

October 2008

Evaluation of Phase I of the Medicare Health Support Pilot Program Under Traditional Fee-for-Service Medicare: 18-Month Interim Analysis

Report to Congress

Prepared for

Mary Kapp, M.Phil.
Centers for Medicare & Medicaid Services
7500 Security Boulevard
Baltimore, MD 21244-1850

Prepared by

Nancy McCall, Sc.D.
Jerry Cromwell, Ph.D.
Carol Urato, M.S.
Donna Rabiner, Ph.D.
RTI International
701 13th Street, NW, Suite 750
Washington, DC 20005-3967

CMS Contract No. 500-00-0022
RTI Project Number 0207960.002

EVALUATION OF PHASE I OF THE MEDICARE HEALTH SUPPORT PILOT PROGRAM
UNDER TRADITIONAL FEE-FOR-SERVICE MEDICARE:
18-MONTH INTERIM ANALYSIS

by Nancy McCall, Sc.D.
Jerry Cromwell, Ph.D.
Carol Urato, M.S.
Donna Rabiner, Ph.D.

Federal Project Officer: Mary Kapp, M.Phil.

RTI International

CMS Contract No. 500-00-0022

October 2008

This project was funded by the Centers for Medicare & Medicaid Services under contract no. 500-00-0022. 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.

ACKNOWLEDGEMENTS

This second Report to Congress is a synthesis of findings from a more exhaustive set of analyses conducted for CMS under the contract titled “Evaluation of Phase I of Voluntary Chronic Care Improvement Pilot Program under Traditional Fee-for-Service Medicare.” Authors of related reports include Nancy McCall, Sc.D.; Jerry Cromwell, Ph.D.; Carol Urato, M.S.; Galina Khatutsky, M.S.; Shulamit Bernard, Ph.D.; Kevin Smith, M.A.; Donna Rabiner, Ph.D.; Joseph Burton, M.S.; and Erica Brody, M.S. Ann Larsen of RTI International made an invaluable contribution in the conduct of this second set of analyses through expert computer programming. Edward Drozd, Ph.D., contributed expert statistical consultation. Janet Mitchell, Ph.D., provided expert scientific review. Kate Bare provided excellent copy editing assistance and Norma DiVito provided superb documentation preparation assistance.

We would like to thank Mary Kapp, M. Phil., our CMS Project Officer, for the guidance she provided in the conduct of our second set of analyses. We would also like to thank her for careful and thoughtful review of our second annual report. She has made an invaluable contribution to this report.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
CHAPTER 1 INTRODUCTION	11
1.1 Overview of the Phase I MHS Program Design and Pilot Launch.....	13
1.2 MHS Phase I Pilot Requested Early Terminations	14
CHAPTER 2 MEDICARE HEALTH SUPPORT PROGRAM DESIGN FEATURES AND ONGOING IMPLEMENTATION EXPERIENCE	16
2.1 Overview of the Evolution of the MHS Pilot Programs	16
2.2 Changes in Outreach	17
2.3 Changes in MHS Elements	18
2.4 Implementation Challenges Reported by the MHSOs	22
2.5 Reasons for Requested Early Termination from the MHS Phase I Pilot.....	24
CHAPTER 3 PARTICIPATION RATES IN THE MEDICARE HEALTH SUPPORT (MHS) PROGRAMS AND LEVEL OF INTERVENTION.....	26
3.1 Introduction and Methods	26
3.2 Participation Rates During the First 18-months of the MHS Pilot Programs for the Original Populations and the first 6-months for the Refresh Populations	28
3.3 Original Populations' Level of Intervention in the MHS Programs	32
3.4 Characteristics of Participants and Never Participants	35
CHAPTER 4 MEDICARE HEALTH SUPPORT BENEFICIARY SURVEY OF SATISFACTION, CARE EXPERIENCE, SELF-MANAGEMENT BEHAVIORS, AND SELF-REPORTD PHYSICAL AND MENTAL HEALTH	37
4.1 Introduction.....	37
4.2 Survey Domains.....	38
4.3 Survey Design and Implementation.....	39
4.4 Analytic Methods.....	40
4.5 Findings	41
4.6 Conclusion	44
CHAPTER 5 INTERIM FINDINGS ON QUALITY OF CARE IMPROVEMENT AND HEALTH OUTCOMES	45
5.1 Methodology	45
5.1.1 Quality of Care Measures	45
5.1.2 Health Outcomes.....	47
5.2 Clinical Quality and Health Outcomes Results for the Original Populations.....	48
5.2.1 Quality of Care.....	48
5.2.2 Health Outcomes.....	50
5.3 Health Outcomes for the Refresh Populations.....	52
5.4 Conclusion	55

CHAPTER 6 INTERIM FINANCIAL OUTCOMES	57
6.1 MHS Pilot Payment Arrangements.....	57
6.2 Financial Analysis Data and Methods	58
6.3 Results for the Original Population.....	61
6.4 Monthly Fee Budget Neutrality	67
6.5 Results for the Refresh Populations.....	68
6.6 Summary of Findings.....	69
 CHAPTER 7 KEY FINDINGS BASED ON 18 MONTHS OF PHASE I OF THE MEDICARE HEALTH SUPPORT PILOT	 72
 REFERENCES	 82

LIST OF TABLES

Table 1-1 Medicare Health Support Organizations (MHSO)	15
Table 3-1 Participation status of the Medicare Health Support Organizations’ original intervention populations during the first 18 months of the Medicare Health Support pilot.....	29
Table 3-2 Participation status of the Medicare Health Support Organizations’ refresh intervention populations during the first 6 months	30
Table 3-3 Characteristics of the Original and Refresh Medicare Health Support Intervention Populations	31
Table 3-4 Percent distribution of original population participants across number of months of telephonic support during Months 7 – 18 of the Medicare Health Support Pilot	32
Table 3-5 Percent distribution of fully participating Medicare Health Support original populations’ intervention beneficiaries by risk stratum and average number of months of telephonic contacts during months 7-18 of Phase I Pilot by risk stratum.....	34
Table 3-6 Differences in the original populations’ demographic characteristics, health status, utilization, and costs at baseline between beneficiaries who agreed to participate within the first 18-month Medicare Health Support Phase I pilot period and those who never agreed to participate	36
Table 4-1 Completed Medicare Health Support original populations’ follow-up surveys and response rates by Medicare Health Support Organization	39
Table 4-2 Summary of Medicare Health Support original populations’ survey of satisfaction, self-management, and functioning.....	42
Table 5-1 Rate of receipt of guideline concordant care measures per 100 Medicare Health Support intervention beneficiaries during the year prior to launch of the MHS pilot and rate of change during months 7-18 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Original Populations.....	49
Table 5-2 Rate of acute care utilization per 1,000 Medicare Health Support intervention beneficiaries during the year prior to launch of the MHS pilot and rate of change during months 7-18 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Original Populations	51

Table 5-3	Mortality rates during the first 18 months of Medicare Health Support Phase I pilot; original populations	52
Table 5-4	Rate of acute care utilization per 1,000 Medicare Health Support intervention beneficiaries during a 6-month period prior to launch of the MHS pilot and rate of change during months 1-6 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Refresh Populations	54
Table 5-5	Comparison of refresh intervention and comparison group mortality rates during the first 6-month Medicare Health Support Phase I pilot period	55
Table 6-1	Intervention/Comparison differences in Medicare Health Support per-beneficiary-per-month (PBPM) growth rates between 18-month Phase I pilot period and 12-month base year period, by Medicare Health Support Organization(MHSO) for original populations	62
Table 6-2	Differences in Medicare Health Support per beneficiary per month (PBPM) growth rates, adjusted for beneficiary characteristics and regression-to-the-mean effects.....	65
Table 6-3	Comparison of PBPM differences between participants, never participants, & comparison beneficiaries in the 12-month base year and 18-month pilot period, by Medicare Health Support Organization (MHSO), original populations.....	66
Table 6-4	Success in achieving monthly fee budget neutrality through first 18 pilot months, by Medicare Health Support Organization (MHSO), original populations.....	67
Table 6-5	Per-beneficiary-per-month (PBPM) Medicare payment differences between Medicare Health Support first 6-month pilot period and a comparable 6-month prior period, by Medicare Health Support Organization (MHSO), refresh populations.....	69

LIST OF FIGURES

Figure 6-1	Intervention/Comparison differences in Medicare Health Support per-beneficiary-per-month (PBPM) growth rates ¹ between 18-month Phase I pilot period and 12-month base year period, by Medicare Health Support Organization for original populations.....	63
------------	--	----

EXECUTIVE SUMMARY

The purpose of this Report to Congress is to report the results of RTI International's 18-month evaluation of eight Medicare Health Support (MHS) pilot programs implemented under Phase I of the "Voluntary Chronic Care Improvement Program (CCIP) Under Traditional Fee-for-Service (FFS) Medicare" pilot as authorized by Section 721 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (Pub. L. 108-173) (codified as Section 1807 of the Social Security Act, hereafter "the Act"). Section 721 requires the Secretary of Health and Human Services to provide for the phased-in development, testing, evaluation, and implementation of chronic care improvement programs. Prior to program implementation, the name of the initiative was changed from Chronic Care Improvement Program to Medicare Health Support, which we refer to as MHS hereafter.

Subsection (b)(5) of the legislation states that the evaluation shall include an assessment of the following factors for each program:

- quality improvement measures,
- beneficiary and provider satisfaction,
- health outcomes, and
- financial outcomes.

Section 1807(c)(1) of the Act states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I.

The legislation also mandated four Reports to Congress, the first of which was to be provided not later than 2 years after the date of implementation. The first Report to Congress was submitted to Congress June 2007. A second Report to Congress was mandated not later than 3 years and 6 months after the date of implementation and is to contain an update on the scope of implementation of the programs, the design of the programs, and findings with respect to the following measures: quality improvement, such as adherence to evidence-based guidelines and re-hospitalization rates; beneficiary and provider satisfaction; health outcomes; and financial outcomes. This report serves as the second Report to Congress and contains the first 18 months of MHS experience, or the half-way point of the Phase I pilot.

To meet the congressional timeline, the first Report to Congress presented evaluation findings based on the first 6 months of MHS program operations. This report presents evaluation findings based on the first 18 months of the MHS program operations for the original populations and the first 6 months of the refresh populations, and was conducted to provide for an assessment of the degree to which the MHS organizations' (MHSOs) programs or component(s) of programs were meeting the statutory requirements necessary for expansion of Phase II. We include the experiences of both the original and refresh populations in this Report

to Congress to capture the impact of the evolution of the MHS programs on acute care utilization and savings, thus capturing the dynamic nature of the implemented programs. The third Report to Congress will examine the full 3-year experience of the original populations and 2-year experience of the refresh populations.

I. Scope of Implementation of the Medicare Health Support (MHS) Programs

After a competitive solicitation, the Centers for Medicare & Medicaid Services (CMS) selected nine chronic care improvement programs for award. Eight MHSOs launched their programs between August 1, 2005 and January 16, 2006. A ninth program decided not to go forward with finalizing its agreement. Programs are distributed throughout the United States and serve a variety of populations. Several programs serve urban and suburban populations, while others target metropolitan and rural communities. Among the populations served, there are significant minority populations of African American, Native American, and Hispanic beneficiaries.

CMS prospectively identified eligible beneficiaries from each area and randomly assigned 30,000 into intervention and comparison groups in a ratio of 2:1 under an intent-to-treat (ITT) evaluation model¹. Our analyses revealed that the block (stratified) randomization procedure effectively created equivalent intervention and comparison populations at the time of randomization for each of the eight MHSOs for the variables that were used in randomization as well as similar demographic, disease, and economic burden profiles. However, our analyses revealed that an unexpected pattern in per beneficiary per month (PBPM) Medicare payment differences between intervention and comparison groups emerged between the time of randomization and the start of the MHS pilots. As a result, CMS modified its financial reconciliation protocol to allow for an actuarial adjustment in the intervention PBPM for any difference from the comparison group in the 12 months just prior to the start date of each MHSO.

One year after launch of each pilot program, CMS offered all MHSOs the option of supplementing their intervention and comparison populations with additional beneficiaries. CIGNA Health Support (CHS) did not opt for a supplemental population. A total of approximately 47,000 beneficiaries were distributed across 7 of the 8 the MHSOs' intervention and comparison populations through a randomization process similar to that used to assign the original populations with one noted exception related to the targeted clinical condition(s). The MHSOs requested and CMS agreed to first assign beneficiaries with heart failure (HF) and diabetes, then heart failure-only, and finally diabetes-only. The MHSOs believed at that time that they would have greater financial success with a population more heavily weighted with heart failure rather than diabetes. For 4 MHSOs, the randomization process and eligible populations in their geographic areas resulted in their supplemental populations containing only beneficiaries with heart failure (HF). CMS, the MHSOs, and others associated with the MHS pilot refer to the supplemental populations as the "refresh populations."

Through 2007, three organizations requested early termination of their programs. LifeMasters Supported Self Care ended their MHS program December 31, 2006, McKesson

¹ Health Dialog had a 2.5 to 1 ratio of intervention to comparison beneficiaries.

Health Solutions, LLC, ended their MHS program May 31, 2007, and CIGNA Health Support ended their MHS program January 14, 2008. Phase I will cease for the remaining MHSOs between July 31, 2008 and August 31, 2008. After the CMS announcement in January 2008 that Phase I would cease at the end of the three year pilot period, XLHealth and Green Ribbon Health (GRH) requested early termination.

The MHSOs received monthly management fees for the full original and refresh populations for the first six months of engagement of each of the populations. After the initial 6-month outreach period, the MHSOs accrue management fees for only those beneficiaries who verbally consent to participate and only during periods of participation. Participation continues until a beneficiary becomes ineligible for the MHS program or opts out of services provided by the MHSO. Participants may drop out of the program at any time and begin participation again at any time, as long as they are eligible. However, over one-half of all MHS beneficiaries are continuous participants, meaning that once they consent to participate they participate for all of their MHS eligible days. Many of these beneficiaries are fully eligible and continuous participants meaning they meet MHS eligibility criteria for the entire first 18 months of the pilot and participate all days after consent. Never participants are individuals in the intervention group who did not consent to participate or were not reachable by the MHSO for all months in which they were eligible to participate and remain in the MHSO's intervention group "at risk" population. Beneficiaries who decline participation may be re-contacted by the MHSO after a sentinel event, such as a hospitalization or ER visit.

For the original populations, participation rates for the first 18-month period ranged from 74 to 95%. Most of the participants consented in the initial 6-month period of the pilot. Of the beneficiaries who never consented to participate, the percent unable to be contacted ranged from 4 to 15% and the refusal rate ranged from 0.3 to 13%. Participation rates were generally lower for the refresh populations than observed during the first 6-month period for the original populations with the exception of Health Dialog, whose 6-month participation rate for the original population was 95% and 96% for their refresh population. The participation rates for the first 6 months for the remaining 6 MHSOs ranged from 34 to 80% for their refresh populations versus 65 to 83% for their original populations.

We compared baseline characteristics of the original beneficiaries who consented to participate during the first 18 months of the pilot with baseline characteristics of the refresh beneficiaries who consented during the first 6 months of the refresh period. We find that participants as a group remain different from the group of beneficiaries that never consented to participate across numerous demographic, health status, utilization, and payment characteristics reviewed. With the exception of McKesson, the proportion of participating beneficiaries with Medicaid enrollment is between 3 and about 14 percentage points lower than for never participants meaning that most MHSOs have not been as successful at recruiting Medicare/Medicaid dual enrolled beneficiaries to participate. Six of the MHSOs have lower rates of Medicare beneficiaries who are under age 65, or beneficiaries with disabilities, among their participating beneficiaries. Mean HCC risk scores calculated for the 1-year period prior to each MHSO going live range from 20 to 40% lower for participants than for never participating beneficiaries. All cause hospitalization and ER visit rates and average per beneficiary per month (PBPM) Medicare payments during the year prior to going live are all lower for participants than

for the never participants. Thus, we continue to observe participants to be a healthier and less costly subset of the whole intervention group across all MHSOs.

II. Beneficiary and Provider Satisfaction

The Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of beneficiary satisfaction. In addition, the evaluation seeks to answer a broader set of research questions related to whether the programs improved knowledge and self-management skills and led to behavioral change among participants. The evaluation includes these additional foci to better understand the factors for program success.

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 at baseline (Months 4 to 6 of the intervention period) and 12 months later. No further surveying of the original populations or surveying of the refresh populations will be conducted during RTI's evaluation. Thus, these reported results are final results with respect to beneficiary satisfaction.

One of the required conditions for expansion of the programs is improvement in satisfaction among the intervention beneficiaries relative to the comparison beneficiaries. RTI and CMS have defined satisfaction to mean that "beneficiaries were helped by their health care team to cope with their chronic conditions." For the satisfaction measure, a positive intervention effect is observed for 2 of 7 MHSOs, Health Dialog and Aetna.

We observe limited MHS intervention effects on 27 beneficiary survey measures across 7 of the 8 MHSOs². Of the 189 measures (27 x 7), 25 (13%) showed significant positive intervention effect and 4 (2%) showed negative intervention effect. The remaining 160 indicators were not statistically different between intervention and comparison beneficiaries. The focus of the pilot program interventions was largely on impacting beneficiary behavior to better manage their chronic illness. The MHSOs were most successful in helping beneficiaries to set goals to address their care needs; however there was little meaningful improvement in self-efficacy or self-care activities. To positively affect acute care utilization, one would expect to see improvement in self-care behaviors by the mid-way point of the pilot. Both the intervention and control groups within most MHSOs exhibited similar declines in physical and mental functioning between the baseline and follow-up surveys. We did not observe any consistent pattern of positive intervention effects by disease cohort of heart failure only, diabetes only, and heart failure and diabetes.

During initial site visits at each MHSO, we spoke with a small number of randomly selected community-based physicians to gauge their early assessment of their satisfaction with the MHS pilot programs. Universally, the community-based physicians felt that the programs could benefit Medicare FFS beneficiaries with chronic conditions. Not unexpectedly, their exposure had been sufficiently limited that they were unable to provide estimates of their current level of satisfaction with the programs. RTI now is examining more broadly provider exposure to

² LifeMasters terminated its participation prior to RTI's fielding of the follow-up survey. Hence, no survey results are reported.

and satisfaction with the MHS pilot programs and will report the findings in the Third Report to Congress.

III. Quality of Care and Health Outcomes Findings

One of the required conditions for expansion of the MHSO programs is improvement in quality of care and the Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of quality improvement measures and health outcomes. We have defined quality improvement for this evaluation as an increase in each intervention group's rate of receipt of claims-based process-of-care measures (e.g., serum cholesterol testing) and improvement in health outcomes as a reduction in each intervention group's rate of hospitalizations, re-admissions, and ER visits, and a reduction in mortality rates relative to their respective comparison group. We present *interim* results for an 18-month period for the original populations and *preliminary* selected 6-month results for the refresh populations.

Across 40 quality of care measures (five measures for each of the eight MHSOs), there was modest improvement in 16 (or 40%) measures for the original populations, in the range of 2 to 4 percentage points. Seven of the 8 MHSOs demonstrated at least one positive intervention effect. Healthways demonstrated a positive intervention effect across all five process-of-care measures and CHS across four of the five measures. LifeMasters improved cholesterol screening rates among beneficiaries with heart failure as well as diabetes but demonstrated no other positive intervention effects. Aetna and GRH improved the rate of cholesterol screening for beneficiaries with diabetes, and McKesson improved the rates of cholesterol screening and HbA1c testing in beneficiaries in diabetes. None of these three MHSOs demonstrated a positive intervention effect related to beneficiaries with heart failure. Health Dialog demonstrated a positive intervention effect related to cholesterol screening for beneficiaries with heart failure but did not demonstrate any intervention effects related to quality of care for beneficiaries with diabetes. XLHealth did not demonstrate any positive intervention effects on quality of care. Because the process-of-care measures that we study are defined as annual rates of service, we believe that it would be inappropriate to evaluate the performance of the MHSOs using only 6 months of intervention experience for the refresh population.

Changes in rates of hospitalization (all cause, heart failure, and diabetes), all cause rates of readmission, and all cause ER visits per 1,000 MHS beneficiaries were assessed for both the intervention and comparison groups for both the original and refresh populations, in total, and by disease cohorts. For the original populations and across the 104 comparisons of acute care utilization (13 measures for each of the eight MHSOs), there were no statistically significant reductions in hospitalizations, rates of readmission, or emergency room (ER) visits between the intervention and comparison groups. For the refresh populations and across 74 comparisons, there were no statistically significant reductions in hospitalizations or ER visits between the intervention and comparison groups.

We also examined mortality rates that tend to be high among the chronically ill. We found no MHSO to have statistically significant lower mortality rates or longer time to death by mid-point in the pilot for the original or refresh populations.

IV. Cost Savings Findings

Subsection (b)(5) of the MMA legislation called for an independent evaluation of financial outcomes, or Medicare program savings. One of the three Phase I requirements for expansion to Phase II is achievement of budget neutrality with respect to monthly management fees. In the MHS pilot, each MHSO receives from CMS a negotiated monthly administrative fee per participant. Fees are also at risk for improvements in quality and beneficiary satisfaction. During the first 6 months of Phase I, the MHSOs received a monthly management fee for each beneficiary in their assigned intervention group until such time that the beneficiary became ineligible or declined to participate. Beyond the initial 6-month period, management fees are paid only for confirmed participants, and only for eligible pilot periods. MHSOs are held at risk for fees based on the performance of the full population of beneficiaries randomized to the intervention group (an intent-to-treat [ITT] model) compared with beneficiaries randomized to the comparison group.

To keep all their management fees, the MHSOs must reduce Medicare payments for their intervention groups by the dollar amount of accrued fees, i.e., achieve budget neutrality. We conducted a difference-in-difference analysis of trends in PBPMs between the year prior to and the first 18 months of the pilot for each of the 8 MHSOs. Comparing growth rates between the original population intervention and comparison groups, we find that none of the MHSOs showed statistically lower rates of growth in intervention PBPMs relative to the comparison group. Four MHSOs exhibited a trend toward higher rates of growth (LifeMasters, Healthways, Health Dialog, and McKesson) and 4 exhibited a trend toward lower rates (Aetna, CHS, GRH, and XLHealth). Required savings for statistical significance ranged from \$48 to \$73. RTI is able to detect differences in intervention and comparison group growth rates that are as small as 3.4 to 4.5% of the 18-month comparison group PBPM.

