the first 6 months of the demonstration was primarily devoted to engagement, we observed a sizeable increase in care management services being provided in month 4 over months 1 through 3. The only contacts in month 16 were written contacts, which were most likely letters to beneficiaries informing them of TST's decision to terminate participation in the CMHCB demonstration early.

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 CMHCB demonstration program for each month from month 4 to month 15. This is the subset of beneficiaries with whom TST 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 3,607, or 73% of eligible intervention beneficiaries in the demonstration.

4.3 Findings

4.3.1 Participation Rates for the TST 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. The results are based on the full demonstration period for the original population only because TST did not have a refresh population. The number of months for the full demonstration period for TST is 16 months.

Table 4-1 displays the number of beneficiaries included in our participation analyses 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 TST had a

total of 16 months (or 487 days) of possible enrollment. If they died after 90 days, their FTE value would be 90/487 or .18 FTEs. If someone was eligible for all 16 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.

Table 4-1
Number of Medicare FFS beneficiaries eligible for and participating in the TST CMHCB demonstration

Characteristics	Months 1-16	Months 1-12	Months 5-16
Intervention group			
Number eligible[1]	4,966	4,965	4,717
Full time equivalent[2]	4,472	4,600	4,349
Number fully eligible	3,959	4,203	3,963
Participants			
Number eligible	4,499	4,490	4,318
Full time equivalent	4,140	4,247	4,003
Number fully eligible	3,678	3,898	3,648
Participants > 75%			
Number eligible	3,881	3,668	4,233
Full time equivalent	3,584	3,467	3,928
Number fully eligible	3,196	3,190	3,585
Non-participants			
Number eligible	467	475	399
Full time equivalent	331	353	346
Number fully eligible	281	305	315
Comparison group			
Number eligible	5,080	5,079	4,932
Full time equivalent	4,726	4,825	4,634
Number fully eligible	4,327	4,502	4,331

NOTES: FFS = fee-for-service; TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries.

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

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

Data Sources: Medicare claims data, Medicare enrollment database.

Program: /vol1/project/07964/025 hiccup/pgm/larsen/programs/texas/final/9mo/tables/tableTT-1.sas 30JUNE2009.

The ratio of FTEs to the total number of eligible beneficiaries in the original intervention population is 0.90 for the entire intervention period (months 1-16) compared with a slightly higher ratio for the first 12 months (.93) and 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 first 12 months because we have a lower proportion of the population that is dying and for the last 12 months because we have dropped from the eligible population anyone that died or lost eligibility in the first 4 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 percent of the original intervention beneficiaries and 15% 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. TST's nonparticipant group was eligible only 71% of the days—much lower than the 92% of days for participants. Also, the participant group had a much higher rate of beneficiaries being fully eligible for the entire intervention period (82%) compared with 60% for the nonparticipant group.

Table 4-2 presents participation rates for the TST population and displays 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 TST population—Of all TST intervention beneficiaries, 91% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. We previously reported (McCall et al., 2008) that most participants consented in the initial 6-month engagement period (87%), however, in comparison with participation rates previously reported for months 1 through 6 of the demonstration, we see a slight increase over the entire intervention period. Eighty-three percent of beneficiaries were continuous participants (***Table 4-2***), which equates to over 90% of participants. Of TST beneficiaries, 4% refused to participate. The percent not contacted or unable to be located was 5%.

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 TST's intervention beneficiaries, 83% participated for more than 75% of their eligible months, which mirrors the continuous participant percentage. ***Table 4-2*** also reports the number of participants over time (for months 6 and month 16, the last month of CMHCB program operation). The number of participants declined over time as would be expected given the attrition due to loss of eligibility or death.

Table 4-2
Participation in the TST CMHCB demonstration program

Characteristics	Statistic
Number of intervention months	16
Participation rate (entire intervention period)	91%
Length of Participation	
Continuous participation after engagement	83%
After initial participation, became a continuous nonparticipant	6%
Intermittent participation	1%
Nonparticipation (never agreed)	9%
Refused to participate when contacted	4%
Not contacted/unable to be contacted	5%
Beneficiaries participating more than 75% of days	83%
Number of participants in selected months[1]	
Month 6	4,114
Month 16 (Last Month)	3,790

NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries.

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

Data Sources: Medicare claims data, Medicare enrollment database.

Program: /vol1/project/07964/025 hiccup/pgm/larsen/programs/texas/final/9mo/tableTT-2.sas 07JULY2009

4.3.2 Characteristics of the TST Intervention and Comparison Populations

In addition to evaluating the level of initial engagement by TST, 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. 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 assignment 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 assignment procedures for producing similar populations based on the selection stratum 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. ***Table 4-3*** provides the percent of beneficiaries by these characteristics for both the intervention and comparison populations.

Table 4-3
Characteristics of the TST CMHCB demonstration program intervention and comparison populations

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Total number of beneficiaries	4,966	5,080	—	—
Full Time Equivalent	4,472	4,726	—	—
Beneficiary characteristics				
Aged-in (vs. disabled)	81.5	91.9	-10.4	**
In Medicaid (vs. not in Medicaid)	3.7	1.5	2.2	**
Male (vs. female)	38.9	40.4	-1.5	N/S
Urban (vs. rural)	75.0	60.8	14.2	**
Age				
Mean	72.6	75.2	-2.6	**
<65	18.6	8.1	10.5	**
65-69	15.0	14.3	0.7	N/S
70-74	18.9	21.8	-3.0	**
75-79	17.3	23.9	-6.6	**
80-84	15.2	17.8	-2.6	**
85+	15.1	14.1	1.0	N/S
Race				
White	85.5	92.5	-7.0	**
African American	6.6	3.2	3.4	**
Other	7.8	4.2	3.5	**
Unknown	0.1	0.1	0.0	N/S
Health status				
Recalculated HCC score				N/S
Mean	1.8	1.8	0.0	N/S
Low: > 1.35 and < 2.00	64.5	65.6	-1.1	N/S
Medium: > 2.00 and < 3.10	22.3	20.9	1.4	N/S
High: > 3.10	13.3	13.5	-0.3	N/S
Baseline costs >= $6000	67.3	65.8	1.6	N/S
Charlson comorbidity index - Mean	2.5	2.6	-0.2	**
Chronic conditions				
HF	18.6	18.3	0.32	N/S
Coronary artery disease	32.8	37.9	-5.16	**
Other respiratory disease	22.4	23.3	-0.89	N/S
Diabetes without complications	29.0	25.4	3.63	**
Diabetes with complications	6.0	7.3	-1.35	*
Essential hypertension	57.3	52.0	5.24	**
Valve disorders	4.6	6.5	-1.87	**
Cardiomyopathy	2.3	3.8	-1.47	**
Acute and chronic renal disease	10.5	11.4	-0.87	N/S
Renal failure	2.8	4.1	-1.23	**

(continued)

Table 4-3 (continued)
Characteristics of the TST CMHCB demonstration program intervention and comparison populations

Characteristics	Rate per 100[1,2] I	Rate per 100[1,2] C	I vs. C	p[3]
Chronic conditions (continued)				
Peripheral vascular disease	5.1	4.1	1.01	*
Lipid metabolism disorders	21.2	28.4	-7.12	**
Cardiac dysrhythmias and conduction disorders	17.8	21.8	-4.04	**
Dementias	2.0	0.6	1.39	**
Strokes	6.3	5.1	1.16	*
Chest pain	10.6	9.3	1.33	*
Urinary tract infection	15.0	13.1	1.93	*
Anemia	13.2	14.9	-1.7	*
Malaise and fatigue (including CFS)	6.5	7.2	-0.6	N/S
Dizziness, syncope, convulsions	11.3	9.6	1.73	**
Disorders of joint	14.7	12.6	2.07	**
Hypothyroidism	14.4	14.5	-0.09	N/S

NOTES: TST = Texas Senior Trails; 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: H:/project/07964/025 hiccup/pgm/larsen/programs/texas/final/9mo/tableTT-3.sas 30JUNE2009

Beneficiaries for both the intervention and comparison groups were eligible for the demonstration if they had an HCC risk score greater than or equal to 1.7 or Medicare costs greater than or equal to $6,000 in 2004. In order to ensure that the comparison group had Medicare costs similar to the intervention group, the comparison group members were randomly selected from each of five cost strata representing the cost quintiles observed in the intervention population and in proportion to the intervention beneficiaries. We observe both cost and HCC score equivalency between the two groups. The mean HCC score for both the intervention and comparison groups was 1.8, meaning that beneficiaries selected for the demonstration were, on average, predicted to be 80% more expensive than the average Medicare fee-for-service (FFS) beneficiary. Two-thirds of beneficiaries had HCC scores less than 2.0, while only 13% had HCC scores above 3.1.

Based on beneficiary characteristics, there are many statistically significant 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, enrolled in Medicaid, under the age of 65, or a minority. These characteristics are often proxies for poor health status. At the same time, the intervention group had a lower mean Charlson comorbidity index score (an alternative health status measure that uses concurrent disease presence to quantify disease burden), although the actual rates are very similar. We also observe considerable variation in baseline rates of chronic conditions. The intervention group had higher rates of conditions such as diabetes with complications, hypertension, and peripheral vascular disease and lower rates of coronary artery disease, cardiac dysrhythmias, and lipid metabolism disorders. There is no apparent pattern in the differences in chronic conditions between the intervention and comparison groups.

4.3.3 Characteristics of Participants in the TST Population

In this section, we report the beneficiary characteristics that predict participation in the TST CMHCB demonstration program. First, we report the same characteristics from ***Table 4-3*** 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 ***Table 4-4***). There are statistically significant lower percentages of men that were participants compared with nonparticipants and higher percentages of beneficiaries with diabetes with complications and renal disease in the participating group.

Next, we report participation rates by beneficiary demographic characteristics, baseline clinical and financial characteristics, and intervention period health status that we use in the multivariate modeling of participation (***Table 4-5***). 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. Despite these indications, participation rates were high across all characteristics because most eligible beneficiaries participated in the demonstration.

In order to better understand which characteristics predicted participation in the demonstration, we estimated three logistic regression models:

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

Table 4-4
Characteristics of the TST CMHCB demonstration program intervention population by participation status

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	4,499	3,881	467	—	—
Full time equivalent	4,140	3,584	331	—	—
Beneficiary characteristics					
Aged-in (vs. disabled)	81.4	81.7	82.1	-0.7	N/S
In Medicaid (vs. not in Medicaid)	3.7	3.3	3.6	0.1	N/S
Male (vs. female)	38.4	38.8	45.0	-6.6	*
Urban (vs. rural)	75.0	75.1	75.0	0.1	N/S
Age					
Mean	72.6	72.5	72.4	0.2	N/S
<65	18.6	18.3	17.9	0.7	N/S
65-69	15.0	15.5	15.3	-0.3	N/S
70-74	18.8	19.4	19.6	-0.8	N/S
75-79	17.5	17.8	15.2	2.3	N/S
80-84	15.1	15.1	16.8	-1.7	N/S
85+	15.1	13.9	15.3	-0.2	N/S
Race					
White	85.3	85.4	88.5	-3.3	N/S
African American	6.8	6.9	4.9	1.9	N/S
Other	7.8	7.6	6.6	1.2	N/S
Unknown	0.1	0.1	0.0	0.1	N/S
Health status					
Recalculated HCC score					
Mean	1.8	1.8	1.8	0.1	N/S
Low: > 1.35 and < 2.00	64.1	65.3	68.6	-4.5	N/S
Medium: > 2.00 and < 3.10	22.6	22.2	18.5	4.1	N/S
High: > 3.10	13.3	12.5	12.9	0.4	N/S
Baseline costs >= $6000	67.5	67.3	64.7	2.8	N/S
Charlson comorbidity index – Mean	2.5	2.5	2.3	0.1	N/S
Chronic conditions					
HF	18.8	18.2	15.9	2.9	N/S
Coronary artery disease	32.9	33.3	30.9	2	N/S
Other respiratory disease	22.6	22.7	19.7	2.9	N/S
Diabetes without complications	29.3	28.5	26.0	3.3	N/S
Diabetes with complications	6.1	6.1	3.6	2.6	*
Essential hypertension	57.3	57.4	56.4	0.9	N/S
Valve disorders	4.6	4.7	5.0	-0.4	N/S
Cardiomyopathy	2.4	2.5	1.5	0.9	N/S
Acute and chronic renal disease	10.8	10.9	7.3	3.5	*
Renal failure	2.8	2.7	2.7	0.1	N/S

(continued)

Table 4-4 (continued)
Characteristics of the TST CMHCB demonstration program intervention population by participation status

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 (continued)					
Peripheral vascular disease	5.0	5.1	5.4	-0.4	N/S
Lipid metabolism disorders	21.3	22.0	20.0	1.3	N/S
Cardiac dysrhythmias and conduction disorders	17.8	18.1	17.6	0.2	N/S
Dementias	2.0	1.3	1.8	0.1	N/S
Strokes	6.3	6.3	5.3	1	N/S
Chest pain	10.8	10.9	8.1	2.7	N/S
Urinary tract infection	15.2	14.3	12.4	2.8	N/S
Anemia	13.2	12.7	13.5	-0.4	N/S
Malaise and fatigue (including CFS)	6.5	6.3	7.1	-0.6	N/S
Dizziness, syncope, convulsions	11.3	11.0	11.7	-0.5	N/S
Disorders of joint	14.7	14.8	13.6	1.2	N/S
Hypothyroidism	14.5	14.3	13.1	1.4	N/S

NOTES: TST = Texas Senior Trails; 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/texas/final/9mo/tables/tableTT-4.sas 30JUNE2009

Table 4-5
Participation rates during the first 6 months of the TST CMHCB demonstration by beneficiary characteristics, baseline characteristics, and intervention period health status

Characteristics	Participation rate
Overall participation rate [1,2]	89
Beneficiary characteristics	
Male	88
Female	89
White	89
African American/other/unknown	89
Age < 65 years	88
Age 65-74	90
Age 75-84	89
Age 85 + years	87
Medicaid	84
Non-Medicaid	89
Urban	89
Rural	89
Baseline characteristics	
Baseline HCC score low	89
Baseline HCC score high	87
Low baseline PBPM	89
High baseline PBPM	87
Baseline Charlson score low	88
Baseline Charlson score medium	89
Baseline Charlson score high	89
Intervention period health status	
Died	80
Alive	90
Institutionalized	81
Not institutionalized	89
Concurrent HCC score low	88
Concurrent HCC score high	88
Number of participants	4,238
Number of total beneficiaries	4,778

NOTES: TST = Texas Senior Trails; 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/texas/final/9mo/partab2.sas 27JULY2009

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 CMOs 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 most were interested in examining which beneficiary characteristics had on 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***).

Table 4-6 presents the results of the logistic regression analyses that predict participation based on various beneficiary characteristics for the original population 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 reflects the relatively high participation rate and also suggests that TST had cast a fairly wide net when engaging their intervention population. Pseudo R-squares for most of the models were 0.01, with the full Model 3 exhibiting the highest pseudo R-square of 0.03. ***Supplement 4A*** contains tables that present the odds ratios and level of significance for Models 1 and 2.

Overall, there are very few significant effects observed in Model 3, which likely reflects the small size of the TST nonparticipating population. TST had only 467 beneficiaries in its nonparticipant group when evaluating the full intervention period (and once we adjusted for periods of ineligibility and death, the number of participants decreased to 331). Examining the full Model 3, we observe that the very old (85 years of age and older), Medicaid enrollees, institutionalized beneficiaries, and those who died during the intervention period were less likely to be participants. This suggests that TST was not as successful at gaining participation agreements from the sicker beneficiaries. Baseline 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. Comparing the full model to the model containing only beneficiary characteristics, we can see that when adding baseline health status the odds ratios modestly change (increase) for beneficiaries 85 years of age and older and Medicaid enrollees. This confirms that being older and enrolled in Medicaid is correlated with health status. The low variation in the full model, as indicated by the pseudo R-square (0.03), indicates TST did not pursue any restrictive targeting of beneficiaries.

4.3.4 Level of Intervention

In this section, we report the frequency of interaction between TST and intervention beneficiaries for a subset of beneficiaries who were fully eligible and participating in the

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

Characteristics	Model 1 OR	p^3	Model 2 OR	p^3
Intercept	4.77	**	4.98	**
Beneficiary Characteristics				
Male	0.94	N/S	0.99	N/S
African American/other/unknown	1.05	N/S	1.06	N/S
Age < 65 years	0.82	N/S	0.85	N/S
Age 75-84	0.91	N/S	0.96	N/S
Age 85 + years	0.59	**	0.73	**
Medicaid	0.62	**	0.65	*
Urban	1.03	N/S	1.02	N/S
Baseline characteristics				
Baseline HCC score medium	N/I	N/I	1.04	N/S
Baseline HCC score high	N/I	N/I	0.93	N/S
Medium baseline PBPM	N/I	N/I	1.05	N/S
High baseline PBPM	N/I	N/I	0.94	N/S
Baseline Charlson score medium	N/I	N/I	1.15	N/S
Baseline Charlson score high	N/I	N/I	1.09	N/S
Intervention period health status				
Died	N/I	N/I	0.68	N/S
Institutionalized	N/I	N/I	0.30	N/S
Concurrent HCC score medium	N/I	N/I	1.00	N/S
Concurrent HCC score high	N/I	N/I	0.95	N/S
Number of Cases	4,966	N/A	4,966	N/S
Chi-Square (p<)	29.05	**	151.62	**
Pseudo R-square	0.01	N/A	0.03	N/S

NOTES: TST = Texas Senior Trails; 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/S means not statistically significant.

N/I means not included.

N/A means not applicable.

The baseline HCC score reference group is <2. The age reference group is 65-74 years. The PBPM reference group is LT $528. The concurrent HCC score reference group is .315 or less.

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

Program: partab3 27JUL2009, partab4b 29JULY2009

CMHCB program from month 4 through month 15. 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, TST 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 hospitalizations.

Descriptive statistics were performed on beneficiaries participating in the TST demonstration program to determine the breadth and depth of contacts related to care management (CM). The data represent beneficiaries who were fully eligible and participating (unless they died) for months 4 through 15 of the demonstration. A total of 3,607 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 3,428 full-time equivalent beneficiaries.

Table 4-7 displays the overall distribution of CM–related contacts. The mean number of CM contacts for each beneficiary was six and the median was four. Slightly less than 40% of beneficiaries had five or more contacts; 28% of beneficiaries had two or fewer CM contacts.

Table 4-7
Distribution of number of contacts with participants in the TST CMHCB demonstration

Statistic	Number	Percent
Number of beneficiaries[1]	3,607	—
FTE beneficiaries[2]	3,428	—
Mean number of contacts	6	—
Median number of contacts	4	—
Distribution low to high contact variables	FTE beneficiaries	Percent
0-2 contacts	960	28.0%
3-4 contacts	1,152	33.6%
5+ contacts	1,315	38.4%
Total	3,427	100.0%

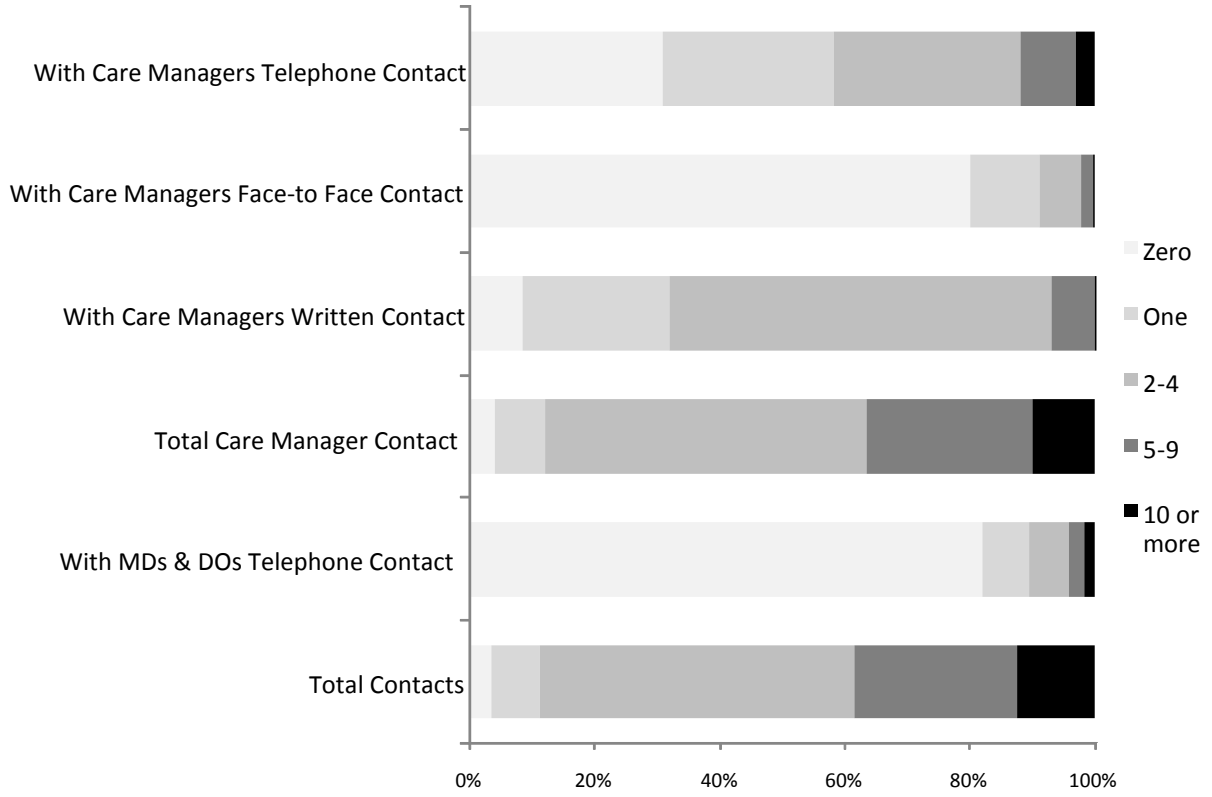
NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries; FTE = full time equivalent.

[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.

Figure 4-1 displays the percent of participants with completed care manager calls, face-to-face visits, written communication as well as physician calls and total contacts by frequency of contact over the 12 months. When all modes of contact are combined (total contacts), 97% of beneficiaries had one or more contacts with a care manager or physician, 50% had two to four contacts and about 25% had five to nine contacts. Written contact was the most frequent form of communication; 91% of beneficiaries received at least one or more written contacts from a care manager. Face-to-face contact was the least frequent form of contact; 80% of beneficiaries had no face-to-face contact with a care manager. Physician CM-related contact is defined as telephone communication with a beneficiary. Of intervention beneficiaries, 82% had no contact with a physician, while 2% had 10 or more CM contacts.

Figure 4-1
Percent distribution of participants[1,2] with telephone calls, face-to-face contact, written communication, provider calls, and total contacts: TST



NOTES: TST = Texas Senior Trails; MD = physician; DO = osteopath.

[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.

TST conducted risk assessments of its participant population to inform the development of individualized care plans and assign participants to one of three risk categories to determine the level of service to be provided to each participant. We hypothesized that the number of CM contacts a beneficiary had would rise with his or her HCC risk score, leading to the next analysis.

Tables 4-8 and ***4-9*** display the frequency of care management contacts (none or 1 or more) by baseline HCC score. Contact rates are stratified by mode of contact by care managers. There was only telephone contact between physicians and participants. Beneficiaries in the low HCC risk group were almost equally likely to have had one telephone conversation with a care manager as beneficiaries in the high HCC risk group. Almost one-third of beneficiaries in the high risk group had face-to-face contact with a care manager; almost twice the percent in the low risk group. Written contact was the more prevalent form of communication between the care manager and beneficiary. High risk beneficiaries were slightly less likely to have any written contact, in contrast to telephone or face-to-face contact, compared with low risk beneficiaries (81% versus 93%).

Table 4-8
Frequency of TST care management contacts by HCC score for full participants during demonstration months 4-15

Clinician and contact mode	HCC score High (>3) frequency N = 473 (13%)	HCC score %	HCC score Medium (2-3) frequency N = 823 (23%)	HCC score %	HCC score Low (<2) frequency N = 2311 (64%)	HCC score %	HCC score Total
Care manager							
Telephone							
0	161	34.0%	277	33.7%	728	31.5%	1,166
≧[1]	312	66.0	546	66.3	1,583	68.5	2,441
Total	—	100.0	—	100.0	—	100.0	3,607
Face-to-face							
0	338	71.5	596	72.4	1,944	84.1	2,878
≧[1]	135	28.5	227	27.6	367	15.9	729
Total	—	100.0	—	100.0	—	100.0	3,607
Written							
0	87	18.4	149	18.1	160	6.9	396
≧[1]	386	81.6	674	81.9	2,151	93.1	3,211
Total	—	100.0	—	100.0	—	100.0	3,607
Physician							
0	340	71.9	643	78.1	1,974	85.4	2,957
≧[1]	133	28.1	180	21.9	337	14.6	650
Total	—	100.0	—	100.0	—	100.0	3,607

NOTES: TST = Texas Senior Trails; HCC = Hierarchical Condition Category.