We also stratified trends in MHSO PBPMs by five non-mutually exclusive disease groups: (1) heart failure (HF) only, (2) diabetes only, (3) HF with or without diabetes, (4) diabetes with or without HF, and (5) HF and diabetes. The MHSOs developed components of their programs to focus upon disease-specific cohorts of beneficiaries (e.g., telemonitoring for beneficiaries with heart failure). Because we do not have detailed intervention data – that go beyond number of visits or telephone contacts - we wanted to indirectly evaluate whether program savings occurred for components of the MHS programs and whether program savings occurred for beneficiaries with only a single disease or multiple co-morbid conditions. No pattern was found within any of the five disease groups among the eight MHSOs that might imply successful targeting of intervention efforts. Of the 40 statistical tests we conducted of differential growth rates by the five disease groups, only one of the differences was statistically significant. Twenty-one of 40 comparisons showed intervention PBPMs trending at a slower rate but 19 PBPMs were trending at a faster rate.

Average monthly management fees during the first 18 months of the pilot ranged from \$67 to \$118³. As a proportion of comparison group PBPMs, average monthly management fees ranged from 4.7% to 9.3%. The 8 MHSOs over the first 18 months have had limited success in

³ For the remaining 18 months of the pilot, monthly management fees will range from \$0 to \$169.

covering accrued fees. In fact, Healthways, LifeMasters, Health Dialog and McKesson have diverged even further in their attempt to recover accrued fees than observed at 6-months. At best, GRH has offset 26% of their accrued fees through Medicare savings after 18 months. XLHealth has offset, or saved, 23% of its accrued fees, Aetna, 21%, and CHS, just 12%.

For the refresh populations after 6-months of their initial intervention period, none of the 7 MHSOs that accepted a refresh population experienced statistically lower PBPM growth in their intervention versus comparison group. The implication of these findings is that the remaining 5 MHSOs (CHS, LifeMasters, and McKesson elected to terminate early) need to generate savings over the pilot's last 18-month period at several times their current rate to reach budget neutrality.

V. Summary of Key Findings

Section 1807(c)(1) of the Act states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves the clinical quality of care and beneficiary satisfaction and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I. In this report, we present key findings based upon the first 18 months of MHS operations, the mid-point of Phase I. Our findings are based on the experience of approximately 240,000 chronically ill Medicare beneficiaries randomized to an intervention or a comparison group in eight geographic areas in the original populations and approximately 47,000 beneficiaries in the refresh populations. To date, this is the largest randomized experiment in population-based case management ever conducted and was designed to test the scalability of such programs in Medicare FFS. Five key findings on participation, beneficiary satisfaction, clinical quality and health outcomes, and financial success have important policy implications for the Centers for Medicare & Medicaid Services (CMS) and future disease management or care coordination efforts among Medicare fee-for-service (FFS) beneficiaries.

Key Finding #1: Several vulnerable sub-populations of Medicare FFS beneficiaries were less likely to agree to participate in the MHS pilot program.

We find that the participant populations continue to be healthier, less costly, and lower users of acute care services than beneficiaries who never participated during any of the first 18 months. With the exception of McKesson, the proportion of participating beneficiaries with Medicaid enrollment is between 3 and 14 percentage points lower than for never participants. Six of the MHSOs have lower rates of Medicare beneficiaries who are under age 65, or beneficiaries with disabilities, among their participating beneficiaries.

The MHS Phase I pilot was designed to be a broad population-based FFS program. If CMS desires broadly focused care management programs, these interim findings suggest alternative recruiting and outreach strategies are needed to reach the sicker and more costly beneficiaries as well as dual Medicare/Medicaid enrollees and beneficiaries with disabilities as the current MHS recruitment strategies are not reaching these populations to the degree they are reaching other FFS beneficiaries. These populations likely include a high proportion of

beneficiaries residing in nursing homes or other institutional settings. During RTI's site visits, MHSOs reported that they found locating and engaging these populations very difficult.

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

Although there was no pre-determined expected number of contacts, the MHS beneficiaries are a sick and costly group of FFS beneficiaries averaging over 1 hospitalization annually in the year prior to program launch, and the MHSOs reported significant unmet clinical and psychosocial need. The majority of fully eligible and participating beneficiaries during months 7 -18 of the pilot received between 2 and 5 months of telephonic support. Given the lack of consistent monthly or bimonthly interaction with many of the MHS participants, it is unlikely that the MHSOs will be successful at changing beneficiary behavior with respect to self-management of their chronic illness. Findings from the beneficiary survey shows there has been little meaningful improvement in self-care activities. To positively affect acute care utilization, one would expect to see improvement in self-care behaviors by the mid-way point of the pilot, and savings have proved illusive to date. Further examination is warranted in how the disease management strategies were implemented and whether there is evidence of successful selective targeting of beneficiaries for intervention contacts that are associated with positive outcomes.

Key Finding #3: There was limited effect in improving beneficiary satisfaction, care experience, self-management, and physical and mental health functioning during the first 18-months of the Phase I pilot.

The Medicare Health Support authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction, and achieves targets for savings, the program (or its components) may be expanded to additional geographic areas. Only 2 of the MHSOs, Health Dialog and Aetna, improved beneficiary satisfaction as measured by beneficiary assessment that their health care team helped them cope with their chronic condition, our principal measure of satisfaction. None of the seven MHSOs included in the beneficiary survey analyses demonstrated consistent positive intervention effects across the four domains of satisfaction, care experience, self-management activities, and physical and mental health functioning. The focus of the pilot program interventions was largely on impacting beneficiary behavior to better manage their chronic illness. Yet these results show little evidence of changes in self-efficacy or self-care. We did not observe any consistent pattern of positive intervention effects by disease cohort of heart failure only, diabetes only, and heart failure and diabetes.

Key Finding #4: Seven of the MHSOs had a positive intervention effect on one or more process-of-care measures but no positive intervention effect on reduction in acute care utilization or mortality.

A second required condition for expansion of the MHS programs is improvement in quality of care and the Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of quality improvement measures and health outcomes. Across 40 quality of care measures (five measures for each of the eight MHSOs), there was modest improvement in 16 (or 40%) of the measures for the original populations. Seven of the 8 MHSOs

demonstrated at least one positive intervention effect. However, rates of improvement in the quality of care measures were relatively modest; 2 to 4 percentage points. We did not examine clinical quality for the refresh populations.

For both the original and refresh populations, none of the 8 MHSOs demonstrated positive intervention effects related to health outcomes. We observe no reduction in mortality rates or time to death during the first 18 months of the pilot for the original populations and the first 6 months of engagement of the refresh populations.

Key Finding #5: Fees accrued to date far exceed savings produced.

A third required condition for expansion of the MHS programs is Medicare program financial savings defined as budget neutrality with respect to the MHSOs accrued management fees. RTI's findings through 18 months, or halfway through Phase I of the pilot, show that none of the 8 MHSOs achieved gross savings rates that were statistically different from zero for their original and refresh populations. Further, no evidence of savings was found among only those beneficiaries who agreed to participate in the intervention. Evidence was found of engagement of less costly, healthier, beneficiaries into the participant pool. Intervention beneficiaries who never consented to participate were much more expensive in the base year and experienced higher rates of cost increases than the participant group. Participant cost increases, alone, were no different than for the entire comparison group. These findings were insensitive to MHSO differences in participation rates.

Savings one-half way through the Phase I pilot period have offset 12-26% of estimated accrued fees for 4 of 8 MHSOs⁴. The remaining 4 MHSOs show no evidence of savings. The 4 MHSOs with modest savings would have to increase their rate of savings by roughly 3-to-7-fold while the second group of 4 "non-savers" has a significantly higher hurdle. This seems unlikely given performance over the first 18 months.

Conclusion

The Medicare Health Support authorizing legislation states that if the results of an independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction, and achieves targets for savings, the program (or its components) may be expanded to additional geographic areas. None of the MHS pilot programs at the mid-point of the pilot have yet to meet the three statutory requirements to improve clinical quality of care and beneficiary satisfaction and achieve budget neutrality with respect to their fees.

Among their original populations, seven of the MHS pilot programs modestly improved rates of receipt of at least one aspect of guideline-concordant care but none reduced rates of acute care hospitalization, readmission, or ER visits. None reduced the rate of mortality. Two of the

⁴ RTI did not factor in a reduction in the negotiated fee requested by one MHSO after the 18-month period. The MHSO requested that its fee be reduced to \$0. Incorporating such a reduction would have resulted in this MHSO appearing more successful on the budget neutrality criterion than they actually were at the mid-point of the pilot.

MHSOs improved beneficiary satisfaction. None of the MHSOs achieved budget neutrality within the first 18 months of program operations within their original populations. The two MHSOs that improved beneficiary satisfaction each had a positive modest intervention effect on one of five process-of-care measures. Neither of these two MHSOs lowered acute care hospitalizations or ER visits nor did they achieve budget neutrality. Another MHSO had no intervention effect on beneficiary satisfaction, quality of care, or budget neutrality. The other five MHSOs modestly improved rates of receipt of guideline-concordant care but none lowered acute care hospitalizations, readmissions, or ER visits or achieved budget neutrality.

Among their refresh populations, none of the seven MHSOs that accepted a refresh population improved health outcomes. We observe no statistically significant reductions in rates of hospitalization or ER visits. Nor do we observe reduction in mortality rates during the first 6 months of engagement of the refresh populations. None of the seven MHSOs experienced statistically lower per beneficiary per month (PBPM) growth in their intervention versus comparison group payments to achieve budget neutrality.

Given the limited gains regarding quality of care and savings to offset accrued monthly management fees, it will be difficult to justify these private disease management models on cost effectiveness grounds—at least for chronically ill Medicare FFS beneficiaries. With 16 statistical successes out of 40 possible improvements in evidence-based process-of-care measures, the cost per successful improvement is approximately \$16 million, based on CMS' estimate of \$250 million in accrued MHS fees through 18 months for the 160,000 original population intervention beneficiaries. The cost would be \$6.4 million per percentage point improvement. Accounting for the 25 (of 189) improved indicators of beneficiary satisfaction and self-management does not materially alter our conclusion. Nor is there any obvious correlation between MHSOs that partially offset fees and their quality of care improvements.

The findings presented in this second Report to Congress are based upon the first 18 months of MHS operations for the original populations, the mid-point of Phase I, and 6 months of MHS operations for the refresh populations. The third Report to Congress will contain the evaluation of the full 3-year Phase I implementation experience and will report on provider satisfaction with the MHS Phase I pilot and the MHSOs' effect on quality of care and health outcomes and Medicare program savings.

CHAPTER 1 INTRODUCTION

The purpose of this Report to Congress is to report the results of RTI International’s 18-month evaluation of eight Medicare Health Support (MHS) pilot programs implemented under Phase I of the “Voluntary Chronic Care Improvement Program (CCIP) Under Traditional Fee-for-Service (FFS) Medicare” pilot as authorized by Section 721 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (Pub. L. 108–173) (codified as Section 1807 of the Social Security Act, hereafter “the Act”). Section 721 requires the Secretary of Health and Human Services to provide for the phased-in development, testing, evaluation, and implementation of chronic care improvement programs. Prior to program implementation, the name of the initiative was changed from Chronic Care Improvement Program to Medicare Health Support, which we refer to as MHS hereafter.

The principal objectives of this initiative are to test a pay-for-performance contracting model and MHS intervention strategies that may be adapted nationally to improve clinical quality, increase beneficiary and provider satisfaction, and achieve Medicare program savings for chronically ill Medicare fee-for-service (FFS) beneficiaries with targeted conditions of heart failure (HF) and/or diabetes. In addition, this initiative provides the opportunity to evaluate the success of the “fee at risk” contracting model, a new pay-for-performance model for the Centers for Medicare & Medicaid Services (CMS). This model provides MHS organizations (MHSOs) with flexibility in their operations and strong incentives to keep evolving toward outreach and intervention strategies that are most effective in improving population outcomes.

The overall design of the MHS pilot follows an intent-to-treat (ITT) model and the MHSOs are held at risk for their monthly management fees based on the performance of the full population of eligible beneficiaries randomized to the intervention group and as compared to all eligible beneficiaries randomized to the comparison group. Beneficiary participation in the MHS programs is voluntary and does not change the scope, duration, or amount of Medicare FFS benefits currently 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 MHS program services.

Each MHSO receives from CMS a monthly administrative fee per participant, contingent on improvements in clinical quality, beneficiary satisfaction, and intervention group savings in Medicare payments being equal to fees paid to the MHSO, or budget neutrality. MHSOs are held at risk for accrued fees based on the performance of the full population of eligible beneficiaries randomized to the intervention group (an ITT model) compared with the comparison group. CMS has developed the MHS initiative with considerable administrative risk as an incentive to reach assigned beneficiaries and their providers and to improve care management. To retain all of its accrued fees, an MHSO must reduce average monthly payments by the proportion of the comparison population payments that the fee comprises. The MHSOs must also meet clinical quality and satisfaction improvement thresholds or pay back negotiated portions of their fees.

Subsection (b)(5) of the legislation states that the evaluation shall include an assessment of the following factors for each program:

- quality improvement measures,
- beneficiary and provider satisfaction,
- health outcomes, and
- financial outcomes.

Section 1807(c)(1) of the Act legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I.

The legislation also mandated four Reports to Congress, the first of which was to be provided not later than 2 years after the date of implementation. The first Report to Congress was submitted to Congress June 2007. A second Report to Congress was mandated not later than 3 years and 6 months after the date of implementation and is to contain an update on the scope of implementation of the programs, the design of the programs, and findings with respect to the following measures: quality improvement, such as adherence to evidence-based guidelines and re-hospitalization rates; beneficiary and provider satisfaction; health outcomes; and financial outcomes. This report serves as the second Report to Congress.

To meet the congressional timeline, the first Report to Congress presented evaluation findings based on the first 6 months of MHS program operations. This report presents evaluation findings based on the first 18 months of the MHS program operations, and reports on analyses that were conducted to provide for an assessment of the degree to which the MHSO programs or component(s) of programs were meeting the statutory requirements necessary for development of Phase II.

RTI International was hired by CMS to be the independent evaluator of the Phase I pilot and has conducted multiple analyses as part of its ongoing evaluation. RTI has made two rounds of site visits to each of the MHSOs. The first set of site visits were conducted approximately 4 months after initial pilot program startup—between November 2005 and May 2006—and a second set between February and August 2007. The first site visit focused on learning about MHS program startup, examining the elements of the MHS programs, determining the nature of the MHSOs' relationship with physicians in each community, learning about ways the MHSOs 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 program implementation (including engagement of the refresh population), program monitoring/outcomes to date, and implementation experience/lessons learned to date. Telephone interviews have occurred every 6 months between site visits and will continue until the end of Phase I. The purpose of the phone calls are to keep apprised of changes in program implementation.

During initial site visits at each MHSO, we spoke with two to four randomly selected community-based physicians to gauge their early assessment of satisfaction with the MHS pilot programs. Universally, the community-based physicians felt that the programs could benefit

Medicare FFS beneficiaries with chronic conditions. Not unexpectedly, their exposure had been sufficiently limited that they were unable to provide estimates of their current level of satisfaction with the programs. RTI now is examining more broadly provider exposure to and satisfaction with the MHS pilot programs and will report the findings in a later Report to Congress.

RTI conducted a beneficiary survey at approximately 6 and 18 months after the start of each pilot program's launch to assess the programs' effects on changes in beneficiary satisfaction with their care experience, self-management behaviors, and self-reported physical and mental health functioning. RTI has also conducted a set of Medicare claims and MHSO participation and intervention data analyses of the first 18 months of the Phase I pilot intervention.

1.1 Overview of the Phase I MHS Program Design and Pilot Launch

After a competitive solicitation, the Centers for Medicare & Medicaid Services (CMS) selected nine chronic care improvement programs for award. Eight MHSOs launched their programs between August 1, 2005 and January 16, 2006. A ninth program decided not to go forward with finalizing its agreement. Programs are distributed throughout the United States and serve a variety of populations. Several programs serve urban and suburban populations, while others target metropolitan and rural communities. Among the populations served, there are significant minority populations of African American, Native American, and Hispanic beneficiaries.

CMS prospectively identified eligible beneficiaries from each area and randomly assigned 30,000 into intervention and comparison groups in a ratio of 2:1 under an intent-to-treat (ITT) evaluation model⁵. Randomization occurred on May 11, 2005. Our earlier analyses contained in the first Report to Congress (McCall *et al.*, June 2007), revealed that the block (stratified) randomization procedure effectively created equivalent intervention and comparison populations at the time of randomization for each of the eight MHSOs for the variables that were used in randomization (i.e., three Hierarchical Condition Categories [HCC] risk score ranges, Medicaid enrollment, and proportion with heart failure [HF]). We also confirmed that the randomization procedure produced similar demographic, disease, and economic burden profiles between the intervention and comparison groups at the time of randomization. However, our analyses revealed that an unexpected pattern in per beneficiary per month (PBPM) Medicare payment differences between intervention and comparison groups emerged between the time of randomization and the start of the MHS pilots. As a result, CMS modified its financial reconciliation protocol to allow for an actuarial adjustment in the intervention PBPM for any difference from the comparison group in the 12 months just prior to the start date of each MHSO.

The MHS pilot targets beneficiaries with the threshold condition(s) of heart failure and/or diabetes from among the diagnoses listed on Medicare claims. However, the level of comorbidity and rates of acute care utilization during the year prior to randomization is very high among MHS beneficiaries. Given that the MHSOs are at financial risk for all Medicare program

⁵ Health Dialog had a 2.5 to 1 ratio of intervention to comparison beneficiaries.

costs, and not just those related to the threshold conditions, MHSOs implemented holistic approaches to care management.

One year after launch of each pilot program, CMS offered all MHSOs the option of supplementing their intervention and comparison populations with additional beneficiaries. This permitted MHSOs to offset the impact of attrition⁶ and achieve roughly 20,000 beneficiaries at the start of Year 2. Groups of 2,000–5,000 beneficiaries, depending on the MHSO, were added to the intervention groups through a randomization process similar to that used to assign the original populations with one noted exception related to the targeted clinical condition(s). The MHSOs requested and CMS agreed to first assign beneficiaries with heart failure (HF) and diabetes, then heart failure-only, and finally diabetes-only. The MHSOs believed at that time that they would have greater financial success with a population more heavily weighted with heart failure rather than diabetes. For four MHSOs, the randomization process and eligible populations in their geographic areas resulted in their supplemental populations containing only beneficiaries with HF. CIGNA Health Support (CHS) did not opt for a supplemental population. CMS, the MHSOs, and others associated with the MHS pilot refer to the supplemental populations as the “refresh populations.”

The MHSOs received monthly management fees for the full original and refresh populations for the first six months of engagement of each of the populations. After the initial 6-month outreach period, the MHSOs accrue management fees for only those beneficiaries who verbally consent to participate and only during periods of participation. Participation continues until a beneficiary becomes ineligible for the MHS program or opts out of services provided by the MHSO. Participants may drop out of the program at any time and begin participation again at any time, as long as they are eligible. However, over one-half of all MHS beneficiaries are continuous participants, meaning that once they consent to participate they participate for all of their MHS eligible days. Many of these beneficiaries are fully eligible and continuous participants meaning they meet MHS eligibility criteria for the entire first 18 months of the pilot and participate all days after consent. Never participants are individuals in the intervention group who did not consent to participate or were not reachable by the MHSO for all months in which they were eligible to participate and remain in the MHSO’s intervention group “at risk” population. Beneficiaries who decline participation may be re-contacted by the MHSO after a sentinel event, such as a hospitalization or ER visit.

1.2 MHS Phase I Pilot Requested Early Terminations

During the second year of operations, three organizations requested early termination of their programs. LifeMasters Supported Self Care ended their MHS program December 31, 2006, McKesson Health Solutions, LLC, ended their MHS program May 31, 2007, and CIGNA Health Support ended their MHS program January 14, 2008. Their primary stated reason for early termination was concern that the 5% savings requirement in addition to achieving savings to

⁶ Attrition was primarily due to the high mortality rate within this population. Beneficiaries also become ineligible for participation in the MHS program if they join a Medicare Advantage plan, enter hospice, develop ESRD, lose Medicare Part A or B eligibility, or Medicare becomes a secondary payer. Several state initiatives, Pennsylvania in particular, focusing upon Medicare and Medicaid dual enrollees also rendered some MHS beneficiaries ineligible for continuing participation.

cover their accrued fees was too ambitious a goal. CMS had not yet made the decision to remove the 5% savings criterion when they requested early termination⁷. Phase I will cease for the remaining MHSOs between July 31, 2008 and August 31, 2008. After the CMS announcement in January 2008 that Phase I would cease at the end of the three year pilot period, XLHealth and Green Ribbon Health requested early termination. **Table 1-1** displays the eight Phase I MHSOs, their geographic MHS service areas, program launch dates, early Phase I program termination dates for the MHSOs that requested early termination, and the original 3-year Phase I pilot program termination dates for the MHSOs that have not requested early termination.

Table 1-1
Medicare Health Support Organizations (MHSO)

MHSO	Target geography	MHSO launch date	MHSO revised termination date	MHSO original termination date
Healthways	Maryland and District of Columbia	8/1/2005	n/a	7/31/2008
LifeMasters Supported SelfCare	Oklahoma	8/1/2005	12/31/2006	7/31/2008
Health Dialog Services Corporation	Pennsylvania (western region)	8/15/2005	n/a	8/14/2008
McKesson Health Solutions, LLC	Mississippi	8/22/2005	5/31/2007	8/21/2008
Aetna Life Insurance Company	Chicago, IL (surrounding area)	9/1/2005	n/a	8/31/2008
CIGNA Health Support	Georgia (northern region)	9/12/2005	1/14/2008	9/11/2008
Green Ribbon Health	Florida (west-central region)	11/1/2005	8/15/2008	10/31/2008
XLHealth Corporation	Tennessee (selected counties)	1/16/2006	7/31/2008	12/31/2008

⁷ Under the original contract terms, there was a 5% savings requirement net of monthly management fees. In December 2007, CMS rescinded the 5% requirement, thereby requiring MHSOs only to achieve fee budget neutrality.

CHAPTER 2

MEDICARE HEALTH SUPPORT PROGRAM DESIGN FEATURES AND ONGOING IMPLEMENTATION EXPERIENCE

2.1 Overview of the Evolution of the MHS Pilot Programs

Our discussion in this Chapter reflects all eight MHSOs through Month 18 of their Phase I pilot programs⁸ and information obtained through two rounds of site visits to each of the MHS programs as well as 6-month interim telephone calls to keep informed of program evolution. In this chapter, we describe changes in program implementation (including engagement of the refresh population), implementation challenges, and MHSO cited reasons for requesting early termination from the MHS Phase I pilot for the three MHSOs that terminated prior to the writing of this report.