[1]≧ = one or more interventions

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

Table 4-9
Frequency of TST care management services by HCC score, full participants during months 4-15

Content	HCC score High (>3) frequency N = 473 (13%)	HCC score: %	HCC score Medium (2-3) frequency N = 823 (23%)	HCC score %	HCC score: Low (<2) frequency N = 2311 (64%)	HCC score %	HCC score Total
Care plan							
0	39	8.2%	48	5.8%	92	4.0%	179
≥[1]	434	91.8	775	94.2	2,219	96.0	3,428
Total	—	100.0	—	100.0	—	100.0	3,607
Discharge plan							
0	340	71.9	607	73.8	1,854	80.2	2,801
≥[1]	133	28.1	216	26.2	457	19.8	806
Total	—	100.0	—	100.0	—	100.0	3,607
Health risk assessment							
0	380	80.3	670	81.4	1,951	84.4	3,001
≥[1]	93	19.7	153	18.6	360	15.6	606
Total	—	100.0	—	100.0	—	100.0	3,607
Advanced directive							
0	403	85.2	712	86.5	1,975	85.5	3,090
≥[1]	70	14.8	111	13.5	336	14.5	517
Total	—	100.0	—	100.0	—	100.0	3,607
Education (diet, exercise, medication, self-care)							
0	406	85.8	702	85.3	1,918	83.0	3,026
≥[1]	67	14.2	121	14.7	393	17.0	581
Total	—	100.0	—	100.0	—	100.0	3,607
Care management adherence–guidelines							
0	445	94.1	768	93.3	2,106	91.1	3,319
≥[1]	28	5.9	55	6.7	205	8.9	288
Total	—	100.0	—	100.0	—	100.0	3,607

NOTES: TST = Texas Senior Trails; HCC = Hierarchical Condition Category.

[1]≥ = one or more interventions

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

When compared with care managers, physicians were less likely to make CM contacts, as evidenced by the high frequency of zero calls in this category. Physicians were more likely than care managers to concentrate their CM telephone contacts in the high risk group. In the high risk group, 28% of beneficiaries had at least one contact with a physician, compared with only 15% in the low risk group; this is a 2-to-1 difference. Most of the CM contacts with beneficiaries were with care managers, and when there was a CM-related contact with a physician, high risk beneficiaries were usually involved.

We also examined the frequency with which different care management services were provided, stratified by HCC score. We focus upon the six services that occurred most frequently; many of the other services were provided to only 1 or 2% of participants. The development of a care plan and discharge planning were the two most frequently utilized care management interventions; respectively, 96% and 22% of beneficiaries received one or more of each of these interventions over the 12-month period. Health risk assessments were conducted on 17% of participants; advanced directives were discussed with 14% of participants; education about diet, exercise, medication, and self-care was provided to 16% of participants; and adherence to medication guidelines was evaluated for 8% of participants.

In the low HCC score group, 96% of beneficiaries had some care planning, compared with 92% in the high risk group. One would expect beneficiaries in the high risk group to have more care planning, but that did not appear to be the case. Beneficiaries in the low risk group were 8 percentage points less likely to have had any discharge planning assistance than beneficiaries in the high risk group. This result is not surprising if the high severity group had more admissions than the low risk group. The third most frequent care management intervention involved health risk assessments; the data provided to us indicated that only 17% of all participants received a documented assessment. There is only a 4 percentage point difference in rate of receipt between the high and low risk groups. About 15% of both high and low risk groups had counseling about the development of an advanced directive. Beneficiary education for diet, exercise, medications, self–management, and adherence to medication guidelines was relatively low, with the higher risk beneficiaries receiving less than low risk beneficiaries. This is consistent with TST's stated plan to provide education only to those in their lowest risk stratum.

To more directly examine the targeting strategy of TST, a multivariate logistic regression model was estimated using the number of total contacts as the dependent variable. The model estimated the likelihood of a participant receiving a high number of contacts. The medium contact group was omitted, thus comparing the high contact group with the low contact group. ***Table 4-10*** 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 4-15 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 51% greater than those for a beneficiary age 65 and older, adjusting for any baseline difference in HCC score and characteristics. Conversely, beneficiaries who were older than 85 years were 28% less likely ($p < .05$) to be in the high contact category. Living in an urban area also increased the likelihood of being in the high contact category by 66%.

Table 4-10
Logistic regression modeling results comparing the likelihood of being in the TST high contact (5+) category relative to the low (0-2) contact category

Characteristics	Odds ratio
Demographics[1]	
Male	1.02
African American/other/unknown	1.20
Age <65	1.51**
Age 75-84	0.92
Age 85+ years	0.72*
Urban	1.66**
Medicaid beneficiary	0.72
Baseline Health Status and Utilization[2]	
Baseline HCC score medium (2.00 - 3.1)	1.05
Baseline HCC score high (>3.1)	1.22
Baseline Charlson score 2-4	1.19
Baseline Charlson score 5+	1.27
Baseline Medicare Payments[3]	
Medium base PBPM ($530-1612)	1.03
High base PBPM ($1612+)	1.12
Demonstration Period Utilization and Health Status[4]	
One hospitalization	1.82**
Multiple hospitalizations	2.89**
Died	0.62*
NH/LTC/SNF resident	0.10**
Concurrent HCC score medium (0.315-1.121)	1.27*
Concurrent HCC score high (>1.121)	1.19
Number of cases	2,429
Chi-square ($p<$)	324.98**
Pseudo R-square	0.13

NOTES: TST = Texas Senior Trails; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month; NH = nursing home; LTC = long-term care; SNF = skilled nursing facility. Weighted by demonstration fraction of eligible days.

* denotes statistical significance at the 5% level.

** denotes statistical significance at the 1% level.

[1] The age reference group is 65-74 years.

[2] The baseline HCC score reference group is <2.0.

[3] The PBPM reference group is LT $280.

[4] The concurrent HCC score reference group is 0.315 or less.

SOURCE: Medicare 2006-2007 claims.

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 monthly Medicare payments during the baseline year. The presence of high levels of baseline severity or costliness was not a positive predictor of being in the high contact group. In contrast, demonstration period acute care utilization was a powerful explainer of more CM contacts. Beneficiaries who had multiple hospitalizations were almost 3 times more likely to be in the high contact group (2.89), and those who had one hospitalization were almost 2 times more likely (1.82) to be in the high contact group than those with no hospitalizations. The 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 in a nursing home, long term care facility, or skilled nursing facility prior to the demonstration were less likely to be in the high contact category. These findings suggest that TST made a focused effort to contact beneficiaries who were at risk of hospitalization or who had been hospitalized, a key stated component of their program.

4.4 Summary

For TST, we find that participants were healthier and younger than beneficiaries who never participated. 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. Given that TST was both provider- and hospital-based, it is surprising that the program was unable to get the sickest eligible beneficiaries to participate. The concurrent HCC score—calculated for the first 6 months of the intervention period, baseline Charlson score, and baseline expenditures had no predictive power for participation. If CMS desires broadly focused care management programs, these findings suggest alternative recruiting and outreach strategies are needed to reach the sicker beneficiaries as well as dual Medicare/Medicaid enrollees and beneficiaries with disabilities or institutionalized.

Although there was no predetermined expected number of contacts, the TST beneficiaries were a sick and costly group of FFS beneficiaries, and they reported significant unmet clinical and psychosocial needs. Across a 12-month intervention period and for fully eligible and participating beneficiaries, the median number of contacts was four. We found some evidence that TST targeted their intervention contacts to beneficiaries who were at risk of hospitalization or who had been hospitalized. However, other than development of care or discharge plans, we found limited intervention services for all beneficiaries. Given the low level of interaction between TST and many of its participants, it is unlikely that TST could be successful at changing beneficiary behavior with respect to self-management of chronic illnesses. Findings from the beneficiary survey show that there was little meaningful improvement in self-care activities during the demonstration period.

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 the Texas Senior Trails (TST) CMHCB demonstration program improve quality of care, as measured by improvement in the rates of beneficiaries receiving guideline concordant care?

In this chapter, we present analyses related to clinical quality performance during the TST 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 that 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 the care management organization (CMO) go-live date and for the last 12 months of the demonstration. Only beneficiaries who had at least 1 day of eligibility in both baseline and the last 12 months of the demonstration 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 described in Chapter 2, we elected a somewhat different approach for the quality of care

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

Statistics	All	COPD	Diabetes	Ischemic vascular disease
Original population				
Months 5-16				
Intervention				
Total number of beneficiaries	4,717	867	1,483	1,687
Full time equivalents[1]	4,708	866	1,479	1,684
Comparison				
Total number of beneficiaries	4,932	967	1,442	1,999
Full time equivalents[1]	4,922	965	1,417	1,994

NOTES: TST = Texas Senior Trails; 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.

measures. Medicare claims for the full baseline and intervention period were included regardless of beneficiary eligibility for the CMHCB demonstration, for example, nonpayment of the Part B premium for 1 or 2 months. This allowed us to provide credit to the 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 last 12 months of the intervention period, 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, and 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 TST 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 not statistically different from 1.0. We report the odds ratio associated with the β 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 TST CMHCB demonstration 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.[6] Only beneficiaries who had at least one 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.

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

5.3 Findings

Process-of-care rates per 100 TST CMHCB demonstration 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.

To better understand the movement underlying the reported difference-in-differences rates, we stratified the TST CMHCB demonstration 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.[7] Only beneficiaries who had at least one 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.

At baseline, rates for the five measures in the comparison group ranged from a low of 21% for oxygen saturation for beneficiaries with COPD to a high of 88% for HbA1c testing for beneficiaries with diabetes. Rates in the comparison group either remained the same or modestly declined over the course of the 16-month demonstration period. At baseline, we also observe that rates for four of the five measures were 3 to 9 percentage points lower in the intervention group than in the comparison group. During our site visit, Texas Senior Trails told us that they found a larger than expected percentage of their intervention beneficiaries without a usual source of primary care. One of the primary goals of the TST CMHCB program was to facilitate patient relationships with primary care physicians and specialists so that patients could receive appropriate preventive care to manage their illnesses and become informed consumers of health care. Further, access to transportation for primary care services was also a major concern. TST participants were eligible to receive support from one of three social workers who addressed issues such as transportation. Alternately, participants residing in Lubbock who were unable to travel to a physician's office could receive home-based care from a program called Senior House Calls, which was run by the Texas Tech University Health Sciences Center (TTUHSC) School of Nursing.

---

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

Table 5-2
Comparison of rates of guideline concordant care for the last 12 months of the TST CMHCB Demonstration period with rates for a 1-year period prior to the start of the TST demonstration

Process of care measures	Rate per 100 Baseline I[1]	Rate per 100 Baseline C[1]	Rate per 100 Demo period I[1]	Rate per 100 Demo period C[1]	D-in-D Rate per 100[1]	D-in-D OR	D-in-D p	D-in-D CI Low	D-in-D CI High
Months 5–16									
All Beneficiaries									
Influenza Vaccine	48	56	53	57	4	1.17	0.01	1.04	1.31
Beneficiaries with COPD									
Oxygen Saturation Test	22	21	13	19	-6	0.63	0.01	0.44	0.89
Beneficiaries with diabetes									
HbA1c Test	85	88	82	84	0	1.07	0.66	0.88	1.43
LDL-C Test	74	83	74	81	2	1.18	0.22	0.91	1.52
Beneficiaries with IVD[2]									
LDL-C Test	72	79	70	75	2	1.14	0.23	0.92	1.41

NOTES: TST = Texas Senior Trails; CMHCB = 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.

[1] All rates are per 100 beneficiaries and are adjusted for periods of demonstration eligibility during the 1-year period prior to the start of the demonstration and 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 Quality Forum definition.

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

Between baseline and the last 12 months of the CMHCB demonstration, there is no pattern of quality improvement across the five measures. The rate of influenza vaccine increased in the intervention population by 4 percentage points more than the comparison group. At the same time, the rate of beneficiaries with COPD who had an annual oxygen saturation test declined by 6 percentage points more than the comparison group. There were no improvements in rates of HbA1c or LDL-C testing among the intervention beneficiaries with diabetes, or rates of LDL-C testing among the intervention beneficiaries with ischemic vascular disease.

Table 5-3 displays the percentages of TST CMHCB demonstration beneficiaries who did or did not receive one of the four process-of-care measures during the baseline and last 12 months of the demonstration. We display the distribution of their 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. Over two-thirds of 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 roughly 40% of beneficiaries were compliant in both time periods for influenza vaccination. However, across the five measures beneficiaries in the intervention group were less likely to be always compliant and more likely to be never compliant. For both groups, beneficiaries 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; nearly 90% 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 their 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 too large of a denominator population—or inclusion of beneficiaries that have a higher level of pulmonary function than—than envisioned by the AMA-PCPI. On the other hand, it is difficult to imagine that nearly 85% of an ill Medicare fee-for-service (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.

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 TST 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 TST CMHCB demonstration program. Out of five measures, there was only one observed increase in

Table 5-3
Percentage of comparison and intervention beneficiaries meeting process of care standards in the baseline year and last 12 months of the TST CMHCB demonstration

Original population	HbA1c Testing[1] C	HbA1c Testing[1] I	LDL-C[1,2] C	LDL-C[1,2] I	LDL-C[1,3] C	LDL-C[1,3] I	Oxygen Assessment[1] C	Oxygen Assessment[1] I	Influenza Vaccine[1] C	Influenza Vaccine[1] I
Always compliant	79	75	73	61	66	58	7	5	44	36
Became noncompliant	10	10	12	13	14	15	15	17	13	12
Never compliant	6	8	7	13	11	15	66	70	29	35
Became compliant	5	6	8	13	9	12	12	8	14	17

NOTES: TST = Texas Senior Trails ; CMHCB = Care Management for High Cost Beneficiaries; LDL-C = low-density lipoprotein cholesterol; I= intervention population; C = comparison population

[1] All percentages are adjusted for periods of beneficiary CMHCB demonstration eligibility during the one-year period prior to the start of the demonstration and the last 12 months the CMO was active. Only beneficiaries who had at least one day of eligibility in both the baseline and demonstration periods are included in this analysis.