In general, the MHS interventions continue to vary in a number of important ways (e.g., the presence of on-the-ground nurse support, conduct of nursing home visits, specific programs to support care at the end of life, home monitoring), however, all programs continue to provide MHS participants with telephonic care management services, including

- nurse-based health advice for the management and monitoring of symptoms,
- health education (via health information, videos, online information),
- health coaching to encourage self-care and management of chronic health conditions,
- medication counseling, and
- health promotion and disease prevention coaching.

Each MHS program has a nurse-based health coaching and health support program; however, the MHSOs continue to vary in how they implement the various aspects of their model. While all MHS interventions involve a telephonic nurse function, only a few of the MHSOs are actively engaged in serving an institutionally based population. Most of the MHS programs have an end-of-life care planning intervention. Most of the MHSOs have some form of on-the-ground efforts as part of their MHS interventions. And, most of the MHSO programs provide some type of monitoring devices at home to a small segment of their populations for whom they believed monitoring would be beneficial. Although six of the MHSOs⁹ received prescription drug information on medications filled by MHS participants who had purchased Part D plans, there is considerable variability in the extent to which the MHSOs use this information for medication counseling or other activities.

⁸ LifeMasters terminated at Month 17 and their experience is reflected through that time.

⁹ LifeMasters and McKesson did not receive Part D data because they had ceased operations by the time the Part D data were available from CMS.

We observed evolution of the MHS programs during the first 18 months of the Phase I pilot. While some of evolution reflected gained knowledge about the level of comorbidity among the MHS fee-for-service (FFS) beneficiaries, key substantive changes were made in most of the MHS programs to address unmet need for cognitive/psychological and social support services, and end-of-life care planning. Substantive changes were also made in the engagement and the initial stages of intervention of the refresh populations. Below we summarize key changes in the MHS programs focusing, first, on changes in the outreach to beneficiaries, providers, and the community, and second, on changes in program features. We conclude with implementation challenges reported by the MHSOs.

2.2 Changes in Outreach

Outreach to MHS Refresh Beneficiaries. All MHSOs but one (CHS) recruited refresh population participants during their second year of activity¹⁰. The seven MHSOs that did accept refresh populations conducted a range of activities to engage these beneficiaries, but they typically modified their original approach to more quickly engage the refresh beneficiaries and begin intervention services based on lessons learned from their engagement of the original populations. Since the sizes of the refresh populations were much smaller than the original populations, the MHSOs could streamline operations and handle the refresh outreach process more efficiently.

Most of the MHSOs retrained their internal staff or moved recruitment from external contractors to internal program staff for this second phase of engagement. Some MHSOs relied on external contractors to engage the original cohort, while they relied on their own staff or a blended approach of external contractors and internal coaching staff to engage the refresh cohort. One MHSO changed its staffing model to rely upon communication specialists rather than their MHS intervention nurses to locate the beneficiaries and make a “warm” transfer to health coaches to gain consent and initiate the intervention. A number of the MHSOs also asked MHS participants from the original cohort to help engage new participants and one MHSO even paid them. These veteran participants provided testimonials about their own experiences and shared their perspectives on why the program was helpful to them. The MHSOs that had relied upon social marketing characteristics during engagement of their original cohort – targeting those predicted to be most likely to agree to participate -- changed its recruitment strategy to be more like most of the other MHSOs – focusing first on beneficiaries they believed were at the highest risk for acute care utilization.

While the MHSOs reported engagement of the refresh beneficiaries was more manageable because of a smaller refresh population, they reported also facing some of the same outreach challenges reported with their original outreach efforts, including not having telephone numbers for all beneficiaries in their target population, and difficulty reaching beneficiaries who resided in institutional settings.

¹⁰ CHS expressed concerns that cost savings generated by care coordination activities that prevent complications of diabetes, such as amputations, may result in cost savings realized more than 3 years following intervention. Therefore, a much shorter time period for the refresh population would unlikely capture all of the cost savings actually generated by the intervention.

Outreach to Providers. All eight MHSOs continued to provide outreach and ongoing communication to medical providers that were viewed as the principal physician for their participating beneficiaries. Three MHSOs placed an increased emphasis on physician engagement through the use of breakfast or lunch meetings with MHS provider service representatives, attendance of MHS staff at local professional meetings, provider newsletters, and greater sharing and seeking of clinical information on MHS participating beneficiaries. For example, several MHSOs asked community-based physicians to: (a) verify that participants were indeed part of their panel; (b) confirm or correct participant contact information; and (c) provide clinical information. In exchange, MHSOs sent periodic reports to providers for those MHS participants in their panel, plus feedback on potential issues to help identify areas for future intervention by the physician and/or MHSO. Based upon physician feedback, one MHSO stopped sending participants' principal physician alert messages if the MHSO learned that beneficiaries had not received selected process-of-care measures based upon beneficiary self-report or Medicare claims data analysis. Physician feedback was that the information was often incorrect. Another MHSO modified forms sent to participating physicians requesting clinical information on MHS participants as they were viewed as too long and burdensome to complete.

In addition to physician outreach, two MHSOs increased their efforts to obtain information on acute care interventions on their participants directly from hospitals in a timely manner. This effort was primarily driven by the MHSOs' recognition during the early part of the pilot that CMS provided claims data on a monthly basis were not useful to them to identify acute care utilization in a timely enough manner to meaningfully intervene. The MHSOs also recognized during the early part of the pilot that they could not rely upon MHS beneficiaries or their families to notify them of admission in a timely manner. One MHSO launched a pilot program that linked their information system directly with a local hospital's information system. A second MHSO attempted and abandoned efforts to directly obtain discharge information from hospitals as it found that there was insufficient volume of admissions at most hospitals to pursue this course of action. Several additional MHSOs established data sharing agreements with the FFS claims processors for their geographic area to obtain information on hospitalizations.

Outreach to the Community. Although most of the eight MHSOs provided some type of outreach to the community, the level and emphasis of this outreach activity varied widely by MHS program. At program launch, MHSOs consulted with community advisory boards or tribal leaders, worked with local Area Agencies on Aging (AAAs) and local community agencies to help locate potential MHS participants and to gain support for the MHS program, and developed special relationships with social service agencies and other community service agencies. After launch, most MHSOs increased their outreach to community organizations in efforts to find social support services for participating beneficiaries. Several MHSOs developed comprehensive community service directories to provide extensive point-of-service assistance and referral support to their participants and health coaches.

2.3 Changes in MHS Elements

Individualized Assessment of Refresh Population and Ongoing Assessment of Original Cohort Participants. The seven MHSOs that accepted refresh populations still conduct a health assessment when a new participant agrees to join the program; however, most of these MHSOs streamlined their assessment or moved more aggressively to complete their assessments in a

shorter time frame, i.e., 6 weeks of engagement rather than 6 months. Although the content of the assessments differed somewhat across the MHSOs, nurses generally asked questions to identify the presence of primary and comorbid diseases, recent health care utilization (e.g., hospitalizations and emergency room visits), deficits in activities of daily living, medications, cognitive issues, and current health and social support services. Depression screening remained a key focus of all MHSOs. The information obtained from the individualized assessment was used to help determine the type and level of intervention and to set or modify self-management goals.

Ongoing assessments of the original populations continued across all eight MHSOs but several abbreviated their assessment tools and most MHSOs increased reliance on nurses' clinical judgment and critical thinking abilities to identify at the point of contact each participant's psychosocial and clinical issues that could lead to an acute care encounter. All MHSOs' nurses reported having developed skills to more directly identify pressing clinical or social issues that could lead to clinical deterioration.

Intensive Case Management. A portion of the intervention population continues to include very sick beneficiaries, requiring close monitoring, in-home visits, and additional support services. Most of the eight MHSOs provide additional services to a small segment of MHS beneficiaries for whom they believe additional services are warranted. They do so either with their own staff or via contractual arrangements. One MHSO initiated in-home medical assessments by physicians for very sick beneficiaries who had no physician of record or were too ill to leave their home. Several other MHSOs refined their MHS intervention model to increase the number of field-based nurses or advanced practice nurses for more intensive home monitoring and support even though the percentage of beneficiaries that received home monitoring is relatively low. Four of the MHSOs provided in-home or nursing home support services to less than 10% of their participants and for only several months. Most MHSOs added social workers or enhanced their community resource guides to help the telephonic and field-based nurses to identify psychosocial services or deal with psychosocial issues.

One MHSO contracted with an external entity to provide care management support to their nursing home residents; while several of the other MHSOs used field-based staff nurses for these services. One MHSO ended care coordination support for their nursing home residents as the number of participants per nursing home was smaller than expected and the quality of care they were receiving was higher than expected.

Another MHSO expanded access to complex case management services. Beneficiaries were identified for this service through data mining of their intervention beneficiaries' Medicare claims data. Case managers were provided with lists of participants with high acute care utilization patterns from claims data to target for greater in-person or telephonic intervention. Several MHSOs modified their predictive models to more accurately identify and then target beneficiaries that are likely to have escalating costs in the near term or are at high risk of dying.

Staffing Changes. In addition to modifications to staffing discussed above, most but not all of the MHSOs made additional changes to their core disease/case management staff to deal with staff turnover or to modify or enhance their MHS program intervention. Two MHSOs that experienced between 25 and 35% staff turnover made modifications to their interview screening process and the way they explain the nature of the care manager role for the MHS program. The

high mortality rate among the MHS beneficiary population was viewed as a stress that needed to be better highlighted during the interview process. In addition, one of these MHSOs hired licensed practical nurses that were able to take on the responsibilities required to serve as disease management care managers. Most of the other MHSOs experienced lower rates of turnover. All MHSOs increased the number of social workers or community resource coordinators to address the high level of socioeconomic needs. Most MHSOs increased the number of care managers during year 2, to handle, in part, the refresh population. However, several did so to enhance their service delivery; while another did so in anticipation of a Phase II rollout on a much larger scale. One MHSO hired an additional 50 health coaches, or doubled their staff, of which approximately 30 were hired to handle the refresh population and the remaining 20 were hired both to replace those from normal attrition and to expand their community based coaching team to focus upon smoothing the transition between various care settings. Several MHSOs modified their care team relationship with subcontractors, who were providing purely telephonic disease management support, by pairing them with case managers located in the geographic areas of operation. This allows for more community-based interaction between the care team and the MHS beneficiaries.

End-of-Life Care Planning. Most of the eight MHSOs either developed or enhanced efforts related to advanced care planning and end-of-life care planning. One MHSO developed a special team of nurse care managers and social workers who have experience in end-of-life issues to address advanced care planning with participants at risk of death within a six-month period. Other MHSOs increased training of their nurse coaches and care managers to deal directly with participants or their families with advanced care planning, including the possibility of incorporating hospice into their planning. One MHSO partnered with external experts from several universities to help train their staff in end-of-life care planning issues.

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 making lifestyle changes. All the MHSOs continued to provide a range of educational resources, including literature, videos, Internet resources, and coaching by a nurse or other care manager via telephone or in person. One MHSO expanded its educational efforts to include nutrition classes to help participants adopt and maintain a healthy diet. Another MHSO launched an initiative focused on supporting participants' efforts to reach evidence-based clinical treatment goals. Foci of this initiative are on blood pressure targets, control of lipid levels, and appropriate use of aspirin, angiotensin converting enzyme inhibitors, and lipid lower medications.

Medication Counseling and Support. All eight MHSO programs continued to include efforts to optimize the medication regimens of participating beneficiaries. Interventions ranged from monitoring compliance and assessing the appropriateness of complex pharmaceutical regimens, to face-to-face meetings with pharmacists at assessment centers for the management of complex cases. The MHSO that provided their intervention beneficiaries with face-to-face meetings with pharmacists moved their intervention from the pharmacy location into their assessment centers to increase use of this element of their program. They found some beneficiaries were uncomfortable talking to pharmacists associated with a store that was not their usual prescription filling source or were unwilling to drive the distance to the pharmacy. One MHSO purchased a software tool for use by health coaches to assess their participants' medication regimes enabling the health coaches to easily assess proper usage, dosage, and potential side effects. Another MHSO developed comprehensive lists of pharmaceuticals in

participants' homes and sent these full pharmacy lists to the beneficiaries' principal physician for review and feedback to the MHSO program and beneficiary.

By the middle of Year 2 of the pilot, the operating MHSOs received CMS data on Part D prescription drug events. The data provided were for only a proportion of their entire MHS population (i.e., generally less than 50% per MHSO), which is reflective of the proportion of all Medicare FFS beneficiaries who had purchased Part D plans. These data provided some limited assistance to several of the MHSOs, who used them to identify individuals with previously undisclosed health conditions/problems (such as depression or chronic pain) and modify interventions to better support them. Several MHSOs reported that they found reconciliation between beneficiary self-report usage and Part D data on filled prescriptions time consuming and not wholly successful making the Part D data of less value than they had originally anticipated.

Monitoring, Feedback, and Follow-Up. Several programs continued to offer or increased biomonitoring of beneficiaries by placing scales or other equipment in their homes; while other MHSOs continued to rely primarily upon participants to report their weights, blood sugars, or other measures via e-mail or telephone. There were no major changes in this aspect of the MHSO programs.

After Hours Access to MHS Support Services. One feature of several MHSO programs is round-the-clock availability of support services. In these programs, participants may call and speak to a nurse or other provider at any time they are having a problem or would like to ask a question. Other programs have systems in place so participants may leave a message about a problem or question and receive a return call within a certain period of time, varying from 30 minutes to the next business day. There was limited modification of this feature during the first 18 months of Phase I. One MHSO received feedback from beneficiaries that the telephone menus to access their triage nurses were difficult to navigate. Therefore, the MHSO simplified the phone tree scripting so that beneficiaries who did not press any telephone key were connected directly with a triage nurse.

Referrals for Provision of Community-Based Ancillary Services. Not all of a participant's needs are provided directly by the MHSOs. During the first 18-month period of the pilot, all MHSOs increasingly recognized the need for transportation or other social support services typically provided by a community service organization (e.g., social workers, dietitians). All MHSOs enhanced their relationships with other service providers and programs and helped selected beneficiaries receive these services through their participation in the MHS program. Several MHSOs hired social workers to assist care managers or health coaches to better deal with the large unmet social service needs of the MHS participants. Most MHSOs also increased their efforts to screen a larger proportion of their participants for depression and to encourage participants to seek medical care for depression when they screened positive. One MHSO hired social workers specifically to help coordinate receipt of behavioral health services. Several MHSOs exercised their option for technical assistance in depression care management offered to all the MHSOs by a team of experts from RAND, the University of Washington, and the University of Pittsburgh.

Information Management Systems. Each MHSO relies heavily on the use of information systems that include their own system of electronic health records, automated call center

operations, and data analysis and storage capabilities and facilities. The MHSOs use the CMS provided Medicare claims data and monitoring reports to varying degrees to analyze and modify their interventions. Several MHSOs blend the claims data with clinical data received from physicians or self-reported by participants to help guide the frequency or nature of intervention. One MHSO now develops goal-oriented quantitative data reports that are provided to each health coach, which reflects their coaching activity coupled with outcomes data, to guide their performance.

Risk Stratification. All of the MHSOs continue to use at least one method (i.e., data derived predictive risk score, nurse assessment, etc.) to stratify their populations into various categories of risk for the likelihood of having a high cost event (e.g., hospitalizations, emergency department visits), or deterioration in clinical health status. A number of MHSOs altered their strategy as they analyzed data and considered how to classify MHS beneficiaries who were typically sicker than populations that MHSOs had supported previously. For example, a number of the MHSOs rely on sophisticated predictive models using proprietary logic with more than 100 variables to identify gaps in care, create risk strata scores, and achieve operational efficiency. For two MHS programs, risk is recalculated with every new piece of information obtained on MHS participants. Most MHSOs use internally developed risk stratification systems to subdivide the MHS population into various risk categories. Where MHSOs found their internal stratification models did not adequately discriminate among different risk groups, they have relied on the Hierarchical Condition Categories scoring system to stratify their MHS populations. Several MHSOs continue to modify their predictive risk models to help inform their intervention operations.

Access to and Use of CMS Data. Operating MHSOs continue to receive CMS claims data for their intervention group participants on a monthly basis. In addition, comparison group data are provided to the MHSOs quarterly, both in aggregate reports and as deidentified claims data sets. A number of the MHSOs developed creative strategies to enhance their ability to manage MHS operations by obtaining hospital and nursing home inpatient census, Medicare claims, or other administrative data on a more frequent basis. Some of the MHSOs negotiated data sharing agreements with Medicare carriers, fiscal intermediaries, or other major health care partners, while others rely primarily on the data provided from CMS and its MHS contractors. As noted earlier, by the middle of Year 2 of the pilot, the operating MHSOs received CMS data on Part D prescription drug events and used them to varying degrees to better understand the clinical conditions of their participants and to look for drug-drug interactions, for example.

2.4 Implementation Challenges Reported by the MHSOs

By the end of Month 18, the MHSOs reported several remaining implementation challenges. While the MHSOs reported engagement of the refresh beneficiaries was more manageable because of a smaller refresh population, they reported also facing some of the same outreach challenges reported with their original outreach efforts, including not having telephone numbers for all beneficiaries in their target population, and difficulty reaching beneficiaries who resided in institutional settings. In spite of the smaller number of refresh beneficiaries and changes in their refresh engagement processes, participation rates were lower for the refresh populations than observed during the initial 6-month period for the original populations with the exception of Health Dialog (see Section 3.2 for more detail).

The MHSOs recommend excluding from future programs beneficiaries who have exceptionally high severity risk scores, residents of long-term care facilities, beneficiaries entitled to Medicare because of disabilities, beneficiaries of advanced age, and beneficiaries who receive the majority of their care from the Department of Veterans Affairs. All MHSOs also reported more MHS participants than expected had extensive psychosocial needs requiring significant in-house and on-the-ground resources to manage their behavioral health issues as well as the establishment of an extensive set of relationships with community based organizations to assist in providing social services. The MHSOs recommend that exclusion of beneficiaries with severe cognitive impairment should be considered.

Additionally, all MHSOs stressed the continuing challenges of receiving hospitalization information on a timely basis to allow for transitional care assistance by the MHSO. The MHSOs found the time lag of 1 to 2 months from time of hospital discharge to receipt of CMS provided claims data did not allow them to identify when a beneficiary had been hospitalized in a timely manner. The MHSOs also reported that beneficiaries or their families did not consistently contact them around the time of an admission thus additional information was necessary. As noted earlier, several MHSOs negotiated with fiscal intermediaries or hospitals to receive discharge information on a more frequent basis; however, our evaluation of changes in hospitalization rates during the pilot did not differ between those that had more timely data and those that relied upon the CMS provided data.

A related challenge that was raised by several MHSOs was their inability to obtain laboratory data directly from clinical laboratories and strongly recommended that future disease management programs be structured so that beneficiaries be considered members of an MHS program similar to a Medicare Advantage plan thereby allowing the MHS programs to obtain clinical information directly from health care providers. In addition, several MHSOs faced challenges implementing their pharmacy intervention due to the availability of Part D data for only the portion of the MHS intervention population that purchased a Part D plan and for which the Part D plans agreed to share the data. One MHSO noted the challenge of reconciling beneficiary provided prescription drug information with data provided by the Part D plans making the Part D data less valuable than originally anticipated.

Another design feature that has continued to be challenging is randomization at the individual beneficiary level, rather than at the physician practice level. Although randomization at the beneficiary level was specified in the program solicitation, MHSOs have offered that randomization at the practice level would have been more favorable to practice-level rather than beneficiary-level interventions and they believe greater physician involvement.

During the second round of site visits, all the MHSOs mentioned continuing problems with the system CMS uses to pay MHSOs and track individual beneficiary eligibility for MHS. For the MHS program, CMS utilizes the same system that is used to pay Medicare Advantage plans; effective January 2006 that system was modified to accommodate Part D plan payments. A number of MHSOs continued to relay problems related to the system modifications as well as concerns that neither the old nor new system is sufficiently tailored to the MHSO program. Specific concerns cited include that the system is difficult to use for financial monitoring of program payments because it does not easily allow the MHSOs to link specific beneficiaries with payments. CMS provides the MHSOs with aggregate payment statements, but individual

beneficiary records must be queried one at a time to confirm eligibility and payment, including retroactive changes.

2.5 Reasons for Requested Early Termination from the MHS Phase I Pilot

Through 2007, three organizations requested early termination of their programs. LifeMasters Supported Self Care ended their MHS program December 31, 2006, McKesson Health Solutions, LLC, ended their MHS program May 31, 2007, and CIGNA Health Support ended their MHS program January 14, 2008. RTI conducted in-person or telephonic interviews with key MHS program staff with each of these MHSOs and report their cited reasons for requesting early termination from the MHS Phase I pilot. After the CMS announcement in January 2008 that Phase I would cease at the end of the three year pilot period, Green Ribbon Health and XLHealth requested early termination by 2½ and 5 months, respectively.

LifeMasters Supported Self Care

LifeMasters terminated after 17 months of operations and reported their three primary reasons for deciding to terminate early from the MHS intervention:

- After program launch, LifeMasters found that the rural population in Oklahoma (which made up the bulk of the intervention group), had much lower utilization of evidence-based care than they had expected. LifeMasters believes that the efforts of their nurses to increase adherence to evidence-based care increased utilization for their MHS participants during the first 17 months of the program. Over time, LifeMasters believed that their program would result in a decrease in preventable hospitalizations and, thus, costs, but did not expect to achieve sufficient reductions in the remaining 1 ½ years of the pilot.
- LifeMasters voiced a concern that from the program's outset reports from CMS showed that the control group's per member per month cost and utilization were several percentage points lower than the intervention group's per member per month cost and utilization at the time of randomization. At the time LifeMasters was assessing operating options, CMS was unable to commit to adjusting this difference either in real time or at reconciliation.
- LifeMasters also believed that a 5% savings criterion was too ambitious a goal for a new, unproven pilot program; however, its MHS application originally estimated a 5.2% savings at the end of the three year period.

McKesson Health Solutions

McKesson terminated after 21 months of operations and reported that a combination of factors contributed to its decision to terminate its MHS program early:

- McKesson believed that the CMS quarterly reports comparing its intervention and control group performance were not sufficient to monitor its progress toward meeting its Medicare savings requirements and effectively manage the financial risk associated with the MHS program. By December 2006, CMS provided McKesson

with beneficiary-level claims data for the control group but McKesson felt that these data were overly de-identified (e.g., no dates of service were provided) and not sufficiently timely to be of use in its program operations.

- McKesson was concerned about its ability to achieve sufficient Medicare savings to cover its fees and the 5% net savings requirement based on program outcomes information available as of the middle of March 2007. At the time McKesson was assessing its options, CMS was unable to commit to adjusting this difference either in real time or at reconciliation. McKesson had estimated a 3-year 5% savings net of fees in its original application.
- McKesson perceived there to be a lack of collaboration and transparency in its relationship with CMS during the interim reconciliation process. McKesson noted that the success of risk contracts such as the one in place for MHS relies on collaboration and transparency between the care management vendor and the payer, so that the two parties come together and mutually agree on the results obtained, rather than a situation where one party determines the results independently, whether or not the second party can analyze the same data and obtain the same result.

CIGNA Health Support (CHS)

CIGNA Health Support terminated its MHS program after 28 months of operations and reported that a combination of factors contributed to its decision:

- At the time CHS was assessing its options, the lack of a decision by OMB on changing the 5% savings criterion to budget neutrality was a major factor in deciding to terminate early. The decision to remove the 5% savings requirement was made in December 2007 and after CHS had requested early termination. CHS had estimated a 3-year 5% savings net of fees in its original application.
- CHS also expressed concern that they had on-going questions about how the financial performance metrics were being calculated and the influence of the volatility on the projections that were never resolved.
- The knowledge that they did not have to stay in Phase I to the end in order to be eligible to bid on Phase II.

CHAPTER 3

PARTICIPATION RATES IN THE MEDICARE HEALTH SUPPORT (MHS) PROGRAMS AND LEVEL OF INTERVENTION

3.1 Introduction and Methods

Our first analysis in this chapter evaluates the rates of participation in the MHS programs and level of intervention. Medicare beneficiaries are confirmed as participants when the MHSO has reached the beneficiary or caregiver, receives verbal consent to participate, and begins services. The MHSOs received monthly management fees for the full original and refresh populations for the first six months of engagement of each of the populations. After the initial 6-month outreach period for each cohort, the MHSOs accrue management fees for only those beneficiaries who verbally consent to participate and only during periods of participation. Participation continues until a beneficiary becomes ineligible for the MHS program or informs the MHSO or CMS that he or she does not want to receive further services from the program. Participants may drop out of the program at any time and begin participation again at any time, as long as they are eligible. However, over one-half of all MHS beneficiaries are continuous participants, meaning that once they consent to participate they participate for all of their MHS eligible days. Many of these beneficiaries are fully eligible and continuous participants meaning they meet MHS eligibility criteria for the entire first 18 months of the pilot and participate all days after consent. Never participants are individuals in the intervention group who did not consent to participate or were not reachable by the MHSO for all months in which they were eligible to participate and remain in the intervention group's "at risk" population. Beneficiaries who decline participation may be re-contacted by the MHSO after a sentinel event, such as a hospitalization or ER visit.

The overall design of the MHS pilot follows an intent-to-treat (ITT) model and the MHSOs are held at risk for their monthly management fees based on the performance of the full population of eligible beneficiaries randomized to the intervention group and as compared to all eligible beneficiaries randomized to the comparison group. The MHS pilot has been designed to provide strong incentives to gain participation by all eligible beneficiaries in the intervention group. In our second round of site visits, most of the MHSOs reported that they had been successful at contacting beneficiaries from the original cohort who were nonparticipants in the initial 6-month period. Thus, we look more closely at participation over time and report participation rates for multiple periods—consented during the initial 6-month outreach period or consented during months 7 through 18. We report the percentage of intervention beneficiaries who never consented to participate and the reason for nonparticipation (refused or were never contacted/unable to be reached). We also report the percentage of beneficiaries, who after initial consent, were continuous participants (while eligible for the MHS program) and the percentage of beneficiaries participating for more than 75% of eligible days. 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, gets a new primary payer (i.e., Medicare becomes secondary payer), dies, elects the Medicare hospice benefit, or develops end stage renal disease (ESRD).

We also examine the level of intervention between the MHSO and its beneficiaries. Although the MHS interventions have a variety of elements (e.g., nursing home visits, specific programs to support care at the end of life, home monitoring, educational classes, and mailings),

all programs provide MHS participants with telephonic care management services that includes nurse-based health advice for the management and monitoring of symptoms, health education (via health information, videos, online information), health coaching to encourage self-care and self-management of chronic health conditions and medications, and health promotion and disease prevention coaching. And, all of the MHSOs have on-the-ground elements of their MHS interventions to varying degrees, ranging from advanced practice nurses who provide intensive case management support to assessment centers where beneficiaries are encouraged to go for in-person interaction. Therefore, we examine the number of telephonic and in-person contacts with the MHSO. For each participating beneficiary in the original populations, the MHSOs provided to CMS on a monthly basis a count of the number of completed telephone calls and the number of visits to community-residing beneficiaries or nursing home residents that occurred during each month.

Because beneficiaries may have intermittent periods of eligibility and participation, for this analysis we examine the number of contacts for the original cohort beneficiaries who were eligible and participating for each month of the 7 to 18 month pilot period¹¹, which is a period of maximum intervention by the MHSOs. Beneficiaries who died during the 12 month period but were fully eligible and participating up to their point of death area also included. The number of intervention beneficiaries that meet these criteria range from 8,516 for XLHealth to 13,817 for Green Ribbon Health (GRH). These participants reflect between 72 and 98% of eligible intervention beneficiaries in the 7-18 month period. For this subset of full participants, we examine the percent of beneficiaries with at least one or more completed calls or visits by number of months of telephonic support. We also report the frequency of months of telephonic support by risk stratification. We use either the MHSO's internal stratified risk score, which often determined frequency of interaction and was provided to CMS on a monthly basis, or an RTI calculated HCC score using claims data for the year prior to program launch if the MHSO did not report an internal risk score. If an internal risk score was provided, we assign the beneficiary to the category that represents the most frequently occurring category.

In the first Report to Congress (McCall *et al.*, 2007), we observed a pattern whereby beneficiaries who agreed to participate within the first 6-months of the pilot tended to be considerably healthier and less costly in the prior year compared to those beneficiaries who never consented to participate. Because the pilot design is an ITT model, engagement of less costly intervention beneficiaries will require the MHSOs to have a larger effect on participants to achieve the required savings. We re-examine selected baseline characteristics of participants and never participants defined at the time of each MHSO going live. Medicare claims data for the one year prior to program launch are used to calculate an HCC score, rates of all cause hospitalization and ER visits, and per beneficiary per month (PBPM) total Medicare payments. Utilization and payment rates are weighted by the number of days the beneficiary met the eligibility criteria for the MHS pilot divided by 365 days. We define a participant as a beneficiary with at least one day of eligibility who agreed to participate at least one day during months 1-18 of the pilot.

¹¹ We report results for LifeMasters for months 6-17.

3.2 Participation Rates During the First 18-months of the MHS Pilot Programs for the Original Populations and the first 6-months for the Refresh Populations

Tables 3-1 and *3-2* present the participation rates for the original populations and the refresh populations, respectively. *Table 3-1*, for the original populations, shows the distribution of all intervention beneficiaries by date of initial consent to participate (within the initial 6-month period versus Months 7 to 18) and by their level of participation (continuous participation after consent and participation more than 75% of eligible days). The rate of intervention beneficiaries that never participated during any of the first 18 month period is also displayed as well as the reason for never participating (refused or not contacted/unable to be located). *Table 3-2*, for the refresh populations, shows the participation rate for the first six months of their engagement and reason for never participating.

Participation Rates for the Original Populations. Over three-quarters of all intervention beneficiaries verbally consented to participate in the MHS program during the first 18 months of the pilot. Health Dialog led the MHSOs with a 95% consent rate. In contrast, XLHealth had the lowest consent rate of 74%, followed by LifeMasters at 76%. The remaining MHSOs had consent rates between 81 and 89%. Most of the participants consented in the initial 6-month period of the pilot. Between one-half and two-thirds of beneficiaries who consented were continuous participants. Of the beneficiaries who never consented to participate, the refusal rate ranged from 0.3% for Health Dialog to 13% for Aetna. The percent not contacted or unable to be located ranged from 4 to 15%.

Participation rates are heavily influenced by length of eligibility during the 18-month period. An alternative measure of participation is the percentage of beneficiaries who participated more than 75% of their eligible days. Health Dialog exhibits the greatest level of participation with 85% of their intervention beneficiaries participating for more than 75% of their eligible days. In contrast, XLHealth has just over 50% of their beneficiaries participating more than 75% of their eligible days. The remaining MHSOs range between two-thirds and three-quarters.

Table 3-1
Participation status of the Medicare Health Support Organizations’ original intervention populations during the first 18 months of the Medicare Health Support pilot

Status	Aetna	Healthways	CIGNA Health Support	Health Dialog	Green Ribbon Health	LifeMasters ¹	McKesson	XLHealth
Participation Rate	83%	89%	89%	95%	84%	76%	82%	74%
Agreed to Participate in Months 1-6	81	83	83	92	83	71	75	65
Agreed to Participate in Months 7-18	2	6	6	3	0 ²	5	6	9
Length of Participation								
Continuous Participation after Engagement	59	65	58	66	57	55	57	48
Beneficiaries Participating More than 75% of Eligible Days	66	76	79	85	77	63	75	53
Never Consented to Participate Rate	17%	11%	11%	5%	16%	24%	18%	26%
Refused to participate when contacted by MHSO	13	3	2	0 ³	3	9	4	11
Not contacted/unable to be located	4	9	9	4	13	15	14	15

NOTES:

¹ LifeMasters examines months 7-17.

² The participation rate for Green Ribbon Health in Months 7-18 is 0.5%.

³ The refusal rate for Health Dialog is 0.3 percent.

SOURCE: RTI analysis of Medicare Health Support (MHS) participation data submitted by the MHSOs for the original population for Months 1 – 18 of the Phase I pilot.

Participation Rates for the Refresh Populations. Participation rates were lower for the refresh populations than observed during the initial 6-month period for the original populations (*Table 3-2*) with the exception of Health Dialog, who received verbal consent for participation from 96% of their refresh beneficiaries. McKesson received verbal consent from only about a third of their refresh beneficiaries. During our second site visit, we learned that McKesson had decided to focus its outreach efforts on a limited subset of its refresh population, those who were at the highest risk for acute utilization. Beneficiaries with recent use of acute care services were contacted first. There are also high rates of never participants within XLHealth’s and LifeMasters’ refresh populations, 37% and 43%, respectively. The remaining MHSOs had 6-month refresh population participation rates that were between 3 and 9 percentage points lower than 6-month participation rates for their original populations.

Table 3-2
Participation status of the Medicare Health Support Organizations’¹ refresh intervention populations during the first 6 months

	Aetna	Healthways	Green Ribbon Health	Health Dialog	LifeMasters ²	McKesson	XLHealth
Participant Rate	72%	74%	80%	96%	63%	34%	57%
Never Consented to Participate Rate	28%	26%	20%	4%	37%	66%	43%
Refused to participate when contacted by MHSO	23	12	2	0 ³	4	3	14
Not contacted/unable to be located	5	14	18	4	32	63	29

NOTES:

¹ CIGNA Health Support did not request a refresh population.

² LifeMasters’ participation rates are for a 5-month period because it requested early termination.

³ Health Dialog’s refusal rate is 0.02%.

SOURCE: RTI analysis of Medicare Health Support (MHS) participation data submitted by the MHSOs for the refresh population for the first 6 months of their engagement.

As discussed in Chapter 1, refresh beneficiaries were assigned to each MHSO through a randomization process similar to that used to assign the original populations with one noted exception related to the targeted clinical condition(s). The MHSOs requested and CMS agreed to first assign beneficiaries with heart failure (HF) and diabetes, then heart failure-only, and finally diabetes-only. For four MHSOs, the randomization process and eligible populations in their geographic areas resulted in their supplemental populations containing only beneficiaries with HF. **Table 3-3** displays the demographic characteristics of the original and refresh populations at the time of go-live for each MHSO and estimates of health status, utilization, and Medicare expenditures derived from Medicare claims data for the 12-month period prior to the go-live date for each MHSO. Only beneficiaries who had at least one day of eligibility during the pilot are included in this table. We do not conduct statistical testing of the differences but display the populations’ characteristics for illustrative purposes. We generally observe in the refresh populations a higher rate of females, a higher average HCC score indicating poorer health status, higher rates of all cause hospitalization and ER visits, and higher average per beneficiary per month (PBPM) total and acute care Medicare payments. Observed differences between the two populations may affect the rate of participation to some degree.

**Table 3-3
Characteristics of the Original and Refresh Medicare Health Support Intervention Populations**

	Aetna		Healthways		Health Dialog		Green Ribbon Health		LifeMasters		McKesson		XL Health	
	Original	Refresh	Original	Refresh	Original	Refresh	Original	Refresh	Original	Refresh	Original	Refresh	Original	Refresh
Total (n)	20,220	4,284	19,976	4,248	20,013	5,939	22,590	2,715	20,110	4,602	20,181	4,558	19,451	5,191
Weighted Total (n = FTEs)	20,167	4,271	19,934	4,238	19,977	5,919	22,526	2,708	20,053	4,588	20,110	4,524	19,233	5,166
Threshold Conditions (%)														
Heart Failure Only	26	6	20	30	22	40	24	65	26	64	21	65	20	45
Diabetes Only	51	0	60	0	55	43	55	0	49	0	54	6	58	27
Heart Failure and Diabetes	23	94	20	70	22	18	20	35	25	36	25	28	22	27
HCC score³ (Average)	2.5	3.1	2.5	3.1	2.5	2.6	2.5	2.8	2.4	2.6	2.2	2.5	2.3	2.6
Women (%)	49	57	50	59	47	66	44	51	49	63	54	64	51	61
In Medicaid (%)	16	20	15	20	17	13	14	5	21	25	43	48	33	33
Chronic Conditions (%)														
Coronary artery disease	47	66	42	63	53	57	49	64	47	59	38	51	39	50
Acute & chronic renal disease	17	37	15	32	16	23	15	25	14	23	14	23	16	26
Peripheral vascular disease	10	14	8	11	9	10	7	9	7	8	7	9	6	8
Cardiac dysrhythmias & conduction disorders	33	49	28	48	33	40	34	54	27	43	23	40	25	41
Charlson comorbidity index (Average)	3.9	4.4	4.0	4.0	4.1	3.4	4.0	3.4	3.7	3.2	3.6	3.1	3.5	3.3
Prior Utilization (Rates⁴ per 1,000)														
All cause hospitalization	996	1,347	847	1,214	893	877	688	798	857	1,015	818	989	715	793
All cause ED/Obs visits	742	882	1,011	1,090	929	935	824	765	1,171	1,321	1,484	1,654	1,235	1,156
Prior Medicare Payments: Average PBPM⁵														
Total Medicare payments	1,524	1,973	1,395	1,902	1,268	1,272	1,226	1,420	1,291	1,499	1,209	1,433	1,150	1,262
Acute hospital payments	681	940	682	1,028	520	513	427	538	512	628	486	611	441	492
Physician payments	349	401	342	393	293	270	430	450	274	283	258	236	287	286

NOTES:

¹ Numbers reported for the intervention periods include only persons who had at least one day of eligibility during the pilot.

² Data are weighted by the fraction of days in the year prior to go-live that the beneficiary met the MHS eligibility criteria.

³ RTI calculated Hierarchical Condition Categories (HCC) score based on the 12 month period prior to go-live for each population.

⁴ Utilization rates are from a 12-month period prior to go-live for each population.

⁵ PBPM = per beneficiary per month calculated for the 12 months prior to go-live for each population.

SOURCE: RTI analysis of Medicare Health Support original and refresh population's characteristics using Medicare Part A and B claims data 2004-2006 and CMS' Enrollment Data Base and Daily Eligibility File.

3.3 Original Populations' Level of Intervention in the MHS Programs

We also examine the level of intervention between the MHSOs and their assigned beneficiaries. On a monthly basis, the MHSOs provide to CMS a count of the number of completed calls and visits to their intervention beneficiaries that occurred during the prior month. In **Table 3-4**, we report the *number of months* that a participant received at least 1 or more telephonic contacts. This analysis is restricted to beneficiaries who were eligible and a participant in all of the months of 7-18 of the pilot (Months 6-17 for LifeMasters so that we capture a full 12 month period) or were fully eligible and participating until the time of their death. This represents between 72 and 98% of the MHSOs' full intervention populations. We restrict the analysis to this subgroup to reduce confounding from partial participation, which allows for a more straightforward evaluation of the distribution of months of support provided by the MHSOs and whether there is evidence of selective targeting of beneficiaries for intervention contacts based upon level of perceived need.

Table 3-4
Percent distribution of original population participants¹ across number of months of telephonic support during Months 7 – 18 of the Medicare Health Support Pilot²

	Months of Telephonic Support					
	0	1	2-5	6	7-11	12
Aetna	19	34	27	2	14	4
Healthways	5	4	38	12	39	3
CIGNA Health Support	4	4	30	12	45	6
Health Dialog	3	3	30	18	45	1
Green Ribbon Health	9	9	52	11	19	0
LifeMasters	10	21	25	2	28	14
McKesson	15	9	52	9	15	0
XLHealth	5	11	75	6	3	0

NOTES:

¹ Beneficiaries had to be eligible and participate in all Months 7-18 or until they died to be included in the analysis.

² LifeMasters examines Months 6-17 to capture a full 12-month period of telephonic support.

SOURCE: RTI analysis of Medicare Health Support (MHS) participation data and telephonic and in-person encounter data submitted monthly by the MHSOs for the original population for Months 1 – 18 of the Phase I pilot.

The majority of fully eligible and participating MHS beneficiaries received between 2 and 5 months of telephonic support¹². However, we do observe fairly disparate patterns of months of telephonic contact across the MHSOs. Between 6% (Health Dialog) and 53% (Aetna) of fully participating beneficiaries received no or only one month of telephonic contact during Months 7-18. Twenty-four percent of McKesson's and 10% of LifeMasters' fully participating beneficiaries received no or only one month of telephonic intervention during the one year period. However, LifeMasters has the largest percentage of beneficiaries with 12 months of telephonic support. In contrast, GRH, McKesson, and XLHealth provided no beneficiaries with 12 months of support. CHS and Health Dialog provided between 7 and 11 months of telephonic support to 45% of their beneficiaries. Healthways' distribution was bimodal; almost 40% of their beneficiaries received between 2 and 5 months of telephonic support while another almost 40% received between 7 and 11 months of telephonic support.

The percentage of fully participating beneficiaries who received any in-person support is very low (not displayed) and reflects the telephonic design of most MHS programs. Less than 10% of beneficiaries in the MHS programs run by Aetna, Healthways, CHS, and Health Dialog received any in-person support. In contrast, XLHealth provided in-person support to over one-half of their fully participating beneficiaries. This also reflects their focus upon in-person assessment centers. However, almost one-half of XLHealth's beneficiaries who received in-person support had encounter(s) in only 1 month. Between 12 and 21% of beneficiaries in MHS programs run by LifeMasters, McKesson, and Green Ribbon Health received any in-person support.

We also examined whether there is evidence of selective targeting of beneficiaries for intervention contacts based upon level of perceived need as determined by the MHSO derived risk scores or an RTI calculated HCC score for the year prior to each MHSO's go-live date. The top half of **Table 3-5** displays the distribution of fully eligible and participating beneficiaries during Months 7-18 across three risk strata defined by the MHSOs or by the HCC strata. There is considerable variation across the MHSOs in the distribution of beneficiaries across the three strata. Three-quarters of LifeMasters' participating beneficiaries are assigned to the high stratum and only 3 participants (0%) are assigned to the low stratum. In contrast, XLHealth has 92% of its participants assigned to the low stratum and only 10 participants (0%) to the high stratum. The three MHSOs that did not report to CMS risk strata information have roughly one-half of their participating beneficiaries assigned to the low stratum, 30% to the medium stratum, and 20% to the high stratum.

The bottom half of **Table 3-5** displays the average number of months of telephonic support by risk stratum. There is a general pattern of increasing average number of months of telephonic support as the level of risk score increases. During the site visits, all MHSOs stated that they targeted interventions based upon perceived need for services due to clinical deterioration or risk of hospitalization. Thus, we expected to see a pattern of increasing level of intervention across the risk strata. However, there is limited separation in average number of months of contact between the three levels of risk score for most of the MHSOs.

¹² Monthly support is defined as any number of calls in the month.

Table 3-5
Percent distribution of fully participating Medicare Health Support original populations’
intervention beneficiaries by risk stratum and average number of months of telephonic
contacts during months 7-18 of Phase I Pilot by risk stratum

	Percent of Fully Eligible and Participating Beneficiaries		
	Risk Strata		
	<u>Low</u>	<u>Medium</u>	<u>High</u>
<u>MHSO</u>			
Aetna ¹	29	33	38
Healthways ²	49	29	21
CIGNA Health Support ²	53	28	18
Health Dialog ²	48	30	22
Green Ribbon Health ³	19	53	28
LifeMasters ¹	0	24	76
McKesson ¹	35	31	34
XLHealth ¹	92	8	0

	Average Number of Months of Telephonic Contact			
	Risk Strata			
	<u>Low</u>	<u>Medium</u>	<u>High</u>	<u>All</u>
<u>MHSO</u>				
Aetna ¹	1.6	3.0	4.2	3.1
Healthways ²	5.4	6.2	6.6	5.9
CIGNA Health Support ²	6.2	6.7	7.2	6.5
Health Dialog ²	5.8	6.3	6.5	6.1
Green Ribbon Health ³	3.0	4.1	4.9	4.1
LifeMasters ¹	3.7	4.7	5.6	5.3
McKesson ¹	2.6	3.7	4.6	3.6
XLHealth ¹	3.0	4.1	3.9	3.1

NOTES:

¹ The risk strata is defined by the MHSO. The risk category is the one a beneficiary was in the most in Months 7-18.

² No risk strata was provided by the MHSO, so we used the RTI calculated HCC risk score for the year prior to the MHSO going live.

³ GRH did not provide a risk category for about 48% of their beneficiaries, the RTI HCC score was used for those with missing data.

^a LifeMasters has 3 beneficiaries assigned to the low risk stratum. Months 6-17 evaluated to capture a full 12-month period of intervention support.

^b XLHealth has 10 beneficiaries assigned to the high risk stratum.

SOURCE: RTI analysis of Medicare Health Support (MHS) participation data, telephonic and in-person encounter data, and MHSO risk stratification data submitted monthly by the MHSOs for the original population for Months 1 – 18 of the Phase I pilot.

3.4 Characteristics of Participants and Never Participants

In the first Report to Congress (McCall *et al.*, 2007), we observed a pattern whereby beneficiaries who agreed to participate within the first 6-months of the pilot tended to be considerably healthier and less costly in the prior year compared to those beneficiaries who never consented to participate. We re-examined selected baseline characteristics of participants and never participants; however, the results presented in this Report to Congress cannot be directly compared to the results from the first Report to Congress. The current analyses exclude any beneficiaries who had no eligibility for the MHS pilot during the first 18 months of the pilot and the baseline utilization and payment rates are constructed using the final CMS daily eligibility file used to determine baseline per member per month payment estimates for the purpose of financial reconciliation.