[2] Beneficiaries with diabetes.

[3] Beneficiaries with ischemic vascular disease.

SOURCE: RTI analysis of 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: qcvars, gcctab3.sas.

rate of receipt of evidence-based care (influenza vaccination); while at the same time there was a decrease in the rate of receipt of another process-of-care measure (oxygen saturation assessment).

Over the course of the demonstration, TST 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. With the exception of oxygen saturation assessment, we also observe between roughly one-quarter to one-half of beneficiaries were not compliant during the last year of the CMHCB demonstration despite focused efforts by TST to encourage beneficiaries to become compliant. 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 Texas Senior Trails (TST) Medicare Care Management for High Cost Beneficiaries (CMHCB) demonstration program by answering the following two evaluation questions:

- Did the TST program improve intermediate health outcomes by reducing acute hospitalizations, readmissions, and emergency room (ER) utilization?
- Did the TST 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 the last 12 months of the Texas Senior Trails demonstration relative to a 12-month baseline period. We also examine differences in the rate of mortality between the intervention and comparison beneficiaries during the demonstration.

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 TST demonstration program date and for the last 12 months of the intervention period. We constructed rates of all-cause hospitalization and ER visits as well as 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 last 12 months of the demonstration 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 the last 12 months of the intervention period, 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 TST 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 TST's program. For the demonstration period, we included index admissions from 13 months through 2 months prior to the end of the demonstration. 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 TST 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 TST 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 intervention and comparison groups between the go-live date and the end of the demonstration period. Date of death was obtained from the Medicare enrollment data base (EDB). Statistical comparison of the mortality rates was made using a *t*-test of differences in mean rates between the intervention and comparison groups.

We also conducted a multivariate logistic regression analysis to determine the predictors of mortality controlling for baseline differences in beneficiary demographic and health status characteristics between the intervention and comparison groups. The logistic model used in this analysis is the same as that described in equation 4.1. The variables included in the model are defined as follows:

- male, a dichotomous variable, set at 1 for males;
- African American/other/unknown, a dichotomous variable, set at 1 for beneficiaries whose race code is African American, other, or unknown;
- 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 years, and age greater than or equal to 85 years; age 65-74 years is the reference group;

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

- 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 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 between 2 and 3 (medium) and 4 or greater (high); Charlson score less than 2 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 CMHCB demonstration program was between $528 and less than $1,612 (medium) and $1,612 or greater (high); PBPM cost less than $528 is the reference group; and

- 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.

6.3 Findings

6.3.1 Rates of Hospitalizations and Emergency Room Visits

Rates of hospitalization and ER visits per 1,000 beneficiaries for the year prior to go-live and the last 12 months of the TST demonstration period 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 last 12 months of the demonstration. 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. Positive D-in-D rates as statistically determined through the IRR indicate that the comparison group exhibited lower rates of growth of hospitalization or ER visits or greater decline than the intervention group.

The baseline rates of hospitalization and ER visits were very high in the TST intervention and comparison populations. The baseline rate of all-cause hospitalization was 774 per 1,000 intervention group beneficiaries. And, the baseline rate of all-cause ER visits was even higher at 985 per 1,000 intervention beneficiaries. We observe similar baseline rates of all-cause hospitalization for the intervention and comparison groups; however, the intervention group had a substantially higher rate of all-cause ER visits at baseline than the comparison group.

Table 6-1
Comparison of rates of utilization for the last 12 months of the TST CMHCB demonstration with rates of utilization for a 1-year period prior to the start of the TST demonstration

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 5-16									
Hospitalizations									
All causes	774	780	735	635	106	1.17	0.00	1.06	1.28
10 ACSCs[5]	208	200	242	192	41	1.21	0.02	1.02	1.42
ED/Obs visits									
All causes	985	761	810	701	-115	0.89	0.11	0.78	1.03
10 ACSCs	136	106	115	106	-20	0.85	0.18	0.68	1.08

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

[1] The baseline period is the 1-year period prior to the go-live date of the care management organization (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 the intervention period.

[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 D-in-D 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, chronic obstructive pulmonary disease and chronic bronchitis, dehydration, bacterial pneumonia, septicemia, ischemic stroke, and urinary tract infection.

SOURCE: RTI analysis of 2004-2008 Medicare enrollment, eligibility, claims and encounter data; Computer runs: acsc01 acsc02 acsctab acsc acsctab1 25JUN2009

The 10 ACSC reasons for hospitalization combined accounted for roughly one-quarter of all-cause hospitalizations and about 15% of all-cause ER visits. Thus, Medicare fee-for-service (FFS) beneficiaries in the TST 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.

The rate of all-cause hospitalization declined in both the intervention and the comparison groups between the baseline and demonstration periods. However, there was a statistically significant greater decline in the all-cause hospitalization rate for the comparison group than in the intervention group. The D-in-D rate is 106 per 1,000 beneficiaries higher in the intervention group. There was also a statistically significant greater decline in the 10 ACSC hospitalization rate in the comparison group than in the intervention group. The 10 ACSC hospitalization D-in-D rate is 41 per 1,000 beneficiaries higher in the intervention group.

Simultaneously, we observe a decline in the rate of all-cause and ACSC ER visits in both the intervention and comparison groups with a greater level of decline in the intervention group. However, the declines in the D-in-D rates of -115 per 1,000 beneficiaries for all-cause ER visits and -20 per 1,000 beneficiaries for ACSC ER visits are not statistically significant.

6.3.2 Rates of 90-Day Readmissions

Table 6-2 displays the number of beneficiaries included in the readmission analyses. ***Table 6-3*** displays the percent of beneficiaries with an admission and 90-day readmission and rate of readmission per 1,000 beneficiaries. Data are displayed for all-cause and ACSC admissions and readmissions. As shown in ***Table 6-3***, the change from the baseline to the demonstration period in the percent of beneficiaries with an admission of any type was statistically significant. Although both groups have a decline in the percent of beneficiaries with an all-cause admission, the intervention group's decline was less, resulting in a positive and statistically significant D-in-D rate of a modest 3 per 1,000 beneficiaries.

Table 6-2
Number of beneficiaries included in analyses of readmissions for TST

Counts of beneficiaries	Intervention	Comparison
Original beneficiaries	—	—
Total number of beneficiaries	4,903	5,034
Full time equivalents[1]	4,894	5,026

NOTES: TST = Texas Senior Trails.

[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.

Table 6-3
Change in 90-day readmission[1] rates between the year prior to the TST CMHCB demonstration and months 2-13 of the demonstration

Utilization	Baseline rate per 1,000[1,2,3] I	Baseline rate per 1,000[1,2,3] C	Demo rate per 1,000[1,2,3] Final I	Demo rate per 1,000[1,2,3] Final C	D-in-D	OR/IRR[4]	*p*	Low CI	High CI
Months 2-13									
Hospitalizations									
Percent with an admission	48	47	37	34	3	1.14	0.02	1.02	1.28
Percent with ACSC[4] admission	16	15	15	12	1	1.12	0.16	0.96	1.32
All-cause 90-day readmission									
Percent with readmission	31	31	35	35	-1	0.97	0.72	0.80	1.17
Readmission rate / 1,000	542	529	680	667	1	1.00	0.96	0.84	1.18
ACSC same-cause 90-day readmission									
Percent with readmission	10	11	13	10	3	1.36	0.21	0.84	2.18
Readmission rate / 1,000	146	148	174	168	8	1.05	0.85	0.64	1.73

NOTES: TST = Texas Senior Trails; CMHCB = Medicare Care Management for High Cost Beneficiaries; I= intervention population; C = comparison population; D-in-D = difference-in-differences; OR = odds ratio; IRR = incidence rate ratio; ACSC = ambulatory care sensitive condition.

[1] 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 the demonstration period.

[2] Only beneficiaries who at least 1 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, chronic obstructive pulmonary disease and chronic bronchitis, dehydration, bacterial pneumonia, septicemia, ischemic stroke, and urinary tract infection.

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims, and intervention data; Computer runs: readm04 readm05 tab6_3_1x 25JUN2009

We observe no other statistically significant differences in percent with a readmission or readmission rates. Although the percent of beneficiaries with an all-cause or ACSC readmission did not change materially over time, the readmission rates per 1,000 beneficiaries grew from the baseline to the last 12 months of the demonstration period. The growth in the readmission rate was relatively greater for all-cause readmissions than for same ACSC cause. The increase over time in the readmission rate per 1,000 beneficiaries was greater than the increase in the likelihood of having a readmission, indicating that the likelihood of having more than one readmission grew over time.

6.3.3 Mortality

Table 6-4 displays rates of mortality during the TST CMHCB demonstration period for the intervention and comparison populations. Over the 16-month demonstration period, 14% of the intervention group beneficiaries and 11% of the comparison group beneficiaries died. This is a statistically significant 3 percentage point difference.

Table 6-4
Mortality rates during the TST CMHCB Demonstration

Description	Intervention number of deaths	Percent	Comparison number of deaths	Percent	Difference	*p*-value
Original population (16 months)	705	14.2	547	10.8	3.4	0.00

NOTE: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries.

SOURCE: RTI analysis of 2004-2008 Medicare enrollment database.

Computer runs: mortality.sas 05JUN2009

As noted in Chapter 4, the intervention group at the start of the TST demonstration program had higher rates of beneficiaries eligible for Medicare because they were disabled, enrolled in Medicaid, under the age of 65, or minorities. These characteristics are often proxies for poor health status. We also observed considerable variation in baseline rates of chronic conditions. The intervention group had higher rates of conditions such as diabetes with complications, hypertension, and peripheral vascular disease and lower rates of coronary artery disease, cardiac dysrhythmias, and lipid metabolism disorders. Given the relatively short intervention period, an imbalance in these characteristics between the intervention and comparison group beneficiaries could influence the observed mortality rates in the demonstration period.

To control statistically for any potential imbalance, we estimated a multivariate logistic regression model of likelihood of death with beneficiary demographic and health status characteristics as control variables. ***Table 6-5*** displays the ORs and statistical significance (p value) from the logistic regression model of likelihood of death during the demonstration period. The intervention status variable (listed first) has an OR of 0.89, which is not statistically different

Table 6-5
Logistic regression analysis of likelihood of death during the TST CMHCB demonstration

Characteristics	OR	$p^1<$
Intervention status	0.89	N/S
Beneficiary characteristics		
Male	1.26	**
African American/other/unknown	0.79	N/S
Age < 65 years	0.79	N/S
Age 75-84	1.60	**
Age 85 + years	3.08	**
Medicaid	1.26	**
Urban	1.10	*
Baseline characteristics		
Baseline HCC score medium	1.79	**
Baseline HCC score high	3.16	**
Medium baseline PBPM	1.29	N/S
High baseline PBPM	1.34	N/S
Baseline Charlson score medium	1.17	N/S
Baseline Charlson score high	1.37	N/S
Demonstration period health status		
Institutionalized	2.13	**
Number of cases	10,046	N/A
Chi-square ($p<$)	368.62	**
Pseudo R-square	0.04	N/A

NOTES: TST = Texas Senior Trails; CMHCB = Medicare Care Management for High Cost Beneficiaries; OR = odds ratio; HCC = Hierarchical Condition Category; PBPM = per beneficiary per month.

[1]* denotes statistical significance at the <5% level;** denotes statistical significance at the <1% level; N/S means not statistically significant; N/A means not applicable..

SOURCE: RTI analysis of Medicare enrollment, eligibility, claims and encounter data; Computer runs: diereg1 22AUG2009

from 0. Thus, death rates during the demonstration period were equivalent between the intervention and comparison groups after controlling for poorer baseline and demonstration period health status.

6.4 Conclusions

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

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

We did not find evidence of improvement in health outcomes as measured by a reduction in acute hospitalizations, 90-day readmissions, ER visits, or mortality. In fact, we observe an increase in the rate of all-cause and ACSC hospitalizations among the intervention group relative to the comparison group. Although we observed a higher mortality rate within the intervention group during the demonstration, a multivariate logistic regression model that controlled for imbalances in baseline health status between the two groups equalized the raw mortality rates.

Two key foci of the TST program were to provide to beneficiaries (1) guidance and support to reduce ER utilization and (2) hospital discharge planning support. TST used the Senior Health Profiles® to identify beneficiaries at risk for repeated hospital admissions and placed them in their medium-risk intervention (i.e., telephonic care management support conducted every 2 weeks to every 2 months) or high-risk intervention (i.e., more frequent telephonic care management and in-person visits from a care manager). Through this intervention strategy, TST had expected to be successful in reducing ER usage and readmissions. We find no evidence to support their expectation.