Table 3-6 displays the baseline difference in demographic characteristic percentages, average HCC score, and PBPM Medicare payments, and average rates of hospitalization and ER visits between the participating beneficiaries and the never participating beneficiaries. Negative (positive) differences signify that the participating beneficiaries have lower (higher) observed values than the never participating beneficiaries. We find that the participant populations remain different from the never participant populations across numerous demographic, health status, utilization, and payment characteristics reviewed. With the exception of McKesson, the proportion of participating beneficiaries with Medicaid enrollment is between 3 and about 14 percentage points lower than for never participants meaning that most MHSOs have not been as successful recruiting Medicare/Medicaid dual enrolled beneficiaries to participate. Six of the MHSOs have lower rates of Medicare beneficiaries who are under age 65, or beneficiaries with disabilities, among their participating beneficiaries. Five MHSOs have lower rates of African American beneficiaries among their participants, while LifeMasters, McKesson, and XLHealth have higher rates of African American beneficiaries among their participants.

Mean HCC risk scores calculated for the 1-year period prior to each MHSO going live range from 20 to 40% lower for participants than for never participating beneficiaries. All cause hospitalization and ER visit rates during the year prior to going live range from 196 to 631 and 41 to 568 per 1,000 beneficiaries lower than rates for never participating beneficiaries, respectively. Mean participant PBPM payments range from \$267 to \$792 lower than PBPMs for the never participants. Thus, we continue to observe participants to be a healthier and less costly subset of the whole intervention group across all MHSOs.

Table 3-6
Differences in the original populations' demographic characteristics, health status, utilization, and costs at baseline between beneficiaries who agreed to participate within the first 18-month Medicare Health Support Phase I pilot period^{1,2} and those who never agreed to participate

	Differences Between Participants and Never Participants ³							
	Aetna	Healthways	CIGNA Health Support	Green Ribbon Health	Health Dialog	LifeMasters	McKesson	XLHealth
Dual Medicare/Medicaid Enrollee (%)	-7.9 **	-4.1 **	-6.8 **	-9.1 **	-14.4 **	-2.7 **	0.7 **	-2.6 **
Percent Less than Age 65	-5.1 **	-4.1 **	-8.6 **	-4.8 **	-9.3 **	-0.9	-1.2	-1.5 *
Percent Race is Black	-3.2 **	-3.2 **	-5.9 **	-3.0 **	-4.8 **	1.1 **	7.1 **	1.4 *
Average HCC Score	-0.4 **	-0.3 **	-0.2 **	-0.3 **	-0.6 **	-0.2 **	-0.3 **	-0.2 **
All Cause Hospitalization (per 1,000)	-449 **	-301 **	-332 **	-296 **	-631 **	-196 **	-309 **	-227 **
All Cause ER Visits (per 1,000)	-390 **	-551 **	-568 **	-422 **	-509 **	-437 **	-41 **	-248 **
Total Medicare PBPM ⁴ payments	-\$545 **	-\$430 **	-\$384 **	-\$347 **	-\$792 **	-\$292 **	-\$436 **	-\$267 **

NOTES:

¹ LifeMasters examines months 7-17.

² To be included in this analysis, intervention beneficiaries had to have had at least 1 day of eligibility during the first 18 months of the Phase I pilot period.

³ Percentages, averages, and rates of never participants are subtracted from the observed values for participating beneficiaries.

⁴ PBPM = per beneficiary per month.

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

SOURCE: RTI analysis of Medicare Health Support (MHS) participation data submitted by the MHSOs for the original population for Months 1 – 18 of the Phase I pilot, CMS's daily eligibility file, and Medicare Part A and B claims data for the one year baseline period prior to the start of each MHSO; 2004 - 2006.

CHAPTER 4

MEDICARE HEALTH SUPPORT BENEFICIARY SURVEY OF SATISFACTION, CARE EXPERIENCE, SELF-MANAGEMENT BEHAVIORS, AND SELF-REPORTED PHYSICAL AND MENTAL HEALTH

4.1 Introduction

The Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of beneficiary satisfaction. In addition, the evaluation seeks to answer a broader set of research questions related to whether the programs improved communication with their health care team, knowledge about their chronic condition(s), and self-management skills and led to behavioral change among participants. The evaluation includes these additional foci to better understand the factors for program success.

As part of the independent evaluation of Phase I of the pilot programs, a survey was conducted at two points in time to help measure the programs' effects on the original intervention populations relative to the original comparison populations. This report focuses on RTI's assessment of the effect during the first 18 months of the pilot of the MHS programs on changes in intervention beneficiary satisfaction, care experience, self-management behaviors, and self-reported physical and mental health functioning relative to observed changes in the comparison beneficiaries.

The eight MHS pilot programs largely employ strategies that aim to increase quality of care and decrease costs by increasing beneficiary knowledge of their clinical condition through educational and coaching interventions related to their chronic condition(s), improving communication with their care provider(s), and improving self-management skills. Success in changing beneficiary behavior should result in better "control" of their chronic conditions and a reduction in anticipated functional decline. Better control of their chronic conditions should reduce acute exacerbations that can lead to acute care interventions. We hypothesize that improved communication with their health care team, perceived helpfulness of education and coaching interventions, and increased self-management skills concomitant with a reduction in acute exacerbations will increase the MHS intervention beneficiaries' overall rating of satisfaction that their health care providers are helping them to cope with their chronic condition.

The following four evaluation questions are addressed in this chapter:

- Do the MHS programs lead to higher levels of beneficiary satisfaction that their health care providers are helping them to cope with their chronic condition(s) than is reported by beneficiaries in the comparison group?
- Do the MHS programs improve perceived helpfulness of educational materials and discussions related to medications, diet, and exercise, and the quality of communication with the health care team?
- Do the MHS programs result in greater engagement in health behaviors?

- Do the MHS programs mitigate the anticipated functional decline or improve physical and mental health functioning?

Increasing beneficiary satisfaction is one of three conditions that must be met in Phase I of the pilot for expansion to a Phase II. The other two conditions for expansion are improving clinical quality of care and achieving targets for savings.

4.2 Survey Domains

The overall design of the MHS pilot follows an intent-to-treat model, so that the underlying population for the survey sample included all beneficiaries assigned to the intervention, as well as all comparison group members. For this reason, the survey contains questions relevant to all beneficiaries in the pilot regardless of their intervention or comparison group status. The conceptual framework underlying the development of the survey instrument emphasizes interventions that help chronically ill beneficiaries understand their disease, manage their symptoms, and perform self-care activities leading to better health outcomes and greater satisfaction that their health care providers are helping them to cope with their chronic condition.

Satisfaction Measure. The primary measure of satisfaction was a rating related to levels of beneficiary satisfaction that their health care providers are helping them to cope with their chronic condition(s). Beneficiaries were asked to include health care providers with whom they interact telephonically or face-to-face.

Care Experience Measures. The survey also included two care experience measures that were principal foci of each of the MHS programs during health coaching sessions. Beneficiaries were asked to rate the helpfulness of specific activities such as discussions about medications, diet, exercise, and coping with stress or sadness. These items are consistent with chronic disease management and the MHS program interventions. Beneficiaries were also asked to rate the quality of communication with their health care team. The MHS programs focused on empowering beneficiaries to have more effective communication with their health care providers.

Self-Management Measures. A goal of disease management is to improve compliance with self-care activities that may maintain or improve functioning. The survey included three sets of questions related to self-management: setting goals, self-efficacy, and self-care activities. The survey included two questions that asked if someone from their health care team helped set goals or a plan to take care of their health problems. The questionnaire also included items that assess self-efficacy, or the belief that the beneficiary has confidence that he or she can perform activities that enhance health. And, third, the questionnaire included questions about eight self-care behaviors, adapted from the summary of Diabetes Self-Care Activities instrument (Toobert et al., 2000). Both self-efficacy rating items and questions on self-care activities focus on the same behaviors such as taking medications, managing weight and fluid and salt intake, exercising, and planning meals.

Physical and Mental Health Functioning Measures. The survey instrument included four important physical and mental health functioning outcome measures that cannot be obtained from claims data. Physical and mental health functioning levels were assessed by responses to the Veterans RAND-12 (VR-12) instrument (Kazis, 2004); which consists of 12 items, half of

which reflect physical function (PHC) and half that are indicators of mental health function (MHC). Mental health functioning was also measured by the Patient Health Questionnaire-2 (PHQ-2), a widely used depression screening tool (Kroenke et al., 2003). The questionnaire also included a second set of physical functioning items that assess the beneficiary’s ability to perform with or without assistance six standard activities of daily living (ADLs) including bathing, dressing, eating, getting in and out of chairs, walking, and using the toilet.

4.3 Survey Design and Implementation

A pre-post longitudinal survey design was used, with a baseline survey fielded within six months of each MHSO’s launch and a follow-up survey fielded a year after the baseline survey was fielded. The same beneficiaries were surveyed at two points in time using the same instrument as well as the same multiple-mode, multiple-contact methodology. The overall baseline response rate was 70% (n= 8,360) with MHSO-specific response rates ranging from 63% to 73%. Respondents to the baseline survey who died or became ineligible between baseline and follow-up survey administration (n=469) were removed for the follow-up survey. Respondents from LifeMasters’ baseline survey were not re-surveyed due to that program’s termination of participation prior to the fielding of the second round (n=1,103). The overall response rate for the follow-up survey was 77.8% (n=6,639) with MHSO-specific response rates ranging from 73% to 84%. Among the intervention group respondents included in this analysis, a high proportion of the beneficiaries, 88% to nearly 100%, were participants in the MHS pilot for some period of the first 18 months. **Table 4-1** displays the number of completed follow-up surveys and response rates by intervention and comparison group in each MHSO.

Table 4-1
Completed Medicare Health Support original populations’ follow-up surveys and response rates by Medicare Health Support Organization¹

MHSO	Intervention group	Control group	MHSO response rate
Healthways	341	438	76.1%
Health Dialog	347	391	81.8
McKesson	361	392	74.6
Aetna	308	364	75.9
CIGNA	329	394	73.1
Green Ribbon Health	362	408	84.1
XLHealth	320	412	80.1
Total Surveys	2,368	2,799	5,167
Total Eligible Sample	3,057	3,582	6,639
Response Rate	77.4%	78.1%	77.8%

NOTES:

¹ LifeMasters’ termination occurred prior to the follow-up survey being fielded. LifeMasters is not included in the beneficiary survey reporting.

SOURCE: RTI Analysis of Medicare Health Support original population beneficiary follow-up surveys conducted March 5, 2007 through July 30, 2007.

4.4 Analytic Methods

The MHS programs reflect a dynamic process of system change leading to behavioral change leading to improved health outcomes, and the type of experimental design within this pilot calls for a pre/post, intervention/comparison analytic approach—referred to as a difference-in-difference approach—to provide maximum analytic flexibility. Further, the overall design of the MHS pilot follows an intent-to-treat model, so that, the underlying population for the survey sample included all beneficiaries assigned to the intervention regardless of their willingness to participate in the pilot program, as well as all comparison group members. For this reason, the survey contained measures relevant to all beneficiaries in the pilot regardless of their intervention or comparison group status. In our analyses, we compare the responses of beneficiaries in the intervention and comparison groups who completed both the baseline and follow-up surveys.

There are numerous estimation techniques that may be used to analyze data from before and after studies like the MHS beneficiary survey. One well-known method is the paired t-test. Using this approach, change over time is computed by subtracting the baseline score from the follow-up score for each beneficiary. The estimated intervention effect is the difference between the mean intervention and comparison group change scores. This is a conservative estimation technique.

An alternative method is to conduct an analysis of covariance (ANCOVA). ANCOVA pools information from both groups and provides an estimate of the intervention effect that is statistically adjusted for individual respondents' baseline levels of the outcome. When baseline and follow-up scores are correlated, as they are for most of the behavioral outcomes of interest in the MHS program, ANCOVA produces a more precise estimate of the intervention effect because the group difference has a smaller standard error than observed using a paired t-test method. The ANCOVA approach will capture not only improved outcomes but also slower rates of declines for outcomes such as functional status.

We compared these two methods in this evaluation. Conducted separately for each MHSO, all analyses were based on beneficiaries who completed the same survey questions at both the baseline and follow-up administrations. ANCOVA models were estimated by ordinary least squares regression equations consisting of the baseline score and an indicator distinguishing intervention from comparison respondents as explanatory variables. Dichotomous outcomes were evaluated using logistic regression. In all comparisons, we found that the standard errors were smaller for ANCOVA than for the paired t-test method. As a result, ANCOVA detected more intervention effects (generally favoring the MHS pilots) that were statistically significant at conventional levels.

Controlling for baseline levels may underestimate the effect of the intervention for some MHSOs if they had already achieved improvements by the time the baseline survey was administered. However, at the time of survey implementation most MHSOs were still involved primarily in outreach activities or conducting baseline assessment of health status and beneficiary needs. More importantly, we examined baseline differences between the intervention and comparison groups and found few significant differences; the lack of differences in baseline

scores between groups suggests that the threat of underestimating the intervention effect is small in these analyses.

The data used to estimate intervention effects were weighted to reflect the composition of the surviving eligible intervention and comparison beneficiaries. The weights were based on the three original program stratification criteria or 12 strata defined by the intersection of three Hierarchical Condition Categories (HCC) risk score groups (low: >1.35 and <2.00 , medium: >2.00 and <3.10 , and high: >3.10), two Medicaid status groups (Medicaid or no Medicaid), and two disease classification groups (heart failure or no heart failure). The weights therefore adjust for any differential attrition between the intervention and comparison groups within an MHSO. The approach is also analogous to the eligibility weighting used for the clinical and financial data analyses conducted as part of the independent evaluation of MHS.

4.5 Findings

The MHS pilot programs employ strategies to improve quality of care while reducing costs by empowering Medicare beneficiaries to better manage their care. They do 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 beneficiary use of medications, eating habits, and exercise, as well as interacting more effectively with their health care team. The MHS programs hypothesize that lifestyle changes and better communication with providers will mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions.

A summary of statistically significant intervention effects for all survey outcomes across the seven MHSOs is shown in **Table 4-2**. Because LifeMasters terminated its MHS contract prior to our fielding of the follow-up survey, their beneficiaries are not included in these analyses. Intervention effects are denoted by plus and minus signs with one plus (negative) sign indicating a positive (negative) intervention effect at the 0.05 significance level and two plus (negative) signs indicating a positive (negative) intervention effect at the 0.01 significance level.

Satisfaction Measure. The primary measure of satisfaction was a rating related to levels of beneficiary satisfaction that their health care providers are helping them to cope with their chronic condition(s). Beneficiaries were asked to include health care providers with whom they interact telephonically or face-to-face. Of the seven MHSOs, only Health Dialog and Aetna demonstrated a positive intervention effect related to helping beneficiaries cope with their chronic condition. Aetna's intervention effect was driven primarily by stability of the satisfaction rating among the intervention group versus the observed decline among its comparison group. Prior research has shown that Medicare fee-for-service beneficiaries are generally satisfied with their health care and our baseline findings are consistent with those observations (Bernard and Urig, 2002).

Table 4-2
Summary of Medicare Health Support original populations' survey of satisfaction, self-management, and functioning

	Statistically significant intervention effect						
	AETNA	Healthways	CIGNA	Health Dialog	McKesson	Green Ribbon Health	XLHealth
Beneficiary satisfaction							
Health care team helped beneficiary cope with chronic condition	+			++			
Beneficiary experience with care							
Number of helpful discussion topics			++	++	+		
Quality of communication with health care team			++	++			
Self-management							
Percent helped set goals		+		+	+	+	+
Percent helped make a plan				++	+		
Self-efficacy ratings (level of confidence)							
Take all medication	+						
Plan meals and snacks							
Manage your blood sugar level	+				-		
Check feet for sores or blisters	+						+
Exercise 2 or 3 times weekly	+						
Limit salt							
Weight yourself							
Limit fluids							
Self-care activities (number of days per week)							
Prescribed medications taken							--
Blood sugar tested		++			++		
30 minutes of continuous physical activity	+						
Feet were checked		+					
Followed healthy eating plan							
Weight was measured					+		
Salt was limited							
Fluids were limited				++			
Physical and mental health functioning							
PHC score							
MHC score							
PHQ-2 score							
Percent PHQ-2 score indicating depression	+						
Number of activities of daily living – difficult to do				+			
Number of activities of daily living – receiving help							

NOTES:

1. LifeMasters' termination occurred prior to the follow-up survey being fielded. LifeMasters is not included in the beneficiary survey reporting.
2. Statistical significance determined using Analysis of Covariance: positive intervention effect denoted as + p<.05, ++ p<.01; negative intervention effect denoted as - p<.05, -- p<.01

SOURCE: RTI Analysis of Medicare Health Support original population beneficiary baseline and follow-up surveys conducted between March 1, 2006 and June 30, 2006 and March 5, 2007 through July 30, 2007.

Care Experience Measures. We also asked care experience questions related to two key aspects of the MHS interventions; helpfulness of educational materials and discussions with their health care team about medications, diet, exercise, and coping with stress or sadness, and quality of communication with their health care team. Health Dialog demonstrated a positive intervention effect related to both measures of helpfulness of discussions and educational

materials and communication with their health care providers. CHS showed statistically significant improvement in the number of helpful discussions as well as a strong beneficiary intervention effect in the quality of communication with their health care team. McKesson showed a positive intervention effect on the number of helpful topics. Aetna, GRH, Healthways, and XLHealth showed no statistically significant intervention effects related to these two care experience measures.

Self-Management Measures. Another goal of disease management is to increase compliance of appropriate self-care behaviors among the chronically ill. The survey instrument included measures to capture changes in beneficiary self-management focusing specifically upon willingness to set self-management goals, self-efficacy, and engagement in self-care activities. The MHSOs were most successful in helping beneficiaries to set goals and make plans to address their care needs. Five of the seven MHSOs showed positive intervention effects related to setting goals (Healthways, Health Dialog, GRH, McKesson, and XLHealth) and two MHSOs showed positive intervention effects related to developing a plan (Health Dialog and McKesson).

In spite of positive effects on setting goals, there was little meaningful improvement in self-efficacy or self-care activities. For example, both Health Dialog and McKesson showed statistically significant intervention effects in raising the proportion of beneficiaries who reported receiving help setting goals and making plans. For Health Dialog, this effect did not translate into increases in self-efficacy and an increase in only one of eight self-care activities. For McKesson, there is an observed decrease in one of the self-efficacy items and an increase in two of eight self-care activities. At baseline, mean scores related to self-efficacy were relatively high suggesting that on average beneficiaries were ‘somewhat sure’ of their ability to manage their health and chronic illness.

Self-efficacy assessments are based in part on previous experience with the recommended behavior. Given the selection criteria for inclusion in the pilot, this is not a population of beneficiaries newly diagnosed with heart failure or diabetes. At baseline beneficiaries typically reported levels of self-efficacy with mean ratings averaging from 3.2 to 4.5 (3 = moderately confident of their ability to perform self-care activities) out of a maximum of 5 (extremely confident). Out of 56 significance tests, six statistically significant effects were found; five favored the intervention and one did not. The positive effects include having confidence to appropriately take medications (Aetna), manage blood sugar levels (Aetna), check feet for sores and blisters (Aetna and XLHealth), and exercise 2 or 3 times weekly (Aetna); while the negative effect was related to managing blood sugar level (McKesson).

Beneficiaries were also asked about their compliance with a number of self-care activities relevant to their underlying condition. We found variation by type of activity in the baseline compliance rates among both the intervention and comparison groups for self-care activities. For example, baseline rates for taking medications as prescribed were quite high ranging from 6.7 to 6.8 days per week, leaving little room to detect improvement; on the other hand, daily weights among beneficiaries with HF ranged from 2.4 to 3.8 days per week providing an opportunity for improvement. Given the high level of reported compliance at baseline it was not surprising to find that there was no positive intervention effect on medication adherence; however, there was one negative intervention effect with XLHealth showing a decrease in medication adherence among the intervention group. The behavior with the lowest baseline compliance was sustained

physical activity; there was one observed intervention effect related to this activity. Aetna showed an intervention effect in both increasing self-efficacy and reported compliance with physical activity. Healthways and McKesson showed a positive intervention effect in two areas; blood sugar testing and checking feet for blisters and sores, and blood sugar testing and weight measurement, respectively. GRH and CHS showed no intervention effect in any of the self-care behaviors.

Physical and Mental Health Functioning Measures. Lastly, the survey instrument included four important physical and mental health functioning outcome measures. We examined the effect of the MHS intervention on several measures of physical and mental health functional status. Given the age and the frailty of this population, we did not expect to see significant improvements in physical and mental health function; we did, however, expect to see a mitigation of the slope of the decline for the intervention group. Between baseline and the follow-up surveys, we noted small declines in physical and mental function in nearly all of the sites. This decline is a common finding in elderly populations, and in particular, a chronically ill elderly population. The goal of intervention is often to mitigate the slope of decline rather than to prevent decline altogether. If the MHSOs were to have an impact on these outcomes, we would observe that the rate of decline for the comparison group would be greater than that of the intervention group. We found only one statistically significant intervention effect in the ADL domain: the decline in the number of ADL difficulties at follow up was mitigated for the Health Dialog intervention group compared to its comparison group.

Given the emphasis of many of the MHSOs on depression screening we were expecting to see some improvement in mental health; however, only one of the MHSOs had an impact on beneficiary mental health status with a mitigation of the decline in the PHQ score relative to the comparison group. Both the intervention and comparison groups within most MHSOs exhibited similar declines in physical and mental functioning between the baseline and follow-up surveys.

4.6 Conclusion

The Medicare Health Support (MHS) authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves beneficiary satisfaction and improves clinical quality of care and achieves targets for savings, the Secretary shall enter into agreements to expand the program (or components) to additional geographic areas not covered under the program as conducted in Phase I.

Only two of the MHSOs, Health Dialog and Aetna, improved beneficiary satisfaction as measured by beneficiary assessment that their health care team helped them cope with their chronic condition, our principal measure of satisfaction. Health Dialog also demonstrated positive intervention effects related to two other care experience measures; while the remaining MHSOs showed limited mixed results. None of the seven MHSOs included in the beneficiary survey analyses demonstrated consistent positive intervention effects across the four domains of satisfaction, care experience, self-management activities, and physical and mental health functioning. The focus of the pilot program interventions was largely on impacting beneficiary behavior to better manage their chronic illness. Yet these results show little evidence of changes in self-efficacy or self-care. We did not observe any consistent pattern of positive intervention effects by disease cohort of heart failure only, diabetes only, and heart failure and diabetes.