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 16 months that the Texas Senior Trails (TST) CMHCB program was in operation. The evaluation questions we address are:

- What were the Medicare costs per beneficiary per month (PBPM) in the base year versus the first 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 16-month period compare with the fees that were paid out? How close was TST 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 in Medicare costs for beneficiaries in the intervention and comparison groups?

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

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

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

The rest of this chapter has five sections. The next two sections describe our data sources, variable construction, and analytic methods. ***Section 7.4*** presents our primary findings on trends in PBPM costs between base and demonstration periods. ***Section 7.5*** shows PBPM cost savings in relation to average monthly fees and whether TST achieved budget neutrality using RTI's costing methods. ***Section 7.6*** stratifies PBPM costs and savings by high cost and high risk categories. ***Section 7.7*** 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 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. To be eligible, a beneficiary had to have had (a) a Hierarchical Condition Category (HCC) score greater than 1.7 and/or (b) annual costs greater than $6,000 in the selection year. 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, 25% of beneficiaries had neither annual base year costs greater than $6,000 nor an HCC score above 1.7.

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 CMO's target area. 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 Per Beneficiary Per Month Costs

All financial analyses were conducted on a PBPM cost basis, or the ratio of eligible Medicare costs to eligible months. The baseline period is defined as 365 days (or 1 year) prior to TST's start date. The 16-month demonstration period includes 487 days (16 months × 30.42 days/month) after the start date.

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

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

To statistically test hypotheses regarding *trends* in beneficiary costs, average PBPM costs first must be calculated at the beneficiary level. Constructing individual PBPM costs required dividing a beneficiary's total cost during eligible periods by his or her own fraction of eligible months during the base year and demonstration period. Most beneficiaries had 12 months of base year eligibility and 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 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 $9,460. Beneficiaries eligible for a year or more averaged $1,424 in monthly costs. Thus, beneficiaries with very truncated eligibility averaged monthly costs 6.6 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.)

Variation can be reduced by trimming high PBPM cost 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.

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 10/487 = 0.0205, or about 50 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.

Table 7-1
TST CMHCB PBPM mean costs by eligible days, intervention group, base year

Eligible days[1]	N (%)[2]	PBPM	Range
< 10	15 (0.3%)	$9,460	$0-$42,691
11-30	40 (0.8)	13,179	0-73,762
31-60	51 (1.0)	9,406	0-45,504
61-90	70 (1.4)	8,724	0-101,606
91-365	606 (12.2)	3,508	0-43,254
366+	4,184 (84.3)	1,424	0-48,783
Mean	4,966	1,982	0-101,606

NOTES: TST = Texas Senior Trails; CMHCB = Care Management High Cost Beneficiaries; PBPM = per beneficiary per month.

[1]Number of days beneficiary eligible for intervention.

[2]N(%): number of beneficiaries (percent of all eligibles).

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

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 TST, its negotiated management fee was $117.

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 TST 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 TST's start date and the intervention period to construct a change in costs. This was done for all beneficiaries in both the intervention and comparison groups, thereby producing a paired comparison within group. Next, we determined the mean difference in the differences in PBPM cost growth rates for each group, treating the mean differences as independent samples.[9] The strength of first calculating the change in PBPM costs at the beneficiary level is that it completely controls for any unique clinical and socioeconomic characteristics that might differ between the intervention and comparison groups. Any imbalances in beneficiary characteristics that might produce intertemporal differences in medical utilization or costs are factored out using first-differencing. Our gross savings rate, in equation form, is

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

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

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

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

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 0, or no savings. Gross savings must be sufficiently greater than zero to assure the government that the measured savings rate was not due to chance.[10] A critical evaluation question is the power we had to detect relatively small savings rates. By "detectable" we mean the rate of savings that would force us to reject the null hypothesis of no savings at all. Having completed the demonstration, we now have the information on both the level and variation in savings rates that allows us to calculate the detectable savings threshold for TST.

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

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

and the detectable threshold is

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

---

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

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

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 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 detectable threshold (DT) as a percent of the comparison group's demonstration mean PBPM cost, or $DT/PBPM_c$.

Table 7-2 shows the variation that exists in the (unweighted) PBPM costs in the base year prior to the start date and the demonstration period for TST's intervention and comparison samples. Mean PBPM costs in the base period ranged from a low of $0 to a high of $211,224 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 20-25% greater than mean costs).

Table 7-2
TST CMHCB PBPM cost distribution thresholds, comparison and intervention group, base and demonstration period

Quantiles[1]	Base year comparison	Base year intervention	Demonstration period comparison	Demonstration period intervention
(N)	(5,080)	(4,966)	(5,080)	(4,966)
Minimum	$0	$0	$0	$0
<10%	129	134	99	88
<25%	345	358	259	269
Median	883	949	732	900
>25%	2,031	2,152	2,014	2,427
>10%	3,723	3,968	4,073	4,656
Maximum	20,482	34,483	211,224	101,606
Mean	1,536	1,625	1,694	1,982
CV	1.22	1.25	2.39	1.94

NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; CV = coefficient of variation.

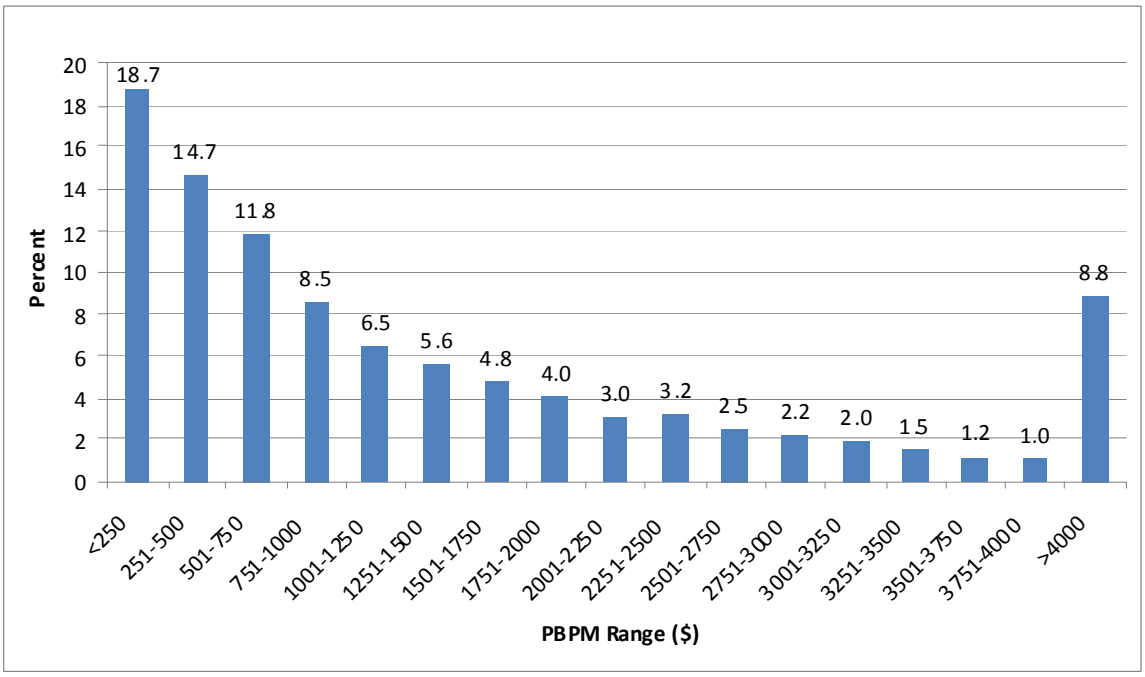
[1]<10%, <25%, >25%, >10%: PBPMs below or above percentage.

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

CVs doubled or more during the demonstration period, implying growing variation in monthly costs across beneficiaries. 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 more than double median costs in the intervention period, indicating a strong right tail of high costs. Maximum values show how high PBPM costs can be before weighting—greater than $100,000 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.

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 much as a 5% difference between the entire intervention and comparison group's PBPM cost trend was still 70%.[12]

Figure 7-1
Frequency distribution of PBPM costs, comparison group, base year: TST



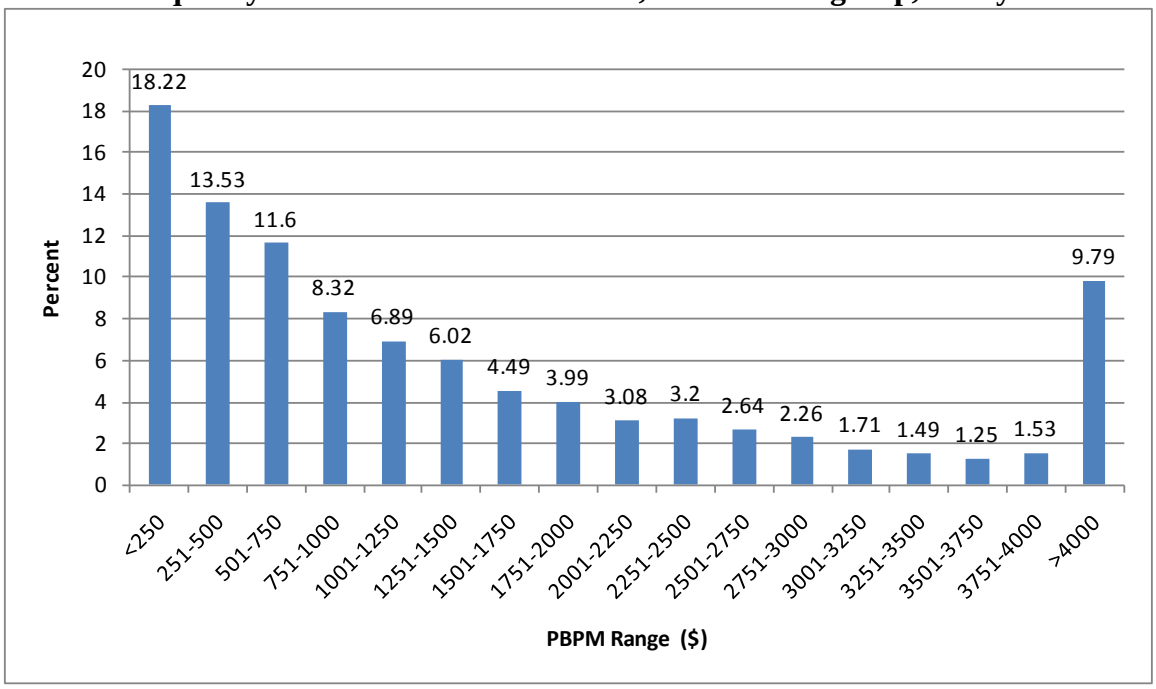
NOTES: PBPM = per beneficiary per month; TST = Texas Senior Trails.

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

SOURCE: Medicare Part A & B claims; COSTRUN2 (July 9, 2009).

12 Power = $\Phi[-z_{.975} + \Delta / SE_{[\Delta I \ \Delta C]}]$. For TST, Power = F [-1.96 + .05*$1,383/$48] = F [-.52] = 0.70. 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)

Figure 7-2
Frequency distribution of PBPM costs, intervention group, base year: TST



NOTES: PBPM = per beneficiary per month; TST = Texas Senior Trails.

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

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

7.3.3 Budget Neutrality

Each CMO is obligated to produce net savings for the Medicare program. Unless the CMO achieves 5% 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 95% 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 4% 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: $PBPM_I - [0.95PBPM_c + MF]$.

CMS also agreed to share a portion of the savings with the CMO, according to the following percentages for disposition of net savings:

- 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%).

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, two budget neutrality tests exist. The most important, and standard, test is how confident we are that a CMO achieved any savings at all: the *t*-test against zero savings. A second test is whether the upper confidence threshold of actual gross savings encompasses the monthly fee. For example, if a CMO saved $50 per beneficiary per month and its monthly fee was $100, we report whether 1.96 times the standard error of the $50 estimate exceeded $100. It is possible for actual savings to be simultaneously insignificant from 0 and the monthly fee, if the monthly fee is low relative to the standard error in PBPM costs. We do not conduct the second test of gross savings against the monthly fee, if gross savings are negative rather than positive.

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 regression-to-the-mean effects (see ***Supplement 7A***). 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 regression-to-the-mean. 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 Per Beneficiary Per Month Cost Levels and Trends

Table 7-3 displays PBPM cost levels and rates of growth in average PBPM costs between the 12-month base year and the 16-month demonstration period. 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.

The weighted base year average PBPM cost was $69 greater ($p$ = insig) in the intervention group versus the comparison group ($1,543 versus $1,474), or 5%.[13] The intervention-comparison gap in PBPM Medicare costs widened to $192 ($p < .01$) in the demonstration period ($1,575 versus $1,383). Between the base year and the 16-month demonstration period, the average comparison group PBPM cost declined significantly by $91 ($p < .01$), while the intervention group's PBPM average Medicare costs actually increased by $32 ($p$ = insig). Thus, the intervention group's PBPM cost grew $123 faster ($p < .01$) than the comparison group's PBPM cost. Intervention beneficiaries, who were 5% more costly at baseline, became 14% more costly, on average, than the comparison group after 16 months.

Participants in the TST intervention group were $68 ($p$ = insig) more costly than the comparison group beneficiaries ($t = 1.8$) in the base period and nonparticipants were $80 more costly ($p$ = insig). The participation rate, based on beneficiaries used in this cost analysis, was 91% (4,491/4,958 – 1). Nonparticipants became $11 less costly ($p$ = insig) during the demonstration period while participants became $36 more costly ($p$ = insig). The participant group's PBPM cost increased $127 faster ($p < .01$) than the comparison group's while the nonparticipant group's PBPM cost grew $80 faster ($p$ = insig) than the comparison group's PBPM cost.

---

13 The intervention group's PBPM cost *2* years prior to going live (or the period used to select the intervention beneficiaries and identify the matched comparison beneficiaries) was roughly 3% higher (not shown).

Table 7-3
TST CMHCB PBPM cost growth rates between base year and 16-month demonstration period, intervention and comparison groups

Study group	Beneficiaries	Base year PBPM Mean[1]	Base year PBPM SE	Demo PBPM Mean[1]	Demo PBPM SE	Differences in means	SE
Intervention	4,966	$1,543**	27.5	$1,575**	33.4	$32.2	37.0
Participants	4,499	1,542**	28.9	1,577**	32.5	35.6	36.9
Non-Participants	467	1,554**	90.0	1,543**	182.4	-11.0	181.0
Comparison	5,075	1,474**	25.2	1,383**	27.6	-91.0**	31.1
Differences							
I – C	N/A	68.6	37.3	191.8**	43.2	123.1**	48.2
Participants – C	N/A	67.7	38.2	194.3**	42.4	126.6**	47.9
Nonparticipants – C	N/A	80.1	98.2	160.0	116.8	80.0	128.3
Participants – Nonparticipants	N/A	-12.3	104.9	34.3	127.5	46.6	141.2

NOTES: TST = Texas Senior Trails; 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; N/A means not applicable.

SOURCE: Medicare Part A&B claims; run texas/final/9mo/costrun2(7/9/09).

7.5 Savings and Budget Neutrality

Table 7-4 presents summary statistics on savings from the TST intervention. 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. Texas Senior Trail's monthly fee is reported also as a percentage of the comparison group's PBPM cost.

Over the course of the 16-month intervention, average monthly costs increased $32 in the intervention group while at the same time falling $91 in the comparison group. The result was a $123 greater relative increase in PBPM cost growth in the intervention group. This positive difference implies dissavings rather than savings, at a rate of 8.9% of the comparison group's PBPM cost.

With roughly 5,000 beneficiaries, each, in the intervention and comparison group, we had the power to detect a savings of $95 or more (in negative terms) at the 95% confidence level. This rate is 6.8% of the comparison group's PBPM cost, implying substantial power to detect relatively modest savings, if they existed.

Table 7-4
TST CMHCB average PBPM gross savings, fees, and budget neutrality status

Description	PBPM cost change
PBPM Outlay Change	
Intervention group	$32
Comparison group	-$91
Difference	$123
Gross (Dis)savings%[1]	8.90%
Minimal Detectable Savings[2]	
Absolute dollars	-$95
% of comparison PBPM	-6.80%
Monthly Fee	
Absolute dollars	$117
% of comparison PBPM	8.50%
Net Fee	
Absolute dollars[3]	$240
% of comparison PBPM	17.40%

NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries.

[1]Gross Savings % = Difference in PBPM outlay changes as % of comparison PBPM. Negative values imply true savings; positive values = dissavings.

[2]Minimum Detectable Savings = 1.96*standard error of difference in mean PBPM changes.

[3] Net Fee absolute dollars = Monthly fee + Difference in PBPM outlay change.

SOURCE: Medicare Part A&B Claims; TST Contract Terms & Conditions; Texas/final/9mo/Cost4b (7/23/09).

TST's monthly fee was $117, which amounted to 8.5% of the comparison group's PBPM during the demonstration period. Thus, TST would have had to achieve 13.5% savings in order to retain all of its fees—at least according to RTI's calculations, which are not official under financial reconciliation.[14]

If one believed that, in fact, TST's intervention actually increased beneficiary costs by $123, then the net effect on Medicare costs would be $123 plus $117, or $240 per beneficiary per month. Instead of reducing the 8.5% fee outlay on PBPM costs, the intervention would appear to have increased Medicare total costs from 8.5% to 17.4% of the comparison group's costs.

7.6 Imbalances between Intervention and Comparison Samples

In the TST demonstration, beneficiaries were eligible if they incurred more than $6,000 in the year prior to selection *or* had an HCC score greater than 1.7. Once eligible intervention beneficiaries were identified, comparison group beneficiaries from the same counties were then "matched" by RTI based on PBPM cost quintiles, as requested by TST. In balancing the two samples on costs alone, the comparison group still could exhibit a different frequency of HCC scores, age, Medicaid status, and other characteristics. Any remaining within-quintile differences in beneficiary characteristics justified our decision to calculate changes in costs for each beneficiary and then compare mean changes in the two samples, thereby controlling for any unique beneficiary-specific differences in base year costs.

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. Beneficiary eligibility was determined using claims from September 1, 2004, through August 31, 2005. The base period for our cost analyses was April 1, 2005, through March 31, 2006. Differential attrition between intervention and comparison groups could produce further deviations in underlying beneficiary characteristics, including their costliness just prior to the start date. The way the two samples were matched in the original sampling period and our recalculation of costs and HCC scores in the subsequent base year raise two possibilities:[15]

- That there were different frequencies of high cost and/or high risk beneficiaries in the intervention and comparison samples in the base year just prior to the start date or

[14] ARC's estimate of the difference in growth rates was 5.5% in an analysis conducted a few months earlier with a shorter run-out period. ARC's interim estimate of higher PBPM growth for Texas Tech was 7% after trimming outliers and adjusting for base year differences (ARC, 2008, Table 6). ARC's final estimate of intervention savings will involve capping costs at the 99th percentile level and using more claims from a longer run-out period.

[15] Solucia Research Company (2008) highlighted these concerns and more in its review of population influences. Their major focus was on the selection of the comparison group and why the one developed by RTI may not accurately portray what cost trends would have been for intervention beneficiaries in lieu of the demonstration.

- That there were different frequencies of several health status and sociodemographic characteristics in the intervention and comparison groups, resulting in different base period trends in costs in the two groups.

It is possible that high cost and high risk beneficiaries exhibit opposing regression-to-the-mean 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 TST'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.[16]

For differences in other beneficiary characteristics to have any effect on intervention savings, two things must happen. First, one or more characteristics must have a statistically important effect on PBPM cost growth rates. Second, unless the same important characteristics also significantly differ, numerically, between the intervention and comparison groups, they will not affect the intervention savings rates. Because most characteristics are simple binary (0, 1) indicators, there must be substantial numbers of "costly" beneficiaries involved and not just a large differences in relative frequencies.

7.6.1 Frequencies of Beneficiary Characteristics

Table 7-5 shows that the intervention and comparison groups were nearly identically distributed by cost and risk during the randomization period. Roughly 50% of beneficiaries who originally qualified for the demonstration had both an HCC score greater than 1.7 and annual costs above $6,000. Only 15% qualified as only high risk beneficiaries. At the point of selection into the demonstration, no beneficiaries failed both criteria. In the subsequent period in which base period costs were determined, the frequency of beneficiaries that continued to have high costs and high HCC scores fell to slightly under 40%, and the overall frequency of high cost–only beneficiaries fell about 5 percentage points. By the year that base costs were determined, roughly one-quarter of beneficiaries did not meet either selection criterion, 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.

Table 7-5 does indicate that, unlike the cost and risk groups, intervention group beneficiaries did differ from the comparison group in having relatively fewer Medicare-aged beneficiaries and more beneficiaries who were either younger than 65 or disabled (19% vs. 8%), African Americans or other minorities (14% vs. 7%), Medicaid enrollees (3.7% vs. 1.5%), or urban residents (75% vs. 61%). No statistically significant differences were found by gender or

16 Solucia's (unidentified) sample of Medicare beneficiaries outside Texas showed a 36% decline in PBPM costs for high cost beneficiaries and a 415% increase for high risk–only beneficiaries between the base and demonstration period. Solucia further shows that simply adjusting demonstration PBPM costs by HCC score would not completely eliminate regression-to-the-mean effects due to varying base year costs.

Table 7-5
Frequency distribution of TST beneficiary characteristics, intervention and comparison groups, base and demonstration years

Characteristics	Intervention (%) Base	Intervention (%) Demo	Comparison (%) Base	Comparison (%) Demo
COSTRISK				
B112 > = 6,000	29.25%	34.9%	28.7%	34.9%
Both	38.1	50.0	37.1	49.3
HCC > 1.7	7.1	15.1	7.4	15.9
Neither	25.6	0.0	26.8	0.0
Agegp				
<65	18.57	N/A	8.11	N/A
65-69	15.00	N/A	14.27	N/A
70-74	18.85	N/A	21.81	N/A
75-79	17.33	N/A	23.91	N/A
80-84	15.18	N/A	17.80	N/A
85+	15.08	N/A	14.09	N/A
Male				
No	61.13	N/A	59.59	N/A
Yes1	38.87	N/A	40.41	N/A
new_race				
Black	6.62	N/A	3.19	N/A
Other	7.75	N/A	4.23	N/A
White	85.63	N/A	92.58	N/A
MEDICAID				
No	96.34	N/A	98.49	N/A
Yes	3.66	N/A	1.51	N/A
DISABLED				
No	81.47	N/A	91.89	N/A
Yes	18.53	N/A	8.11	N/A
Urban				
No	24.97	N/A	39.18	N/A
Yes	75.03	N/A	60.82	N/A
LTCB				
No	97.67	N/A	98.29	N/A
Yes	2.33	N/A	1.71	N/A
SNFB				
No	88.37	N/A	89.9	N/A
Yes	11.63	N/A	10.1	N/A
Nh_B112				
No	90.10	N/A	96.81	N/A
Yes	9.90	N/A	3.19	N/A

NOTES:TST = Texas Senior Trails.
Beneficiaries weighted by fraction of eligible days in demonstration period.
Base Year: 12 months prior to intervention start date (4/1/2005-3/31/2006).
Sample Year: 12 months prior to selection (9/1/2004-8/31/2005). N/A means not applicable.
SOURCE: Medicare Part A and B claims; base year: texas/final/9mo/Cost4b (7/23/09); sample year: Cost4 (11/7/2008).

institutionalized status in a long-term care or skilled nursing facility (SNF). Intervention beneficiaries, however, were 3 times more likely than comparison beneficiaries to have had a nursing home stay in the year prior to the demonstration. It is possible that systematic differences in beneficiary mix, coupled with sizable effects on PBPM cost trends, could explain the higher rate of cost increases in the intervention group. We explore the effects of any imbalances between the intervention and comparison groups in the next two sections.

7.6.2 PBPM Cost Levels and Trends by Cost and Risk Group

Table 7-6 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. Intervention beneficiaries qualifying only with an HCC score greater than 1.7 averaged PBPM costs of just $282 compared with high cost–only beneficiaries ($1,365; 4-8 times greater) and both high cost and high risk beneficiaries ($2,810; 10 times greater). The gap narrowed considerably in the demonstration period as a result of opposing cost trends. Both high cost groups experienced declines in their PBPM costs while the high risk–only group experienced an average increase of $937 ($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 7-6, we find even larger cost declines in the high cost comparison groups compared with the intervention groups. The high cost and high risk comparison group saw its costs decline $625 ($p < .01$) compared with only $505 ($p < .01$) in the intervention group, net increase in the latter of $120 (p = insig). The difference is even greater in the high cost–only group ($189; $p < .01$). By contrast, costs rose slower in the high risk–only comparison group, resulting in a net cost increase in the intervention group of $126 (p = insig). Intervention effects in the neither group corresponded to those in the high risk–only group ($119; $p < .05$).

Table 7-7 better demonstrates the extent of regression-to-the-mean 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 regression-to-the-mean phenomenon. Unweighted mean costs were $1,536 in the base period, with an overall increase of $158. Cost increases are inversely correlated with a beneficiary's base period PBPM costs. At the extremes, beneficiaries with less than $250 in base period costs saw their average costs increase by $780 while those with initial costs greater than $4,000 experienced average decreases of $2,224. On average, one-half of beneficiaries had base period costs less than $883 versus $732 in the demonstration period; yet, mean PBPM costs increased. Thus, more beneficiaries in the demonstration period had costs lower than the original $883 median. The fact that mean PBPM costs still rose in the demonstration period means that beneficiaries with rising costs from the base to intervention period more than offset the greater number of beneficiaries who experienced lower costs.