CHAPTER 5

INTERIM FINDINGS ON QUALITY OF CARE IMPROVEMENT AND HEALTH OUTCOMES

The Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of quality improvement measures and health outcomes. The legislation also mandated a second Report to Congress that contains findings with respect to quality improvement, such as adherence to evidence-based guidelines, re-hospitalization rates, and health outcomes. In this chapter, we present analyses of changes in the rate of receipt of evidence-based process-of-care measures; changes in the rate of hospitalizations, re-admissions, and ER visits; and changes in mortality rates. We present *interim* results for an 18-month period for the original populations and *preliminary* 6-month results for the refresh populations.

In our evaluation, we have selected measures that may 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 are only available for the intervention beneficiaries that are participating during the pilot. Further, beneficiary self-report is subject to recall error and to the willingness of beneficiaries to provide the information.

As part of the ongoing monitoring of the pilot, CMS receives quarterly performance reports that contain quality of care and health outcomes measures that overlap to some degree with our evaluation measures (e.g., annual HbA1c and low-density lipoprotein (LDL) testing for beneficiaries with diabetes, and rates of acute care utilization and mortality). The MHSOs are at financial risk for performance improvement for a subset of the performance monitoring measures however, the “at risk” measures differ to some degree across the MHSOs and the definitions of improvement for achieving the performance standards differ considerably across the MHSOs. The monitoring and “at risk” quality and health outcomes performance measures are calculated using Medicare administrative data, beneficiary self-report, and physician supplied information. And, some measures are calculated for the intervention group only using beneficiary self-reported data. Additionally, some measures are calculated for only those beneficiaries who are actively participating during the measurement quarter or those for whom the MHSO has obtained the information. Under an intent-to-treat model and our difference-in-differences evaluation approach, we require information for the pre- and pilot periods and for both the intervention and comparison populations.

5.1 Methodology

5.1.1 Quality of Care Measures

We define quality of care as adherence to evidence-based guideline-concordant care, and selected four claims-based measures related to the threshold conditions of heart failure (HF) and diabetes as the focus of our evaluation: rate of annual HbA1c testing (diabetes); rate of dilated retinal eye examination (diabetes), rate of low-density lipoprotein cholesterol (LDL-C) testing (diabetes or HF), and rate of urine protein screening (diabetes). Beneficiaries with both heart

failure and diabetes are included in both disease cohorts (i.e., heart failure with or without diabetes and diabetes with or without heart failure). Because Medicare claims data do not contain laboratory results, we are restricted to assessing rate of compliance rather than level of control. National Quality Forum (NQF)-endorsed National Voluntary Consensus Standards for Physician-Focused Ambulatory Care specifications are used to create the four process-of-care measures and have been endorsed for the reported clinical conditions.

For the original populations, we created process-of-care measures for the 12-month period immediately prior to each MHSO's go-live date and for Months 7 through 18¹³. The latter time period is after the initial 6-month engagement period ended and reflects maximum intensive focus within the first 18-month period. Because the process-of-care measures that we study are defined as annual rates of service, we believe that it would be inappropriate to evaluate the performance of the MHSOs using only 6 months of intervention experience for the refresh population so we do not report any process-of-care quality measures. The first six months of refresh pilot experience also overlaps with the initial engagement period for that cohort during which time the MHSOs are attempting to contact the beneficiaries, conduct initial assessments, and begin intervention services.

Medicare claims for these measures are included regardless of MHS eligibility to ensure that we are fully capturing behavior of intervention and comparison populations that is not subject to Medicare eligibility or payment rules and to provide credit to the MHSOs in case the services occurred after exposure to the MHS intervention and during the intervention period. One could envision that the MHSO 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 is included in the Medicare claims files during a period of ineligibility as a denied claim, it reflects actual receipt of the service and is therefore included in our analyses.

Rates per 100 beneficiaries are calculated for the intervention and comparison groups for the 12-month baseline period and for the 7- to 18-month intervention 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. Statistical testing of the change in the rate of receipt of the quality of care measures is performed at the individual beneficiary level. The standard method for modeling a binary outcome, such as receiving a HbA1c test or not, is logistic regression. The experimental design for the MHS pilot also requires that the variance of the estimates be properly adjusted for the repeated (pre- and post-) measures observed for each beneficiary within a nested experimental design. The MHS pilot is based on eight nested cohorts of Medicare beneficiaries (i.e., one population from each MHSO) who were randomized to intervention and comparison groups within 12 strata defined by the intersection of three Hierarchical Condition Categories (HCC) risk score groups (low: >1.35 and <2.00, medium: >2.00 and <3.10, and high: >3.10), two Medicaid status groups (Medicaid or no Medicaid), and two disease classification

¹³ Months 7 to 17 is the reference period for LifeMasters as they requested early termination at Month 17.

groups (heart failure or no heart failure). In addition, an eligibility fraction¹⁴ ranging from zero to one was assigned to the pre- and post-time periods for each beneficiary. STATA SVY was used to fit the model with robust variance estimation. Operationally, the 12 strata and a beneficiary identifier were included in the SVYSET statement to reflect the stratified sampling design and to denote beneficiaries as the primary sampling unit (PSU) since randomization was conducted at the beneficiary level. The period of eligibility was included as the weight.

5.1.2 Health Outcomes

We focus on three utilization measures to capture the intervention's effectiveness in improving the quality of outpatient care which is expected to reduce exacerbations of the chronic diseases that result in acute institutional care. Three intermediate clinical outcome variables—hospitalizations, re-admissions, and ER visits, including observational bed stays—were constructed for all beneficiaries and separately by the threshold conditions of heart failure and diabetes. For the original populations, beneficiaries with both heart failure and diabetes are included in both disease cohorts and rates are calculated for a 12-month period prior to launch of each MHSO's pilot program and for Months 7–18 of its intervention period¹⁵.

For the refresh populations, the randomization process and eligible populations in their geographic areas resulted in the supplemental populations for four the MHSOs containing only beneficiaries with HF with or without diabetes. These four MHSOs do not have any beneficiaries with diabetes only. Thus, rates were constructed for all beneficiaries and separately for beneficiaries with the threshold conditions of heart failure and diabetes and heart failure only. Hospitalization and ER visit rates are calculated for the first 6-month intervention period and for a comparable 6-month period¹⁶ during the year prior to each MHSO's refresh go-live date. The comparable 6-month period in the prior year was selected to remove the seasonality influence on these measures. We do not report readmission rates due to the short 6-month intervention period.

Medicare claims for these measures are included if the beneficiary was MHS eligible on the day the service started. For statistical testing, STATA SVY was used to fit a negative binomial model with robust variance estimation. The denominator for the hospitalization and ER visit measures are all eligible beneficiaries. The denominator for the readmission measure is all beneficiaries who had a hospitalization in the measurement period. In conducting the difference-in-differences statistical testing of readmission rates only beneficiaries who had a hospitalization in each of the measurement periods are included in the analysis. For each utilization measure, the difference-in-differences rate is reported and reflects the growth (or decline) in the intervention group's mean utilization rate relative to the growth (or decline) in the comparison group's mean utilization rate.

¹⁴ The eligibility fraction is calculated by dividing the number of months eligible for the pilot by the number of total months in the period, i.e., 12.

¹⁵ Note that we examined months 7-17 for LifeMasters for the original population.

¹⁶ Note that we examined months 1-5 for LifeMasters for the refresh population.

Another outcome metric in this evaluation is mortality. We evaluate mortality rates and time to death as both can substantially influence the cost of care for MHS beneficiaries. If the MHS interventions reduce mortality or prolong life, the intervention group beneficiaries may incur higher pilot period costs by living longer. On the other hand, all MHSOs focus upon encouraging end-of-life care planning through the use of advanced directives and consideration of hospice. These interventions may mitigate the potential increase in costs. We compare differences in mortality rates between the intervention and comparison groups from the original and refresh go-live dates and the end of the 6-month period for the refresh populations and the 18-month pilot period for the original populations. Mortality rates are not adjusted for periods of ineligibility because, unlike the claims-based measures, mortality rates are not sensitive to missing data. Statistical comparisons of the rates were made using a t-test of differences in mean rates between the intervention and comparison groups.

5.2 Clinical Quality and Health Outcomes Results for the Original Populations

5.2.1 Quality of Care

Analyses of changes in quality of care during Months 7-18¹⁷ of the Phase I pilot relative to the year prior to the pilot reveal modest improvement in receipt of evidence-based process-of-care measures for the original populations (*Table 5-1*). 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.

Across 40 measures, (five measures for each of the eight MHSOs), there was improvement in 16 (or 40%). For beneficiaries with heart failure (with or without diabetes), rates of cholesterol testing in the year prior to the pilot ranged from 55 to 71%. During Months 7-18 of the pilot, the intervention groups' rates of change of cholesterol testing were 2 to 4 percentage points higher for four MHSOs relative to their comparison groups' rates. For CHS and Healthways, their intervention groups' rate grew more than their comparison groups' rate. For Health Dialog and LifeMasters, their intervention effect was to prevent the downward trend in the rate observed within their comparison groups.

¹⁷ Months 7-17 for LifeMasters is reported.

Table 5-1

Rate of receipt of guideline concordant care measures per 100 Medicare Health Support intervention beneficiaries during the year prior to launch of the MHS pilot and rate of change during months 7-18 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Original Populations

	Aetna	Healthways	CIGNA Health Support	Green Ribbon Health	Health Dialog	LifeMasters	McKesson	XLHealth
	Rate/100	Rate/100	Rate/100	Rate/100	Rate/100	Rate/100	Rate/100	Rate/100
Heart Failure Beneficiaries (with and without diabetes)								
Cholesterol Screening								
Rate one-year prior to go-live	66	71	66	75	63	60	55	66
D-in-D rate during months 7-18	1.2	4.0 *	3.5 *	1.2	2.4 *	2.6 *	0.5	-0.1
Diabetes Beneficiaries (with and without heart failure)								
Cholesterol Screening								
Rate one-year prior to go-live	76	81	76	85	77	69	65	75
D-in-D rate during months 7-18	1.8 *	3.2 *	3.0 *	0.1	0.0	2.1 *	2.4 *	0.6
HbA1c								
Rate one-year prior to go-live	83	88	87	88	85	81	81	87
D-in-D rate during months 7-18	0.3	2.0 *	2.4 *	1.6 *	0.6	0.8	1.5 *	0.5
Urine Protein Screening								
Rate one-year prior to go-live	67	72	72	74	71	65	66	70
D-in-D rate during months 7-18	0.5	1.7 *	2.3 *	0.9	-0.4	1.4	1.0	0.9
Retinal Eye Exam								
Rate one-year prior to go-live	40	38	32	41	42	33	32	32
D-in-D rate during months 7-18	1.7	2.0 *	0.2	-1.0	0.7	1.7	1.0	1.3

NOTES:

1. Statistical significance testing of the difference-in-differences rate is conducted using STATA and a logistic regression model with robust variance estimation. * p<.05
2. Beneficiaries with both heart failure and diabetes are included in both disease cohorts.
3. We examined Months 7-17 for LifeMasters.

SOURCE: RTI Analysis of Medicare Health Support original populations' rate of receipt of guideline concordant care for months 7 to 18 of Phase I pilot compared to one year prior to each program's launch using Medicare Part B claims data 2004–2007 and the MHS daily eligibility file.

For beneficiaries with diabetes (with or without heart failure), four evidence-based process measures were evaluated. Rates of performance in the year prior to the pilot ranged from 65 to 85% for cholesterol screening, 81 to 88% for HbA1c testing, 65-74% for urine protein screening, and 32 to 42% for retinal eye examination. During Months 7-18 of the pilot, intervention groups at six of the MHSOs showed modest positive intervention effects.

Two MHSOs, Health Dialog and XLHealth, did not demonstrate any intervention effects for beneficiaries with diabetes. Further, XLHealth did not demonstrate any intervention effects for beneficiaries with heart failure or diabetes. Of the six MHSOs that demonstrated positive intervention effects for cholesterol screening, three were able to improve screening rates for both clinical cohorts while the other three were successful at improving rates for only one of the clinical cohorts.

5.2.2 Health Outcomes

Acute Care Utilization. Rates of hospitalization, readmission, and emergency room (ER) visits per 1,000 original population intervention beneficiaries for the year prior to go-live for each MHSO are presented for all beneficiaries and stratified by the two threshold conditions (i.e., heart failure and diabetes). Beneficiaries who have both heart failure and diabetes are considered in each of the disease groups. Rates of hospitalization are for all-cause and then for threshold condition-specific principal diagnoses. Rates of re-admissions and ER visits are for all-cause only. Below the utilization rates are the difference-in-differences rates of change observed between the baseline period and Months 7-18 of the pilot. Negative difference-in-differences rates indicate that the intervention group's mean rate of hospitalization or ER visits declined more than the comparison group's mean rate or the intervention group's mean rate of hospitalization or ER visits grew at a lower rate than the comparison group's mean rate. Positive difference-in-differences rates indicate that comparison group exhibited lower rates of growth of hospitalization or ER visits or greater decline than the intervention group.

Across the 120 comparisons (15 measures for each of the eight MHSOs), there were no statistically significant reductions in the rate of growth in hospitalizations, re-admissions, or emergency room (ER) visits in the original population intervention groups relative to the comparison groups (***Table 5-2***). The six statistically significant differences in the displayed measures show the intervention groups with higher rates of growth than their respective comparison group. Across the MHSOs and clinical cohorts, we observe a fairly consistent pattern of increasing rates of hospitalizations, readmissions and ER visits within the comparison groups during the first 18 months of the pilot. To date, the MHSOs have not been successful at reversing the observed acute care utilization trajectory observed in the comparison groups.

Table 5-2
Rate of acute care utilization per 1,000 Medicare Health Support intervention beneficiaries during the year prior to launch of the MHS pilot and rate of change during months 7-18 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Original Populations

	Aetna	Healthways	CIGNA Health Support	Health Dialog	Green Ribbon Health	LifeMasters	McKesson	XLHealth
	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000
All Beneficiaries								
All Cause Hospitalization								
Rate one-year prior to go-live	935	789	666	816	633	809	766	658
D-in-D rate during months 7-18	-21.1	17.7	3.6	20.7	10.9	8.0	2.4	-7.3
Heart Failure Hospitalization								
Rate one-year prior to go-live	146	113	100	132	87	103	116	85
D-in-D rate during months 7-18	-6.2	-2.8	3.8	-7.0	3.1	2.6	5.0	1.7
Diabetes Hospitalization								
Rate one-year prior to go-live	31	29	26	25	19	27	36	24
D-in-D rate during months 7-18	-2.4	-3.7	-0.3	2.2	-0.6	-0.7	4.4	2.8
All Cause Readmission								
Rate one-year prior to go-live	436	402	316	343	270	329	337	312
D-in-D rate during months 7-18	-2.9	7.3	-3.0	51.9 *	38.2	43.8	-4.4	-7.1
All Cause ER Visit								
Rate one-year prior to go-live	732	988	1,214	849	790	1,134	1,448	1,213
D-in-D rate during months 7-18	-9.3	0.3	38.8	11.1	43.0	98.6 *	34.7	31.1
Heart Failure Beneficiaries (with and without diabetes)								
All Cause Hospitalization								
Rate one-year prior to go-live	1,149	1,069	886	1,053	810	994	971	825
D-in-D rate during months 7-18	-47.3	47.0	24.5	33.3	25.7	-19.1	-13.0	-20.1
Heart Failure Hospitalization								
Rate one-year prior to go-live	281	263	233	266	174	189	235	169
D-in-D rate during months 7-18	-14.7	-10.9	5.2	-5.3	11.1	4.3	5.4	3.7
Diabetes Hospitalization								
Rate one-year prior to go-live	22	24	14	17	14	17	26	16
D-in-D rate during months 7-18	-2.4	-3.9	4.2	4.9	0.1	-3.2	-6.1	9.9 *
All Cause Readmission								
Rate one-year prior to go-live	505	492	387	411	302	381	395	368
D-in-D rate during months 7-18	2.1	6.6	8.4	63.9	51.0	50.7	14.5	-8.4
All Cause ER Visit								
Rate one-year prior to go-live	747	1,085	1,363	914	841	1,177	1,531	1,253
D-in-D rate during months 7-18	30.5	25.5	55.2	-8.5	27.7	95.2	-15.6	-4.2
Diabetes Beneficiaries (with and without heart failure)								
All Cause Hospitalization								
Rate one-year prior to go-live	873	727	619	750	594	754	719	634
D-in-D rate during months 7-18	-6.2	18.5	-7.2	36.0	17.3	27.0	-16.1	2.4
Heart Failure Hospitalization								
Rate one-year prior to go-live	96	74	67	86	62	67	83	66
D-in-D rate during months 7-18	-7.2	-9.7	-1.3	-6.7	6.4	8.4	-6.3	2.2
Diabetes Hospitalization								
Rate one-year prior to go-live	40	36	31	32	24	36	45	29
D-in-D rate during months 7-18	-1.3	-4.6	-1.0	2.0	-1.0	-1.1	4.9	3.5
All Cause Readmission								
Rate one-year prior to go-live	420	387	298	318	270	313	328	310
D-in-D rate during months 7-18	2.7	10.5	-17.2	58.9 *	40.5	48.6	-40.7	-4.6
All Cause ER Visit								
Rate one-year prior to go-live	727	982	1,196	828	788	1,155	1,457	1,242
D-in-D rate during months 7-18	-14.8	16.9	44.1	3.6	63.8 *	113.4 *	24.7	35.1

NOTES:

1. Statistical significance testing of the difference-in-differences rate is conducted using STATA and a negative binomial regression model with robust variance estimation. * p<.05
2. Beneficiaries with both heart failure and diabetes are included in both disease cohorts.
3. We examined Months 7-17 for LifeMasters.

SOURCE: RTI Analysis of Medicare Health Support original populations' rate of acute care utilization for months 7 to 18 of Phase I pilot compared to one year prior to each program's launch using Medicare Part A claims data 2004 – 2007 and the MHS daily eligibility file.

Mortality. With one exception, there were no observed differences in mortality between the original intervention and comparison groups during the first 18-month period (**Table 5-3**). For Health Dialog, the intervention group experienced a 1.1 percentage point higher death rate than its comparison group. In our final set of analyses, we will explore more fully the rate of and time to mortality over the full 36-month period. We will examine the rate of take-up of the hospice benefit and the use of services immediately preceding death and the influence that these factors have on the rate of mortality as well as time to death. Most MHSOs have undertaken efforts related to advanced care planning and end-of-life care. These program initiatives could lead to higher rates of death or shorter time to death for a subset of the intervention beneficiaries in the short-run.

Table 5-3
Mortality rates during the first 18 months of Medicare Health Support Phase I pilot; original populations

	Mortality rate		
	Intervention	Comparison	Difference
	(%)	(%)	
Aetna	15.3	15.3	0.0
Healthways	13.4	13.4	0.0
CIGNA Health Support	14.1	14.3	-0.2
Health Dialog	17.1	16.0	1.1*
Green Ribbon Health	15.5	15.7	-0.2
LifeMasters	15.2	15.6	-0.4
McKesson	13.6	13.5	0.1
XLHealth	14.6	14.7	-0.1

NOTES:

1. Statistical significance testing of differences in the original populations' mortality rates between intervention and comparison beneficiaries is conducted using a t-test. * p<.05

SOURCE: RTI Analysis of Medicare Health Support original populations' mortality using the Medicare Enrollment Database and the MHS daily eligibility file.

5.3 Health Outcomes for the Refresh Populations

Refresh Population Acute Care Utilization. All cause acute care utilization for a 6-month baseline period is high ranging roughly from 414 to 688 admissions per 1,000 beneficiaries and 392 to 898 ER visits per 1,000 beneficiaries. Heart failure accounts for roughly one-quarter of all hospitalizations, a rate that is higher than observed for the original MHS population. Diabetes accounts for less than 4% of all cause admissions, a rate that is somewhat lower than observed for the original MHS population. The refresh populations were purposely selected for presence of heart failure. However, it remains that MHS beneficiaries have a significant amount of other clinical co-morbid conditions that result in the use of acute care

services. A review of the frequency of principal diagnoses for non-diabetes or HF acute events reveals that many are related to pneumonia or other respiratory diseases and coronary artery disease and related cardiac conditions.

For three MHSOs, the trend in rates of all cause hospitalization was upward during the first 6-month intervention period relative to a comparable 6-month period in the prior year; for the other four MHSOs, the trend in all cause hospitalization was downward (not displayed). The direction of change is generally observed for both the intervention and comparison populations. The same general patterns hold for ER visits as well.

Table 5-4 displays the rates of hospitalization and emergency room (ER) visits per 1,000 intervention refresh beneficiaries for a 6-month period prior to launch of each MHSO's refresh efforts for all beneficiaries and by two clinical cohorts: heart failure and diabetes, and heart failure only. Rates of hospitalization are for all-cause and then for threshold condition-specific principal diagnosis. Rates of ER visits are for all-cause only. Below the utilization rates are the difference-in-differences rates of change observed between the baseline period and Months 1-6 of the refresh period¹⁸.

Across the 74 comparisons displayed in **Table 5-4** there is only one statistically significant difference in rates of change in hospitalizations or ER visits between the intervention and comparison groups. For McKesson, its rate of all cause hospitalization for heart failure only intervention beneficiaries declined substantially more than its comparison group's rate during Months 1-6 of the pilot. During our second site visit, the MHSOs reported that they modified their original approach to more quickly engage the refresh beneficiaries and begin intervention services based on lessons learned from their engagement of the original populations. Since the sizes of the refresh populations were much smaller than the original populations, the MHSOs could streamline operations and handle the refresh outreach process more efficiently. The quicker engagement of the refresh populations has not yielded any measurable impact yet.

¹⁸ We evaluate Months 1-5 for LifeMasters.