Table 7-6
TST CMHCB PBPM costs by cost and risk group, intervention and comparison groups, base and demonstration periods

Description	High-cost and high-risk PBPM	High-cost and high-risk SE	High-cost only PBPM	High-cost only SE	High-risk only PBPM	High-risk only SE	Neither PBPM	Neither SE
Intervention (N)	(1,986; 40%)	N/A	(1,409; 28%)	N/A	(354; 7%)	N/A	(1,217; 25%)	N/A
Base year	$2,810**	55.4	$1,365**	23.9	$282**	6.6	209**	4.0
Demonstration	2,306**	64.5	1,324**	45.5	1,219**	163.7	872**	40.5
Difference	-505	76.1	-41	48.5	937**	163.6	663**	40.2
% change	-18%	N/A	-3%	N/A	232%	N/A	217%	N/A
Comparison (N)	(1,969; 39%)	N/A	(1,416; 28%)	N/A	(376; 7%)	N/A	(1,319; 26%)	N/A
Base year	2,676**	51.0	1,405**	25.7	275**	6.3	218**	3.9
Demonstration	2,052**	54.7	1,176**	43.6	1,086**	96.5	762**	32.1
Difference	-625**	64.4	-229**	49.3	811**	95.8	544**	32.0
% change	-23%	N/A	-16%	N/A	194%	N/A	150%	N/A
Differences-in-differences								
Intervention	-505**	76.1	-41	48.5	937**	163.6	663**	40.2
Comparison	-625**	64.4	-229**	49.3	811**	95.8	544**	32.0
Differences-in-differences	120	99.6	189	69.2	126	186.2	119	51.0

NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month; SE = standard error. N/A means not applicable.

Beneficiary PBPM weighted by fraction of eligible days in demonstration period.

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

SOURCE: Medicare Part A & B claims; texas/final/9mo/Cost4b (7/23/09).

Table 7-7
TST CMHCB trend in comparison group PBPM costs by base period PBPM

Base year PBPM level	N	Base year	Demonstration period	Change
< 250	905	$121	$901	$780
250-500	672	364	1,280	916
500-750	576	621	1,525	904
750-1,000	413	870	1,580	710
1,000-1,250	342	1,123	1,808	684
1,250-1,500	299	1,370	2,031	661
1,500-1,750	223	1,621	2,299	677
1,750-2,000	198	1,863	1,883	21
2,000-2,250	153	2,118	2,539	412
2,250-2,500	159	2,364	2,358	-6
2,500-2,750	131	2,619	2,697	79
2,750-3,000	112	2,881	4,749	1,867
3,000-3,250	85	3,126	2,779	-347
3,250-3,500	74	3,364	2,932	-432
3,500-3,750	62	3,619	2,752	-867
3,750-4,000	76	3,880	3,434	-445
> 4,000	486	6,329	4,104	-2,224
Mean	5,080	1,536	1,694	158
Median	—	883	732	-151

NOTES: TST = Texas Senior Trails; CMHCB = Care Management for High Cost Beneficiaries; PBPM = per beneficiary per month.

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

7.6.3 Multivariate Regression Tests of Intervention Savings

Three sets of regression coefficients in ***Table 7-8*** test the intervention effect by using the beneficiary's base year PBPM cost (PBPM_base) to explain each beneficiary's demonstration period PBPM cost. Controlling for each beneficiary's base period PBPM cost, the intervention coefficient of 168.7 is highly significant ($p < .01$) implying faster cost growth, on average, in the intervention relative to the intervention group. In other words, given the average change in PBPM costs relative to the base period, intervention beneficiaries experienced higher increases than did the comparison beneficiaries.

The base period PBPM cost coefficient (0.336; $p < .01$), when combined with the intercept coefficient, implies substantial regression-to-the-mean of costs (= 0.336 - 1 = -0.664). Imagine two comparison group beneficiaries, one with a relative low ($500) and another with a relatively high ($2,500) PBPM cost in the base period. The predicted PBPM cost of the initially "low cost" comparison beneficiary would more than double during the intervention period, while that of the "high cost" beneficiary would decline by about one-quarter.[17]

Regression-to-the-mean effects are quite substantial but clearly not in one direction. Including only high cost beneficiaries would clearly have produced sizable declines in PBPM costs during the demonstration. This effect, however, was almost completely offset by rising costs from the high risk but not high cost group (and reinforced by the neither group). The net decline in the comparison group was $91 (see ***Table 7-3***). Also note that the standard error of comparison group costs was slightly higher in the demonstration period.

The second regression model controls for which cost-risk group the beneficiary was in during the base period. The key intervention coefficient is unaffected. This is true even though each of the three cost-risk groups is highly different from the neither group. The high cost plus high risk group also exhibits much greater cost increases than the other groups, i.e., $732 ($p < .01$) over and above the regression-to-the-mean effect. The intervention coefficient remains positive and highly significant because the intervention and comparison groups were balanced with regards to cost and risk in the base period.

In the third model, the coefficients controlling for beneficiary characteristics has the effect of making the intervention coefficient only slightly less positive, but it remains statistically significant ($134.7; $p < .01$). 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 gender, race, and age coefficients are insignificant, as are the Medicaid enrollee, disabled status, and urban coefficients. Statistical insignificance for some characteristics is due to controlling for both the beneficiary's base period PBPM cost as well as

[17] The calculation is as follows:

PBPM[base]	PBPM[demo]	PBPM Change	%Change
$500	$1,225	$725	+145%
$2,500	$1,897	$603	- 24%

Table 7-8
Regression results: TST intervention gross savings controlling for base period PBPM and beneficiary characteristics

Independent variable	PBPM_Demo Coefficient	PBPM_Demo *t*	PBPM_Demo Coefficient	PBPM_Demo *t*	PBPM_Demo Coefficient	PBPM_Demo *t*
Interception	1056.8**	30.9	849.4**	18.6	1,277.0	1.8
Intervention	168.7**	4.1	167.7**	4.1	134.7**	3.2
PBPM_Base	0.336**	30.4	0.248**	18.6	0.285**	17.5
High cost–high risk	N/I	N/I	732.5**	11.8	684.1**	10.3
High cost	N/I	N/I	141.2**	2.5	109.1	1.9
High risk	N/I	N/I	318.8**	3.7	294.1**	3.4
Male	N/I	N/I	N/I	N/I	-37.4	0.9
African American	N/I	N/I	N/I	N/I	56.7	0.6
Other	N/I	N/I	N/I	N/I	-67.6	0.8
Age 65-69	N/I	N/I	N/I	N/I	-637.0	0.9
70-74	N/I	N/I	N/I	N/I	-524.5	0.7
75-79	N/I	N/I	N/I	N/I	-464.7	0.7
80-84	N/I	N/I	N/I	N/I	-418.7	0.6
85+	N/I	N/I	N/I	N/I	-465.5	0.7
Medicaid	N/I	N/I	N/I	N/I	-44.9	0.3
Disabled	N/I	N/I	N/I	N/I	-306.6	0.4
Urban	N/I	N/I	N/I	N/I	58.3	1.3
LTC	N/I	N/I	N/I	N/I	-660.9**	3.9
SNF	N/I	N/I	N/I	N/I	-187.6*	2.4
NH	N/I	N/I	N/I	N/I	218.8*	2.4
R^2	0.086	N/A	0.101	N/A	0.105	N/A
N	10,045	N/A	10,045	N/A	10,045	N/A

NOTES: TST = Texas Senior Trails; PBPM = per beneficiary per month; LTC =long-term care; SNF =skilled nursing facility; NH = nursing home.

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

PBPM_Demo: Beneficiary's average PBPM during demonstration.

PBPM_Base: Beneficiary's average PBPM in base period.

High Cost-High Risk: PBPM > $6,000 and HCC > 1.7 in base year.

High Cost: PBPM > $6,000 and HCC< 1.7.

High Risk: PBPM < $6,000 and HCC > 1.7.

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

NH = 1 if MD visited beneficiary in nursing home within three months of state date.

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

SOURCE: Medicare Part A & B claims; Texas/final/9mo/Cost4b (7/23/09).

which cost-risk group (s)he was in during the base year prior to the start date. All three indicators of being institutionalized are negative and statistically significant. Thus, once the cost-risk group and base period cost of long-term and SNF institutionalized beneficiaries was accounted for, they were actually less costly. Likely, the large regression-to-the-mean effect (0.285 - 1 = -0.715) is understating to some extent the decline in their costs from the base year to the demonstration period. The nursing home population, by contrast, continued to have significant cost increases.

Table 7-9 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 = 0.285 * ($1,543 - $1,474) = $19.70; high cost / risk = 684 * (0.381 - 0.371 = $6.80). 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 $192 (see ***Table 7-3***). The "pure" intervention effect explains 70% ($135 / $192) of the $192 difference, leaving 30% to be explained by unbalanced characteristics. Alternatively, unbalanced characteristics, including the difference in base period PBPM costs, explain only 30% ($57) of the demonstration period cost difference. The difference between base period mean PBPM costs can account for roughly 10% ($19.70) of the $192 difference in demonstration period PBPM costs between the intervention and comparison groups. The slight imbalance in the high cost–high risk group added 3.5% to the overall greater costs in the intervention group. The greatest effects occur between ages 70 and 84 that together explained about 30% of the higher costs—again holding all other variables constant. The relatively high number of Medicare disabled in the intervention group actually lowered costs somewhat (16.7% of the $192 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 fact that nursing home beneficiaries were 3 times as common in the intervention group as in the comparison group added 8% to intervention costs.

7.7 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. Cost inequalities increased substantially between the base period and the demonstration period; thus having large sample sizes to test for intervention effects was important. With roughly 5,000 beneficiaries, each, in the intervention and comparison group, we had the power to detect a savings of $95 or more (in negative terms) at the 95% confidence level. This rate is 6.8% of the comparison group's PBPM cost, implying substantial power to detect relatively modest savings, if they existed.

No savings, however, were found for the intervention. In fact, cost growth was statistically greater ($123 per beneficiary per month), not less, in the intervention group compared with the comparison group. This increase was not due to nonparticipants in the intervention. Costs growth among participants was even slightly higher ($126). Texas Senior

Trails not only failed to achieve budget neutrality using RTI's method; it also doubled the $117 cost of Medicare's per beneficiary monthly management fee.

Table 7-9
Marginal effects of beneficiary characteristics on TST PBPM cost growth

Regression variable	Coefficient	I-value	C-value	I - C impact	% PBPM change
PBPMb	0.285	1543	1474	19.7	10.2
Highcost/risk	684	0.381	0.371	6.8	3.5
High cost	109	0.293	0.287	0.6	0.3
High risk	294	0.071	0.074	-0.9	-0.5
Male	-37.4	0.611	0.596	-0.6	-0.3
African American	56.7	0.066	0.032	1.9	1.0
Other NW	-67.6	0.078	0.042	-2.4	-1.2
Age 65-69	-637	0.150	0.143	-4.7	-2.4
70-74	-525	0.189	0.218	15.5	8.1
75-79	-465	0.173	0.239	30.6	15.9
80-84	-419	0.152	0.178	11.0	5.7
85+	-466	0.151	0.141	-4.6	-2.4
Medicaid	-44.9	0.037	0.015	-1.0	-0.5
Disabled	-307	0.185	0.081	-32.0	-16.7
Urban	58.3	0.750	0.608	8.3	4.3
LTC	-661	0.023	0.017	-4.1	-2.1
SNF	-188	0.116	0.101	-2.9	-1.5
NH	219	0.099	0.032	14.7	7.7
Sum Total	—	—	—	56.0	29.2

NOTES: TST = Texas Senior Trails; PBPM = per beneficiary per month; LTC =long-term care; SNF =skilled nursing facility; NH = nursing home.

SOURCE: Medicare Part A & B claims.

It does appear that intervention beneficiaries, at least during the base year just prior to the demonstration start date, were 4.5% more expensive, even though both samples were selected within five cost quintiles using data from a year prior to the demonstration period. Consequently, in the greater than 1-year gap between selecting the beneficiaries and the start date, base year costs widened somewhat. This justified CMS' decision to make a statistical adjustment in its financial reconciliation process by adjusting each beneficiary's demonstration period cost by their own base year cost, which was somewhat higher on average in the intervention group.

Despite this cost difference, no evidence was found that the intervention and comparison groups were particularly unbalanced, or dissimilar, with respect to cost and risk group. Any

residual difference that existed in cost occurred within each cost-risk group. We did find that matching on cost quintiles did not balance the two groups on several other patient characteristics. The intervention group had 2.5 times the percentage of disabled beneficiaries, twice the percentage of minorities, more than twice the percentage of dual Medicare-Medicaid enrollees, and more than 3 times the number of nursing home beneficiaries prior to the start date.

Concerns over these imbalances prompted an examination of cost trends, first, within cost and risk groups. In the two groups in which beneficiaries qualified because of high base year costs, average costs fell less in the intervention group than in the comparison group, not more. Those qualifying only due to high risk had average costs rising more rapidly in the intervention than in the comparison group. We found, therefore, that changes in costs during the demonstration rose substantially when they had been low initially and fell substantially when they had been high in the base year. This suggests that marked regression-to-the-mean was taking place.

Regression methods were used to statistically control for both regression-to-the-mean and any biases against the intervention due to initial imbalances in several beneficiary characteristics. The control variables had no material effect on the difference in growth rates in the two groups. Intervention cost growth remained statistically greater, not less, than in the comparison group. PBPM costs were nearly 14% greater in the intervention group than in the comparison group during the 16-month intervention period. Only 30% of the difference could be explained by regression to the mean or varying beneficiary characteristics. Fully 70% must be attributable to the more costly intervention. Where imbalances were large in percentage terms, they were small, numerically (e.g., 1%-3% of the population), had minor effects on cost *growth* (as distinct from initial cost *levels*), or were in favor of the intervention.

CHAPTER 8
KEY FINDINGS FROM THE TST CMHCB DEMONSTRATION EVALUATION

The purpose of this report is to present the findings from RTI International's evaluation of Texas Tech University Health Sciences Center and its Texas Senior Trails (TST) Medicare Care Management for High Cost Beneficiaries (CMHCB) demonstration program. Our evaluation focuses upon three broad domains of inquiry:

- **Implementation.** To what extent was TST able to implement is program?
- **Reach.** How well did TST engage its intended audience?
- **Effectiveness.** To what degree was TST 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 16 months of TST operations. Our findings are based on the experience of approximately 10,000 ill Medicare beneficiaries assigned to an intervention or a comparison group. Five key findings on participation, beneficiary satisfaction, clinical quality and 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 TST demonstration program.

Of all TST intervention beneficiaries, 91% verbally consented to participate in the CMHCB demonstration at some point during the intervention period. In spite of this high participation rate, we found that the participant population was healthier and younger than beneficiaries who never participated. 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. Given that TST was both provider- and hospital-based, it is surprising that they were unable to get the sickest eligible beneficiaries to participate. These findings suggest alternative recruiting and outreach strategies are needed to reach the sicker beneficiaries as well as dual Medicare/Medicaid enrollees and beneficiaries who have disabilities or are institutionalized.

Key Finding #2: The intensity of intervention with the participating beneficiaries is unlikely to produce significant behavioral change and savings.

Although there was no pre-determined expected number of contacts, the TST beneficiaries were a sick and costly group of FFS beneficiaries, and they reported significant

unmet clinical and psychosocial needs. Across a 12-month intervention period that reflects maximum intervention months for fully eligible and participating beneficiaries, the median number of contacts was four. We found some evidence that TST targeted their intervention contacts to beneficiaries who were at risk of hospitalization or who had been hospitalized. However, we found limited intervention services for all beneficiaries other than developing a care plan or a discharge plan. Given the low level of interaction with many of their participants, it is unlikely that they could be successful at changing beneficiary behavior with respect to self-management of their chronic illness leading to a reduction in use of acute care settings and savings.

Key Finding #3: There was no improvement in beneficiary satisfaction, experience with care, self-management, or physical and mental health functioning.

The TST CMHCB demonstration program employed strategies to improve quality of care while reducing costs by empowering Medicare beneficiaries to better manage their care. They did so 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 and should encourage more effective interactions between beneficiaries and their primary health care providers. TST hypothesized that lifestyle changes and better communication with providers would mitigate acute flare-ups in 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.

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. Among the 19 outcomes covered by the survey, only one statistically significant, counterintuitive, group difference was found: members of TST's intervention group were *less* (not more) certain that they could take their medications as prescribed. There was no improvement in beneficiary satisfaction, care experience, self-management, or physical and mental health functioning.

Key Finding #4: TST had a positive intervention effect on one of five quality of care process measures but no positive intervention effect on reduction in rates of hospitalization, emergency room visits, 90-day readmissions, or mortality.

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 TST CMHCB demonstration program. Out of five measures, there was only one observed increase in rate of receipt of evidence-based care (influenza vaccination) while at the same time there was a decrease in the rate of receipt of another process-of-care measure (oxygen saturation assessment).

Over the course of the demonstration, TST had expected to increase beneficiaries' rate of adherence to evidence-based care. Yet, between one-quarter to one-half of beneficiaries were not compliant with evidence-based guidelines during the last year of the CMHCB demonstration despite focused efforts by TST to encourage beneficiaries to become compliant. 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.

We did not find evidence of improvement in health outcomes as measured by a reduction in acute hospitalizations, 90-day readmissions, ER visits, or mortality. In fact, there was a statistically significant increase in the rate of all cause and ambulatory care sensitive condition (ACSC) hospitalizations among the intervention group relative to the comparison group. Although we observed a statistically higher mortality rate in the intervention group during the demonstration, multivariate analysis that controlled for imbalances in baseline health status between the two groups equalized the raw mortality rates.

Key Finding #5: Medicare cost growth was greater in the intervention group than the comparison group.

No Medicare savings were found for the intervention. In fact, cost growth was statistically greater ($123 per beneficiary per month), not less, in the intervention group compared with the comparison group. This increase was not due to nonparticipants in the intervention as participants had even slightly higher cost growth, on average ($126).

Not only did TST fail to achieve budget neutrality using RTI's evaluation method, it also doubled the $117 cost of Medicare's monthly per beneficiary management fee. If one believed that, in fact, TST's intervention actually increased beneficiary costs by $123, then the net effect on Medicare costs would be $123 plus $117, or $240 per beneficiary per month. Instead of reducing the 8.5% fee outlay on per beneficiary per month (PBPM) costs of the comparison group, the intervention would appear to have increased Medicare total costs from 8.5% to 17.4% of the comparison group's costs.

Multivariate regression methods were used to statistically control for both regression-to-the-mean and any biases against the intervention caused by initial imbalances in several beneficiary characteristics. The control variables had no material effect on the difference in growth rates in the two groups. Intervention cost growth remained statistically greater, not less, than in the comparison group. PBPM costs were nearly 14% greater in the intervention group than in the comparison group during the 16-month intervention period, compared with being only 5% higher in the base year. Because demonstration beneficiaries were selected, in part, because of their high pre-demonstration costs, both intervention and comparison groups experienced significant regression-to-the mean effects. Beneficiaries with initially high costs experienced large declines during the demonstration period while initially lower cost beneficiaries saw even larger increases, percentagewise. Yet, only 30% of the higher demonstration period costs could be explained by regression-to-the-mean or imbalances in beneficiary characteristics. Fully 70% would appear to be attributable to the more costly intervention. Where imbalances were large in relative terms, they were small percentagewise (e.g., 1% versus 3% of the population on Medicaid), had minor effects on cost *growth* (as distinct from initial cost *levels*), or were in favor of the intervention.

Conclusion

Based on extensive qualitative and quantitative analysis of performance, we find that

TST had limited success in improving key processes of care and no success in improving beneficiary satisfaction, self-management, or functional status and reducing hospital admissions, readmissions, emergency room visits, or mortality. The one process improvement success was done so at substantial cost to the Medicare program in the form of monthly management fees ($7.9 million) with no demonstrable savings in program outlays on health services. With only one statistical success, in the rate of influenza vaccination, the cost per successful improvement is $7.9 million or $2 million per significant percentage point improvement in a population of 5,000 FFS beneficiaries, or $400 per intervention beneficiary.

Despite a limited gain in one process-of-care measure, the lack of program savings to offset monthly management fees cannot justify the TST model for chronically ill Medicare fee-for-service beneficiaries on cost effectiveness grounds.

What might explain the lack of success in TST's demonstration?

Ineffective Targeting. 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). TST selected two large geographic areas in Texas that were medically underserved and where Texas Tech University Health Sciences Center (TTUHSC) and Texas Tech Physician Associates (TTPA) provided health care to a large proportion of the population in this area that had limited primary care access and high emergency room usage. Emergency departments are a significant source for hospital admissions in the target areas and therefore make important contributions to hospital profit margins. Thus, the TST program focused on (1) reducing emergency room utilization and (2) providing hospital discharge planning support to reduce readmissions.

To implement their care management strategy, each participant program was assigned to a care team composed of nurses who provided care management support via telephone from Dallas, Texas, and nurse care managers located in Lubbock and Amarillo, who conducted on-site visits with beneficiaries at their homes, physician offices, or in the hospital, as needed. When TST learned that one of its participants was admitted to the hospital, it reassigned this individual to its high-risk intervention, and when appropriate, a care manager visited the beneficiary in the hospital to determine the cause of the hospitalization and identify any new health or social issues to be addressed. Not surprisingly, TST adopted a strategy of targeting beneficiaries at greatest risk of a hospitalization and higher costs. Their targeting strategy was unsuccessful—and costly.

Using the total number of contacts with their intervention population reported by TST and $7.9 million in management fees, we calculate that each contact cost was roughly $332, or over 5 times the national average payment for a face-to-face office visit with an established patient with moderate complexity under the Medicare Fee Schedule.

TST's lack of success is not surprising in light of the substantial regression-to-the-mean behavior that naturally occurs with the elderly chronically ill. Armed with data on beneficiary

disease, utilization, and cost profiles in the base period, health coaches were instructed to focus 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. Targeting this group 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 program was unable to predict future complications with any precision for those with initially stable, less costly, conditions. Lacking direct access to patients' medical records, the health coaches often began working with beneficiaries with incomplete information. Further, the health coaches 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 TST was unable to successfully improve patient self-management.

Because real time information on health status was not available on beneficiaries as their health declined, TST care managers focused on identifying beneficiaries at the time of discharge from the hospital. At that point, major inpatient costs had already been incurred along with a pre-ordained stream of post-acute care institutional and physician services. Unfortunately, even then, we find no evidence of success in reducing readmissions. It is not clear that any *ex post* success after discharge would have been sufficient, by itself, to save Medicare enough to justify the management fees negotiated by TST.

"Remote" Nurse-Patient Communication. 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 TST was ineffectual has to do with the limitations of remote nurse call centers in managing beneficiary utilization of health services. By complementing, not substituting, for the primary care physician, health coaches worked to change beneficiaries' lifestyles and encouraging them to take their medications, to track key health indicators such as blood sugar levels, and to follow physician orders. While all clinicians agree these are valuable adjuncts to successful self-care management, having nurses act as health coaches, listen on telephones, and make suggestions is not the same as directly determining whether a patient is admitted to a hospital and having a personal relationship cultivated over several years. Nor can call-center nurses have much effect on service intensity once patients are admitted for care. Nurses on telephones cannot admit patients, tell them to fill a prescription right away, or instruct them to "start exercising now or risk a heart attack in the next few months." Moreover, communicating by telephone with elderly and disabled patients is complicated by the relatively high frequency of cognitive impairments.

Latent Health Needs. A final possible reason why TST may have not been successful in reducing acute care utilization and Medicare program costs is the selection of medically underserved geographic areas. It may be reasonable to assume that once the health coaches conducted health assessments, they uncovered substantial unmet health care needs. By facilitating the patient-physician relationship, the health coaches may have actually increased health care utilization.

In light of the results of other CMS disease management demonstrations that relied to varying degrees on this model of care management, it is unlikely that simply managing elderly patients “at a distance” via the telephone or an occasional in-person visit will produce the kinds of savings for which policymakers had hoped. To the extent that detailed health assessments identify significant unmet health care needs, costs are likely to increase not decrease. It also suggests that to be effective at all an intervention may require intensive personal clinical attention, which costs money. A win-win scenario, in which better quality is achieved at lower overall cost, is not likely.

REFERENCES

Actuarial Research Report (ARC): Third Monitoring Report for Texas Tech, Jan. 11, 2008.

Anderson, G.: Chronic Conditions: Making the Case for Ongoing Care. Partnership for Solutions, Johns Hopkins University and the Robert Wood Johnson Foundation. Baltimore. Johns Hopkins University, Dec. 2002.

Brody, E.R., and Bernard, S.L.: Evaluation of the Medicare care management for high cost beneficiaries demonstration: Site visit report on Texas Senior Trails. Prepared for the Centers for Medicare & Medicaid Services, Feb. 2007.

Centers for Disease Control and Prevention Diabetes Cost-Effectiveness Group: Cost-effectiveness of intensive glycemic control, intensified hypertension control, and serum cholesterol level reduction for type 2 diabetes. JAMA 287:2542-2551, 2002.

Centers for Medicare & Medicaid Services (CMS): Medicare to award contracts for demonstration projects to improve care for beneficiaries with high medical costs. Press release from CMS Office of Public Affairs, Jul. 1, 2005. Available at http://www.cms.hhs.gov/apps/media/press_releases.asp

Coleman, E.A., Parry, C., Chalmers, S., and Min, S.: The care transitions intervention. Results of a randomized controlled trial. Arch. Int. Med. 166:1822–1828, 2006.

Congressional Budget Office (CBO): An Analysis of the Literature on Disease Management Programs. Washington. U.S. Congress, Oct. 2004.

Congressional Budget Office (CBO): High-Cost Medicare Beneficiaries. Washington. U.S. Congress, May 2005.

Glasgow, R.E., Orleans, C.T., and Wagner, E.H.: Does the chronic care model serve also as a template for improving prevention? Milbank Q. 79(4):579-612, 2001.

Greene, W.: 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.

Solucia Research Company: Texas Senior Trails Medicare High Cost Demonstration Program: A Review of Population Influences. Unpublished report submitted to CMS on behalf of Texas Senior Trails, Oct. 1, 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.