Table 5-4
Rate of acute care utilization per 1,000 Medicare Health Support intervention beneficiaries during a 6-month period prior to launch of the MHS pilot and rate of change during months 1-6 of the pilot: Intervention vs. Comparison/Pilot vs. Baseline (Difference-in-Differences); Refresh Populations

	Aetna	Healthways	Health Dialog	Green Ribbon Health	LifeMasters	McKesson	XLHealth
	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000	Rate/1,000
All Beneficiaries							
All Cause Hospitalization							
6-month rate prior to go-live	688	632	460	414	563	549	398
D-in-D rate during months 1-6	-24.5	-29.2	-4.3	18.0	26.2	-53.6	-5.9
Heart Failure Hospitalization							
6-month rate prior to go-live	193	180	81	109	123	156	72
D-in-D rate during months 1-6	10.3	-8.3	-10.0	-0.8	-2.1	-12.3	7.3
Diabetes Hospitalization							
6-month rate prior to go-live	29	12	8	5	5	7	8
D-in-D rate during months 1-6	-11.6	1.2	2.1	0.8	0.5	0.5	1.4
All Cause ER Visit							
6-month rate prior to go-live	424	568	470	392	687	898	558
D-in-D rate during months 1-6	-36.2	77.8	23.5	42.0	62.9	89.1	85.8
Heart Failure and Diabetes Beneficiaries							
All Cause Hospitalization							
6-month rate prior to go-live	698	669	686	459	653	615	515
D-in-D rate during months 1-6	-8.1	-35.9	78.7	76.8	17.1	-6.4	-26.3
Heart Failure Hospitalization							
6-month rate prior to go-live	199	201	214	131	161	207	112
D-in-D rate during months 1-6	10.0	-15.8	4.6	31.1	11.7	17.2	-9.2
Diabetes Hospitalization							
6-month rate prior to go-live	31	18	20	13	13	16	18
D-in-D rate during months 1-6	-12.3	1.7	10.1	0.6	0.3	5.6	1.9
All Cause ER Visit							
6-month rate prior to go-live	429	604	524	386	756	1021	646
D-in-D rate during months 1-6	-42.5	89.4	-70.6	14.1	71.8	107.0	39.9
Heart Failure Only Beneficiaries							
All Cause Hospitalization							
6-month rate prior to go-live	531	546	488	390	513	540	399
D-in-D rate during months 1-6	n/r ²	-14.2	-66.6	-14.4	31.5	-79.0 *	-9.1
Heart Failure Hospitalization							
6-month rate prior to go-live	99	131	102	97	102	147	80
D-in-D rate during months 1-6	n/r	9.3	-35.0	-18.1	-9.5	-27.4	16.8
All Cause ER Visit							
6-month rate prior to go-live	341	482	500	395	649	858	547
D-in-D rate during months 1-6	n/r	53.0	-21.5	56.7	57.8	58.8	62.1

NOTES:

1. Statistical significance testing of the difference-in-differences rate is conducted using STATA and a negative binomial regression model with robust variance estimation.
2. Results are not reported (n/r) for Aetna for heart failure only beneficiaries due to small numbers; only 252 beneficiaries are in the intervention heart failure only refresh population.
3. CIGNA Health Support did not request a refresh population.
4. We evaluated Months 1-5 for LifeMasters.

SOURCE: RTI Analysis of Medicare Health Support refresh populations' rate of acute care utilization for months 1 to 6 of their initial Phase I pilot experience compared to the same 6-month period during the year prior to each program's launch using Medicare Part A claims data 2006 – 2007 and the MHS daily eligibility file.

Refresh Population Mortality. Mortality rates during the first 6-month pilot period are displayed in **Table 5-5**. Mortality rates during the first 6-months of the refresh pilot period ranged between 6 and 8%. We do not observe any statistically significant differences in rates of mortality between intervention and comparison populations.

Table 5-5
Comparison of refresh intervention and comparison group mortality rates during the first 6-month Medicare Health Support Phase I pilot period

	Mortality rate		
	Intervention	Comparison	Difference
	(%)	(%)	
Aetna	6.9	6.9	0.0
Healthways	6.6	6.9	-0.3
Health Dialog	6.0	6.0	0.0
Green Ribbon Health	7.2	6.6	0.6
LifeMasters	7.1	6.3	0.8
McKesson	6.9	8.0	-1.1
XLHealth	6.2	5.7	0.5

NOTES:

1. Statistical significance testing of differences in the refresh populations’ mortality rates between intervention and comparison beneficiaries is conducted using a t-test. * p<.05
2. CIGNA Health Support did not request a refresh population.
3. We evaluated Months 1-5 for LifeMasters.

SOURCE: RTI Analysis of Medicare Health Support refresh populations’ mortality using the Medicare Enrollment Database and the MHS daily eligibility file.

5.4 Conclusion

The Medicare Health Support (MHS) authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I. In this chapter, we present analyses of changes in the rate of receipt of evidence-based process-of-care measures; changes in the rate of hospitalizations, re-admissions, and ER visits; and changes in mortality rates. We present *interim* results for an 18-month period for the original populations and *preliminary* selected 6-month results for the refresh populations.

Across 40 quality of care measures (five measures for each of the eight MHSOs), there was modest improvement in 16 (or 40%) measures for the original populations. Seven of the 8

MHSOs demonstrated at least one positive intervention effect. Healthways demonstrated a positive intervention effect across all five process-of-care measures and CHS across four of the five measures. LifeMasters improved cholesterol screening rates among beneficiaries with heart failure as well as diabetes but demonstrated no other positive intervention effects. Aetna and GRH improved the rate of cholesterol screening for beneficiaries with diabetes, and McKesson improved the rates of cholesterol screening and HbA1c testing in beneficiaries in diabetes. Health Dialog demonstrated a positive intervention effect related to cholesterol screening for beneficiaries with heart failure but did not demonstrate any intervention effects related to quality of care for beneficiaries with diabetes. XLHealth did not demonstrate any positive intervention effects on quality of care.

Rates of improvement in the quality of care measures were relatively modest; 2 to 4 percentage points. The MHSOs were most successful improving cholesterol screening. Examination of the underlying trends in rates of change in the comparison populations show a very clear pattern of declining rates of cholesterol screening and HbA1c testing over time suggesting a possible ceiling effect in the 70 to 80% range for a chronically ill population. The positive intervention effects observed for these two measures were primarily driven by the MHSOs mitigating the observed downward trend in the comparison groups or modestly improving the intervention groups' rates. In contrast, we observe substantial increases in rates of retinal eye examination across the comparison groups, generally a 20 percentage point improvement, over relatively low baseline rates, and, modest increases in rates of urine protein screening over baseline rates that generally were lower than those observed for cholesterol screening and HbA1c testing. Only Healthways was successful at outperforming the comparison populations' improvement in both of these measures and CHS was successful having a positive intervention effect for one of these two measures.

For both the original and refresh populations, none of the 8 MHSOs demonstrated positive intervention effects related to health outcomes. Across the 120 comparisons for the original populations, there were no statistically significant reductions in the rate of growth in hospitalizations, re-admissions, or ER visits in the intervention groups relative to the comparison groups. Nor do we observe any statistically significant reductions in rates of hospitalization or ER visits across 74 comparisons for the refresh populations. We observe no reduction in mortality rates or time to death during the first 18 months of the pilot for the original populations and the first 6 months of engagement of the refresh populations.

CHAPTER 6

INTERIM FINANCIAL OUTCOMES

The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (Pub. L. 108–173) authorizing the Medicare Health Support (MHS) pilot program called for an independent evaluation of financial outcomes, or savings, along with improvements in beneficiary clinical quality of care and satisfaction with their health care. Interim, 18-month findings based on RTI’s evaluation are intended to provide input to CMS’ recommendation to the Secretary regarding the Phase II expansion of the pilot for successful program(s) or components of programs. To meet the conditions for expansion, MHS Organizations (MHSOs) must achieve targets for savings subject to budget neutrality. That is, they must lower Medicare expenditures on health services by an amount equal to the fees that the MHSOs have accrued for managing their assigned beneficiaries.

RTI’s findings are based on the experience of approximately 290,000 chronically ill Medicare beneficiaries (approximately 30,000 in each of 8 MHSOs’ original populations and between 4,000 and 8,000 in each of 7 MHSOs’ refresh populations) randomized to an intervention or a comparison group in eight geographic areas. To date, this is the largest randomized experiment ever conducted of population-based care management.

In this chapter, we present evaluation findings for the original populations on levels and trends in Medicare payments for the year prior to the start date and over the first 18 months for each of 8 Medicare Health Support Organizations. We also present preliminary 6-month results for the “refresh” populations that were offered by CMS to all 8 MHSOs and accepted by 7 of the organizations to partially offset attrition primarily due to death during the first year. First, however, we summarize the payment arrangements negotiated with each of the MHSOs.

6.1 MHS Pilot Payment Arrangements

In the MHS pilot, each MHSO receives from CMS a negotiated monthly administrative fee per participant. Fees are at risk for performance, including

- budget neutrality with respect to fees
- improvement in clinical quality and beneficiary satisfaction.

During the first 6 months of each cohort, the MHSOs received a monthly management fee for each beneficiary in their assigned intervention group until such time that the beneficiary became ineligible or declined to participate. Beyond the initial 6-month period, management fees are paid only for confirmed participants, and only for eligible pilot periods. MHSOs are held at risk for fees based on the performance of the full population of beneficiaries randomized to the intervention group (an intent-to-treat [ITT] model) compared with beneficiaries randomized to the comparison group. To keep all their management fees, MHSOs must reduce Medicare payments for the intervention group by the amount of accrued fees, i.e., achieve budget neutrality. To the extent that the MHSOs do not fully engage their assigned populations, the percentage savings on those that they do actively manage (the participants) must be even greater for them to be financially successful. This assumes no impact on Medicare expenditures among those that never participated. CMS designed the MHS initiative to encourage participating

organizations to actively engage as many beneficiaries as possible—especially those in greatest need of care management – and to intervene as they perceived the need. CMS also required MHSOs to put a portion of their fees at risk for several clinical processes of care and one patient satisfaction measure.

The financial results presented in this chapter are based on RTI’s analysis of financial outcomes and does not assess the financial implications of the quality and satisfaction measures incorporated in each MHSO’s performance standards, which is the responsibility of another CMS contractor. Furthermore, RTI’s evaluative approach differs from the CMS financial reconciliation protocol negotiated between CMS and the MHSOs. The evaluator’s responsibility includes assessing financial outcomes, including any cost savings to the program. This assessment necessarily requires hypothesis testing using various statistical methods, as described in the next section.

6.2 Financial Analysis Data and Methods

Data. Data for the financial analysis is based on the approximately 240,000 original beneficiaries and approximately 47,000 refresh beneficiaries. RTI extracted Medicare claims for all eligible beneficiaries in the 8 MHSOs between 2004 and 2007. Any claims in periods in which a beneficiary was ineligible (e.g., joined a Medicare Advantage plan, entered the ESRD program, dropped Part B coverage) were deleted as were a few beneficiaries (less than 1-in-10,000) with no claims in the base year or intervention period. Beneficiaries who died during the pilot period had their spending averaged over eligible days prior to death.

In our first analysis of 6-month impacts, RTI developed its own base year claims database to calculate changes in spending at the beneficiary level using monthly indicators of Part A and B fee-for-service (FFS) eligibility. At the time of this set of analyses, CMS has produced a baseline daily eligibility file for each MHSO. For this Report to Congress, we reconstructed the 12-month baseline claims file using the daily eligibility file rather than the CMS Denominator file. We also changed the selection algorithm for home health claims for three of the MHSOs; we use *from* date rather than *start* date¹⁹. Baseline PBPM expenditures were then re-calculated using the new baseline data file and beneficiary level weights reflecting the proportion of time each beneficiary was eligible during the first 18 months of the pilot.

Analytic Approach. Medicare payments for health care services are based on claims for services during the 18-month²⁰ pilot period and for the 12-month base year prior to each MHSO’s start date for the original populations. We also analyzed claims for a preliminary 6-month period for the “refresh” populations along with a corresponding 6-month period prior to each MHSO’s start date. The prior year’s claim file has a longer “run-out,” and therefore is more complete than the 6- or 18-month pilot periods. This should not bias RTI’s comparisons of

¹⁹ During the analysis phase, CMS modified its method of selecting eligible home health claims, changing the selection of claims using **from date** rather than **start date**. RTI received revised intervention-period utilization files for three of the MHSOs (McKesson, LifeMasters, and XLHealth) within the time period of our 18-month analyses. We modified the baseline files for these MHSOs to also select claims based upon the **from date**. For the remaining five MHSOs, baseline and intervention claims files were constructed using **start date**.

²⁰ For LifeMasters, the pilot period is 17 months due to their requested early termination.

changes in spending between the two large, randomly determined, intervention and comparison populations.

RTI's analytic approach is based on a *comparison of growth rates in PBPMs 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 PBPMs between the base year and the pilot period that are not accounted for through randomization.
- Second, by calculating changes in PBPMs at the beneficiary level (i.e., “paired” base-pilot period PBPMs), we can conduct statistical t-tests of the differences in spending growth rates between intervention and comparison groups.

In addition to answering the Congressional question whether any or all of the eight MHS pilot programs or components of programs have achieved budget neutrality (or even any savings), we also are interested in *generalizing* MHS results to future program expansions by answering the question: “What savings are likely to be realized if the pilot 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 due to chance with no long-run implications.

Constructing PBPMs. In conducting the analyses, RTI first had to develop base and pilot period spending measures. Average per beneficiary per month (PBPM) spending was derived by summing all eligible Medicare claims payments for services used by a beneficiary divided by the number of MHS eligible months. Eligible months, in turn, was calculated as the number of beneficiary eligible days divided by 30.42. Next, we calculated beneficiary-specific *changes* in average PBPMs between 12 months (6 months for refresh) prior and 18 months (6 months for refresh) during the pilot period. RTI then weighted each beneficiary's change in PBPM by the fraction of days they were eligible during the 18-month pilot period so as not to overstate the importance of beneficiaries eligible for the intervention for very short time periods. This method effectively weights a beneficiary's pilot and base period PBPM equally. We chose not to trim outlier spending. We assumed all spending in the claims files represented true program costs.

Overall PBPM Analyses. In the results Section 6.3, we begin by presenting overall *changes* in intervention PBPMs between the base year and the first 18 months of the pilot relative to changes in the comparison group PBPMs using a pre/post, intervention/comparison, difference-in-differences, analytic design. We also provide estimates of “statistically detectable” savings as a percent of the comparison group's PBPM. An upper 5% confidence interval threshold is used for this purpose. Statistically detectable thresholds illustrate the power we have to detect small, but likely true, savings given the large sample sizes at the MHSO level. Thresholds are critical as a benchmark when evaluating the budget neutrality savings criterion as well as the level of savings that could be expected in a Phase II expansion of the program.

We also stratify trends in PBPMs by five disease groups: (1) heart failure (HF)-only, (2) diabetes-only, (3) HF with or without diabetes, (4) diabetes with or without HF, and (5) HF and diabetes. This provides 40 additional savings tests of intervention effects (8 MHSOs, 5 strata) for specific disease subgroups, albeit at somewhat reduced sample sizes and statistical power. The

MHSOs developed components of their programs to focus upon disease-specific cohorts of beneficiaries (e.g., telemonitoring for beneficiaries with heart failure)²¹. Because we do not have detailed intervention data – that go beyond number of visits or telephone contacts - we wanted to indirectly evaluate whether program savings occurred for components of the MHS programs and whether program savings occurred for beneficiaries with only a single disease or multiple comorbid conditions.

ANCOVA Regressions. Besides conducting t-tests of mean differences in changes in PBPMs, RTI also performed analyses of covariance (ANCOVA) on beneficiary PBPMs using ANCOVA regression models. We controlled for base year beneficiary characteristics, such as beneficiary age, gender, threshold chronic condition, and two comorbidity measures, the Hierarchical Condition Categories (HCC) risk score and the Charlson comorbidity index. While randomization eliminated practically all of the between-group differences in beneficiary characteristics, some residual differences might remain. Statistically controlling for beneficiary characteristics not only tests for any possible biases remaining in our paired t-tests but also provides a more precise estimate of intervention effects on Medicare payments. Regression modeling also provides an overall average regression-to-the-mean effect from base year to the pilot years. It is possible that high-cost beneficiaries in the base year become less costly in the pilot period (and low-cost beneficiaries more costly). While regression-to-the-mean effects should be similar between intervention and comparison groups, by using multivariate regression methods we are able to control for any random differences that might remain between the two groups. Our “paired” t-tests of PBPM changes capture regression-to-the-mean effects as well but do not provide any independent measure of its importance for a chronically ill population more generally.²²

Participant Effects. We next compare pilot period PBPM spending growth rates for participants and never participants, separately, with the entire comparison group. For policy makers, it is important to know if MHSOs are having more success with beneficiaries who are actually participating in the intervention. This may help target beneficiary populations for future programs.

Budget Neutrality. Given the interest in how successful MHSOs have been in achieving budget neutrality halfway through the pilot period, we present an analysis of progress toward budget neutrality based on our comparison of PBPM growth rates.²³ If an MHSO achieves an intervention PBPM after 18 months that is 2% lower than its comparison group’s PBPM, and its

²¹ We also evaluated these five cohorts for changes in quality of care and health outcomes. We did not report those results as they did not differ substantively from the reported results.

²² Disease management contracts often employ a simpler, pre/post, measure of intervention success. Without a matched control group, regression-to-the-mean effects could be explaining some of their apparent intervention success. ANCOVA analysis provides a test of the bias that might be inherent in simpler experimental designs.

²³ CMS’ financial reconciliator is responsible for the final determination of budget neutrality for each MHSO and uses a somewhat different methodology, e.g., trimming outliers, adjusting base year PBPMs using an aggregate actuarial factor as opposed to using each beneficiary’s own base year PBPM as is necessary in RTI’s methodology to allow for statistical testing at the beneficiary level.

monthly fee is 8% of the comparison group's PBPM, then the MHSO is roughly one-quarter of the way to meeting budget neutrality with one-half of the time left in the pilot period. The MHSO would have to triple its rate of savings in the last 18 months to achieve overall budget neutrality.

Average monthly fees were estimated as a weighted average of negotiated fees through the pilot's first 18 months; weights were the proportion of eligible months that fees were in effect across all intervention beneficiaries during the pilot period. Fees were based on those negotiated between MHSOs and CMS and reflect any changes made through the 18-month analysis period.²⁴ In several MHSOs, a different fee(s) was (were) applicable to the post-6-month period, in which case we use a month-weighted average of the two (or more) fees for full 18-month period. We down-weighted the applicable fee for months 7-18 to account for the fact that fees were accrued only on participating beneficiaries after six months. A 1.5% monthly attrition factor was also applied to the monthly fee weights to account for deaths and attrition for other reasons.

Refresh Analyses. Finally, we conclude with a preliminary analysis of financial performance for the 7 MHSOs that accepted CMS's offer of a "refresh" population. The analysis covers only a 6-month pilot period for refresh beneficiaries relative to a comparable 6-month base period prior to the go-live of the refresh populations. At 13% to 33% the size of the original populations, the precision of the refresh population estimates is slightly less than 50% of those in the original populations. Consequently, fewer statistically significant differences are expected in the refresh populations.

6.3 Results for the Original Population

Overall Spending Differences. In *Table 6-1* and *Figure 6-1*, we present differences in PBPM growth rates for each of the 8 MHSOs for their original populations. The first two columns give average PBPM growth rates for the intervention and comparison beneficiaries. Both groups show positive spending growth in all MHSOs that was significant at the 95% confidence level. Column (3) shows the differences in growth rates between the two groups. It is our test of intervention cost savings. Four MHSOs exhibited faster PBPM growth rates in the intervention group (led by LifeMasters at +\$38) and four exhibited slower rates (led by XLH at -\$29). None of the MHSOs showed either statistically slower or faster rates of growth in intervention PBPMs relative to their comparison group. Rates of growth in PBPMs (see column 4) varied from -2.1% of the comparison group PBPM (XLHealth) to a positive 2.7% (LifeMasters).

²⁴ RTI did not factor in a reduction in the negotiated fee requested by one MHSO after the 18-month period. The MHSO requested that its fee be reduced to \$0. Incorporating such a reduction would have had this MHSO appearing more successful on the budget neutrality criterion than they actually were at the mid-point of the pilot.

Table 6-1

Intervention/Comparison differences in Medicare Health Support per-beneficiary-per-month (PBPM) growth rates between 18-month Phase I pilot period and 12-month base year period, by Medicare Health Support Organization(MHSO) for original populations

MHSO	Intervention ¹	Comparison ¹	Difference-in-growth rates ²	Growth rate differences as % of comparison PBPM	95% CI ³	Detectable % of comparison PBPM ⁴
Aetna	\$319	\$345	-\$26	-1.5%	\$66	3.9%
Healthways	331	305	+26	+1.6	73	4.5
CIGNA Health Support	218	230	-13	-1.0	53	4.1
Health Dialog	253	228	+26	+1.9	61	4.3
Green Ribbon Health	295	312	-17	-1.2	48	3.4
Life Masters	260	222	+38	+2.7	54	3.8
McKesson	235	234	+1	+0.0	52	3.8
XLHealth	302	331	-29	-2.1	52	3.7

NOTES: *p<.05; **p<.01.

¹ Per beneficiary per month (PBPM) payments in the first 18 pilot months minus payments for same beneficiaries in 12 months prior to MHSO’s start date. Differences weighted by beneficiary’s fraction of eligible days in the 18-month pilot period. No outlier adjustments have been made.

² Column 1 minus column 2. Pairwise t-test of differences using 5% confidence level.

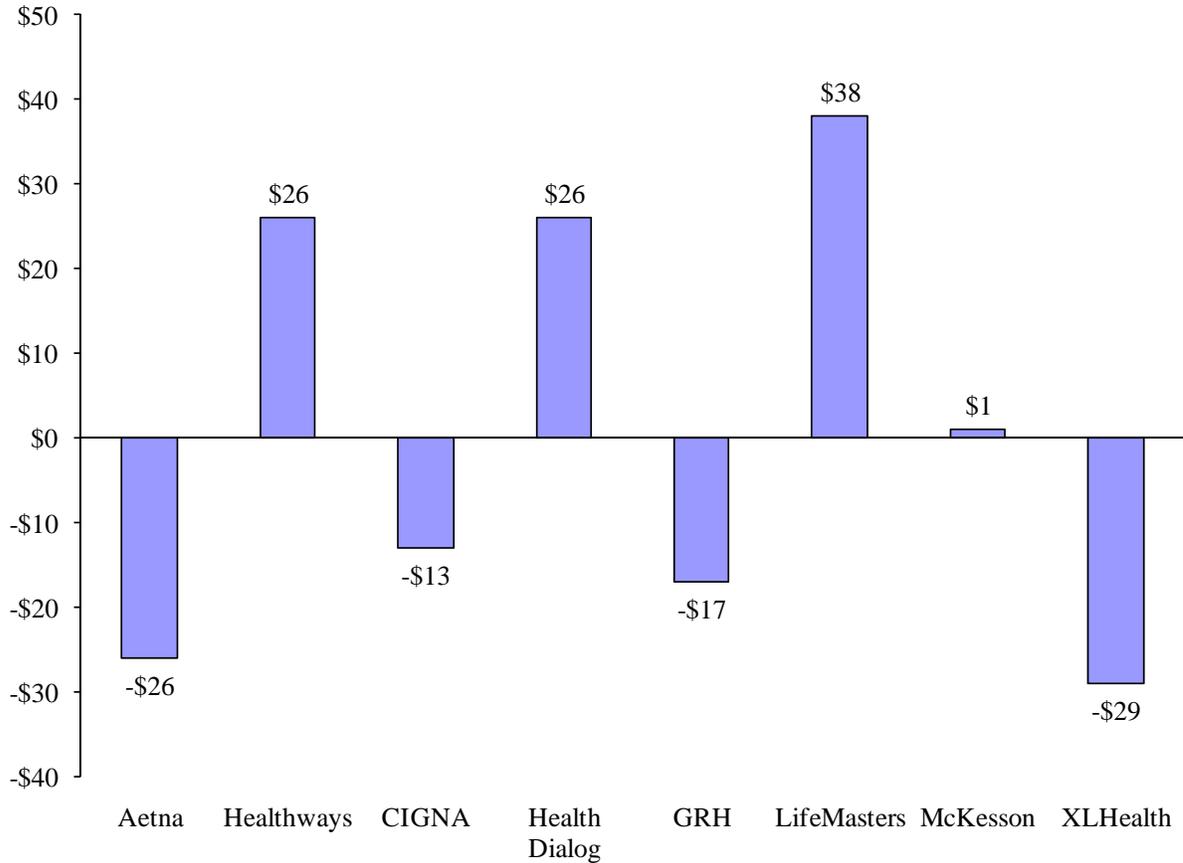
³ 95% upper confidence threshold (CI) = 1.96*standard error of differences in intervention & comparison PBPM means.

⁴ 95% CI divided by 18-month comparison PBPM.

SOURCE: RTI Analysis of Medicare Health Support original populations’ Medicare expenditures using Medicare Part A and B claims data 2004 – 2007 and MHS daily eligibility file.

Column (5) reports the upper 95% confidence threshold for the differences in growth rates. Differences in column 3 larger than the threshold would be considered statistically different from zero implying real Medicare savings. Only 4-of-8 MHSOs achieved lower PBPM growth, but the rates were well below the threshold to be considered statistically significant. Column (6) reports the 95% upper threshold as a percent of the comparison group PBPM. This column indicates how small a difference in growth rates we could confidently report as truly meaningful, expressed as a “detectable” percent of the comparison group’s PBPM. Smaller percentages imply greater accuracy in measuring the true intervention effect. Thresholds can be higher in MHSOs with greater variation in beneficiary PBPMs. According to column 6, through the first half of the Phase I pilot we are able to detect differences in intervention and comparison group growth rates as small as 3.4% (GRH) to 4.5% (Healthways) of the comparison group PBPM. In the original agreements between CMS and the MHSOs, they had to achieve 5% gross savings first before any fees could be retained. Consequently, if an MHSO had achieved 5% gross savings over the first 18 months, RTI would have reported it as a statistically significant finding. XLHealth’s intervention had the largest negative growth effect through 18 months relative to its comparison group (-2.1%); yet, this percentage was well below the 3.7%, or \$52, difference required for statistical significance (and well below the original 5% savings criterion).

Figure 6-1
Intervention/Comparison differences in Medicare Health Support per-beneficiary-per-month (PBPM) growth rates¹ between 18-month Phase I pilot period and 12-month base year period, by Medicare Health Support Organization for original populations



NOTE:

¹ Average change in intervention PBPMs between 18-month pilot and 12-month base period prior to start date minus averages change in comparison group. Weighted by beneficiary pilot fraction of eligible days. No outlier trims.

SOURCE: Based on Table 6-1 data and technical notes.

Disease Groups. We also stratified trends in MHSO PBPMs by five disease groups: (1) heart failure (HF) only, (2) diabetes only, (3) HF with or without diabetes, (4) diabetes with or without HF, and (5) HF and diabetes.²⁵ No pattern was found within any of the five disease groups in any of the eight MHSOs that might imply successful targeting of intervention efforts. Twenty-one of 40 comparisons showed intervention PBPMs trending at a slower rate but 19 PBPMs trending at a faster rate. Of the 40 statistical tests we conducted of differential growth rates by the five disease groups, only one was statistically significant. XLHealth’s intervention heart failure-only PBPM increased by \$124 less than the comparison group (\$381 v. \$257).

Beneficiary Characteristics and Regression-to-the-Mean. *Table 6-2* summarizes the key results of multivariate comparisons of intervention and comparison group PBPMs between the base year and 18-month pilot period. Column 1 reproduces the differences in PBPM growth rates shown in column 3 of *Table 6-1*. Using regression techniques, the second column tests the robustness of the findings after controlling for each beneficiary’s base year demographic and clinical characteristics. The difference in growth rates are practically identical, implying that intervention and comparison group beneficiaries are equivalent on the key characteristics that affect their costliness to the Medicare program.²⁶

Column 3 adds each beneficiary’s base year average PBPM to their demographic and clinical characteristics. The coefficients reflect differences in growth rates adjusting for any systematic differences in base year PBPMs between intervention and comparison beneficiaries. Adjusting as well for base year PBPM does not change the previous finding of no statistically significant differences in intervention and comparison group growth rates.

Regression-to-the-mean effects in column 4 are highly significant and vary only marginally across the 8 MHSOs. This is to be expected as the phenomenon affects practically all beneficiaries regardless of which group they are in. One way to interpret the effects is to predict the average intervention period PBPM for two beneficiaries, one costing \$500 per month during the base year and another costing \$2,500, or 5-times as much. Based on predicted PBPMs during the 18-month pilot, the cost difference between these two beneficiaries would narrow to 1.5-1.7 after the base year, holding beneficiary characteristics and intervention effects constant. Such large regression-to-the-mean effects imply that many high-cost chronically ill beneficiaries cost Medicare far less in a subsequent period—and vice-versa for low-cost beneficiaries. This reinforces the need for a scientifically matched or randomized comparison group in benchmarking intervention performance in lieu of a simpler pre/post design; otherwise, the natural trend to lower costs among high-cost beneficiaries could be misinterpreted as intervention success.

²⁵ A sixth group of beneficiaries diagnosed with only one claim for heart failure or diabetes was too small for rigorous statistical testing (comparison group samples were usually less than 100 per MHSO). In this sixth group, 4-of-8 MHSOs exhibited slower PBPM growth relative to their comparison group and another 4 had positive growth. Including them in the overall analysis of PBPM growth rates changed intervention growth rates by less than one percent (e.g., from -\$26 to -\$25.8).

²⁶ Beneficiary age, comorbid risk scores, and gender were always highly significant positive predictors of a beneficiary’s PBPM—especially the HCC score. That controlling for them has no effect on the intervention-comparison group differences implies that beneficiary characteristics do not vary systematically in the two groups.

Table 6-2
Differences in Medicare Health Support per beneficiary per month (PBPM) growth rates, adjusted for beneficiary characteristics and regression-to-the-mean effects

MHSO	Intervention - Comparison PBPM growth rates unadjusted	Intervention - Comparison PBPM growth rates adjusted for beneficiary characteristics	Intervention - Comparison PBPM growth rates adjusted for beneficiary characteristics & regression-to-mean	Regression-to-mean effect
Aetna	-\$26	-\$26	-\$2	0.42 **
Healthways	26	26	-18	0.35 **
CIGNA Health Support	-13	-12	17	0.36 **
Health Dialog	26	29	-19	0.34 **
Green Ribbon Health	-17	-17	14	0.38 **
LifeMasters	38	38	-37	0.38 **
McKesson	1	1	3	0.40 **
XLHealth	-29	-29	35	0.39 **

NOTE: **p< .01

Column 1: Intervention impact coefficient in ANCOVA regression of pooled base year and 18-month intervention PBPMs for roughly 30,000 beneficiaries per MHSO. Equivalent to column 3, Table 6-1.

Columns 2&3: Intervention impact coefficient controlling, first, for beneficiary characteristics, then for regression-to-the-mean based on base year PBPMs.

Column 4: ANCOVA base year PBPM coefficient.

SOURCE: RTI Analysis of Medicare Health Support original populations' Medicare expenditures using Medicare Part A and B claims data 2004 – 2007 and MHS daily eligibility file.

Participants & Never Participants. *Table 6-3* compares differences in PBPMs between two sub-groups of intervention beneficiaries, participants and never participants, and the entire comparison group. The first two columns show differences in PBPMs between participants and comparison beneficiaries in the base and pilot periods. The third column gives the change in the differences between participant and comparison PBPMs that occurred during the pilot period. The last three columns show differences and changes in differences over time for the never participants who are still part of the intervention group. For example, Aetna's PBPM for participants in the base year was \$16 greater, on average, than for the entire comparison group. Over the pilot period, the positive difference became a negative difference of \$31, implying that Aetna's participant PBPM grew \$47 slower compared with the entire comparison group. By contrast, Aetna's never participant group started out \$250 greater than the comparison group, a gap that increased to \$355, implying a \$105 faster increase relative to the comparison group.

All but McKesson engaged in the intervention beneficiaries (i.e., participants) who were financially similar to comparison group beneficiaries in general, as evidenced by the

insignificant PBPM differences in column 1²⁷. However, 5-of-8 MHSOs were also financially responsible for never participants who averaged a statistically significant \$86-\$250 more costly prior to the intervention (column 4). Columns 3 and 6 summarize growth rates in the two intervention sub-groups benchmarked against the comparison group. None of the 8 MHSOs experienced participant PBPM growth that was statistically less than their full comparison group. By contrast, 3-of-8 MHSOs had never participant group growth in PBPMs significantly in excess of their comparison group (last column, *Table 6-3*).

Table 6-3
Comparison of PBPM differences between participants, never participants, & comparison beneficiaries in the 12-month base year and 18-month pilot period, by Medicare Health Support Organization (MHSO), original populations

MHSO	Participants minus comparison			Never participants minus comparison		
	Base year ¹	18-months ¹	Difference ²	Base year ¹	18-months ¹	Difference ²
Aetna	\$16	-\$31	-\$47	\$250 **	\$355 **	\$105
Healthways	-15	-12	3	26	272 **	246 **
CIGNA Health Support	-16	-39	-23	86 *	161 **	75
Health Dialog	-18	1	19	189 *	515 **	326 **
Green Ribbon Health	-11	-43	-32	125 **	207 **	82
LifeMasters	-17	0	17	59	177 **	118 **
McKesson	-38 *	-49 *	-11	168 **	231 **	63
XLHealth	-16	-67 **	-51	20	66	46

NOTE: *p<.05; **p<.01

¹ Numbers in column represent the difference in average PBPMs in a given period between intervention group and comparison group. Beneficiary PBPMs weighted by fraction of eligible days in 18-month pilot period. No outlier adjustments.

² Difference represents the change in the difference in PBPMs between the base and 18-month pilot periods.

SOURCE: RTI Analysis of Medicare Health Support original populations' Medicare expenditures using Medicare Part A and B claims data 2004 – 2007 and MHS daily eligibility file.

²⁷ In the financial savings analyses, beneficiary PBPMs are weighted by the fraction of eligible days in the 18-month pilot period. When doing so, the observed differences in PBPMs between participants and never participants are substantially smaller and no longer statistically significant with the exception of McKesson. Thus, weighting by pilot period eligibility reduces the potential bias in the financial savings analyses from lesser regression-to-the mean effect in the participant groups than we observe in the full comparison groups with higher PBPMs at baseline.

6.4 Monthly Fee Budget Neutrality

Table 6-4 compares the difference in PBPM growth rates of Medicare spending with the estimate of each MHSO's average monthly fee as a percent of the comparison group PBPM through the first 18-months of the pilot program. Success in lowering Medicare claims costs is required if MHSOs are to be able to retain their accrued fees.

Table 6-4
Success in achieving monthly fee budget neutrality through first 18 pilot months, by Medicare Health Support Organization (MHSO), original populations

MHSO	% Difference intervention/comparison PBPMs	Average Monthly Fee (\$)	Monthly fee % of comparison PBPM	Monthly fee % net of PBPM differences	% Monthly fee retained
Aetna	-1.5	118	6.9	5.4	22
Healthways	1.6	94	5.9	7.5	-27
CIGNA Health Support	-1.0	104	8.2	7.2	12
Health Dialog	1.9	104	7.5	9.4	-25
Green Ribbon Health	-1.2	67	4.7	3.5	26
LifeMasters	2.7	76	5.4	8.1	-50
McKesson	0.0	114	8.4	8.4	0
XLHealth	-2.1	131	9.3	7.2	23

NOTE:

Column 1: Difference in 18-month intervention/comparison PBPM growth rates, taken from column 4, Table 6-1. PBPMs weighted by fraction of eligible days in 18-month pilot.

Column 2: Weighted average of negotiated fees during pilot's first 18 months; weights are proportion of 18 months fees in effect.

Column 3: Average monthly fee as percent of 18-month comparison PBPM. Fees after first 6 months weighted by participation rate and adjusted for attrition.

Column 4: Column 3 plus column 1.

Column 5: Column 4 divided by column 3 minus 1 times 100.

SOURCE: RTI Analysis of Medicare Health Support original populations' Medicare expenditures using Medicare Part A and B claims data 2004 – 2007 and MHS daily eligibility file.

Column 3 expresses the estimated average effective monthly fee accrued by MHSOs as a percent of their comparison group PBPM through 18 months. Monthly fees are accrued on the full intervention population during the 6-month outreach period, then only on participants. This figure adjusts for attrition and participation rates to express the estimate of what the monthly fee would be across all eligible person-months for the intervention group as a percent of the comparison group PBPM Medicare claims costs. Percentages represent the budget neutral level of financial performance required of MHSOs. Fee percentages range from a low of 4.7% (Green Ribbon Health) to a high of 9.3% (XLHealth). MHSOs would need to attain at least this level of savings to reach budget neutrality.

Column 4 adjusts the fee percentages in column 3 for any gross savings in Medicare claims costs in column 1, thereby producing net monthly fee costs to Medicare. Net fee costs range from a low of 3.5% (GRH) to a high of 9.4% (Health Dialog). A zero or negative percentage in column 4 would imply that an MHSO had achieved budget neutrality (or better) halfway through the three year pilot period. All percentages are positive, however.

Column 5 estimates the amount of fees that MHSOs had offset through decreased Medicare claims costs as of month 18 of the pilot.²⁸ Positive percentages imply fees that have been “covered,” through lower Medicare claims costs. Four-of-eight MHSOs exhibit positive percentages ranging from 12% (CHS) to 26% (GRH). These 4 MHSOs, assuming that gross savings through 18 months were an accurate indicator of intervention impacts, would have to increase their rate of savings by 2.6- to 7.2-fold over the second half of the pilot period to achieve overall budget neutrality. Four other MHSOs, because they have higher, not lower, claims costs relative to the comparison group, would have to save more than 100% of their fee costs in lower Medicare health care outlays with just half of the pilot period remaining.

6.5 Results for the Refresh Populations

Table 6-5 summarizes the differences in intervention and comparison group PBPM growth rates for the refresh populations. Between the base year and the 6-month pilot period, 4-of-7 MHSOs experienced statistically significant increases in both their intervention and comparison group PBPMs (columns 1 and 3). None of the 7 MHSOs experienced statistically slower PBPM *growth* in their intervention versus comparison group (see the Difference-in-growth rates column). Four MHSOs exhibited slower intervention PBPM growth while 3 exhibited faster growth than their comparison group. Aetna experienced the largest gross savings of the 7 MHSOs (-\$155; 6.3% of comparison PBPM), although this difference, like the other six differences, was not statistically significant. Whether slower PBPM growth of 4-6% becomes statistically significant depends upon trends in PBPM spending during the remaining pilot period for the refresh populations. It is possible that regression-to-the-mean effects over time might narrow the variation in beneficiary average PBPMs and provide a clearer understanding of the level of success with the refresh populations.

²⁸ Statistics in column 5 are only approximate. CMS’s financial reconciliation process will ultimately determine each MHSO’s refund obligation.

Table 6-5
Per-beneficiary-per-month (PBPM) Medicare payment differences between Medicare Health Support first 6-month pilot period and a comparable 6-month prior period, by Medicare Health Support Organization (MHSO), refresh populations

MHSO	Intervention		Comparison		Difference-in-growth rates ³	Growth rate differences as % of comparison PBPM ⁴
	Mean ¹	N ²	Mean ¹	N ²		
Aetna	\$368	** 4,283	\$524	** 2,142	-\$155	-6.3%
Healthways	328	** 4,245	375	** 2,113	-48	-2.2
Health Dialog	101	* 5,938	167	* 2,409	-66	-4.4
Green Ribbon Health	113	2,715	6	1,361	106	+7.5
LifeMasters	46	4,599	-71	2,310	117	+7.7
McKesson	36	4,557	84	2,270	-48	-3.2
XLHealth	259	** 5,187	177	** 2,609	82	+5.9

NOTES: * p<.05; ** p<.01

¹ Mean differences = change in per beneficiary per month (PBPM) payments between first 6 months of engagement and same 6 months prior to each Medicare Health Support Organization's start date. Differences weighted by beneficiary's eligible fraction of days in the 6-month pilot period. No outlier adjustments have been made.

² N= number of eligible beneficiaries at each MHSO refresh start date, excluding beneficiaries with zero costs at baseline.

³ Difference between intervention and comparison group paired PBPM growth rates; column 1 minus column 3.

⁴ Ratio of difference-in-growth rates to 6-month comparison group refresh PBPM. Negative values reflect lower claims cost increases in intervention group.

6.6 Summary of Findings

Subsection (b)(5) of the MMA legislation called for an independent evaluation of financial outcomes, including any cost savings, along with analyses of changes in clinical quality of care and beneficiary and provider satisfaction. Section 1807(c)(1) of the authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves the clinical quality of care, and improves beneficiary satisfaction, and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I.

Interim findings based on this report are intended to provide input to CMS' recommendation to the Secretary regarding the expansion of the pilot for successful program(s) or components of programs. Financial success under the original terms and conditions required MHSOs to save Medicare at least 5% on Medicare health care outlays, called gross savings; otherwise, organizations would have to return all of their monthly management fees. This would have assured a positive return on Medicare's investment. In December, 2007, CMS waived the 5% requirement with OMB approval; MHSOs are now only required to achieve gross savings equal to its management fees: the so-called budget neutrality criterion. Refunds are required for

savings less than fees, thereby assuring the pilot program will be at least budget neutral with respect to management fees.

RTI, as the independent evaluator, is responsible for evaluating the financial outcomes of MHSOs and informing policy makers on the likely future success that could be expected with program expansion. RTI's findings through 18 months, or halfway through the pilot, are based on the experience of approximately 240,000 chronically ill Medicare beneficiaries randomized to an intervention or comparison group. To date, this is the largest randomized experiment in disease management ever conducted. Key findings are the following.

1. **None of the 8 MHSOs achieved gross savings rates that were statistically different from zero. At least during the first half of the MHS pilot, RTI cannot assert that the MHSOs, individually or as a group, had any cost-saving impact on Medicare claims costs.** Lack of statistical success was not due to small sample sizes. RTI was able to detect savings rates well below 5% of monthly fees as a percent of the monthly PBPM. The fact that 4-of-8 MHSOs exhibited greater increases in spending than in the randomized comparison group reinforces our conclusion of no discernable effect of these MHSOs' interventions on Medicare claims costs.
2. **The lack of financial success was uniform across five broad disease groups. Of the 40 possible statistical tests, only one was significant.** Using a 95% two-sided confidence interval, we should expect to observe one statistically positive or negative difference simply at random that could disappear by the end of the pilot period. Moreover, no systematic pattern of success was found by disease group that could be indicative of intervention effects at least with one disease group. It is questionable, statistically, to focus on a single positive finding when the overwhelming pattern of results is contradictory.
3. **Controlling for beneficiary characteristics including beneficiary age, gender, type of chronic disease, comorbid risk factors, and fraction of time in the pilot had no impact on the lack of demonstrable savings.** As each of these factors were strong predictors of higher costs by themselves, the fact that they had no effect on the differences in intervention and comparison PBPM growth rates implies that they were equally distributed between the intervention and comparison groups.
4. **No evidence of savings was found among beneficiaries who agreed to participate in the intervention.** Evidence was found of engagement of less costly, healthier, beneficiaries into the participant pool. Intervention beneficiaries who never consented to participate were much more expensive in the base year and experienced higher rates of cost increases than the participant group. Nevertheless, participant cost increases, alone, were no different than for the entire comparison group. These findings were insensitive to MHSO differences in participation rates. Health Dialog led the MHSOs with a 95% participation rate within its original population and a 96% participation rate within its refresh population, yet it was one of the least successful MHSOs in slowing the growth in Medicare spending. Difficulty in engaging sicker,

more costly beneficiaries raises material questions about the future success of a broad, population-based, approach to Medicare chronic disease management.

5. **Savings one-half way through the Phase I pilot period have offset 12-26% of estimated accrued fees for 4 of 8 MHSOs. The remaining 4 MHSOs show no evidence of savings.** Halfway through the pilot, Medicare's return on investment, at best, is minus 74% and, at worse, greater than minus 100%. The final financial reconciliation of the amount of accrued fees each MHSO can retain will be computed using a methodology that incorporates additional steps, such as trimming outlier costs in the base and pilot periods, fees at risk for performance, and only applies baseline adjustment factors favorable to MHSOs. However the results presented here show fees accrued far exceed savings produced.
6. **With just one-half of the time remaining in the Phase I pilot period, the 4 MHSOs with modest savings would have to increase their rate of savings by roughly 3-to-7-fold while the second group of 4 "non-savers" would have to reduce Medicare spending on services more than the entire average monthly fee they have accrued.** This seems unlikely given performance over the first 18 months.
7. **Four-of-seven MHSOs exhibited slower intervention spending growth rates in their refresh populations after the first six months, but none of these savings were statistically significant.** Seven MHSOs accepted "refresh" populations beginning in the pilot's second year to partially offset the attrition rate in the original chronically ill intervention populations. Given the inconsistency of financial performance between the first 6 and next 12 months in the much larger original populations, we believe it is too early to make inferences from the "refresh" experience.

CHAPTER 7

KEY FINDINGS BASED ON 18 MONTHS OF PHASE I OF THE MEDICARE HEALTH SUPPORT PILOT

The purpose of this Report to Congress is to report the results of RTI International’s 18-month evaluation of eight Medicare Health Support (MHS) pilot programs implemented under Phase I of the “Voluntary Chronic Care Improvement Program (CCIP) Under Traditional Fee-for-Service (FFS) Medicare,” pilot as authorized by Section 721 of the Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (Pub. L. 108–173) (codified as Section 1807 of the Social Security Act, hereafter “the Act”). Section 721 requires the Secretary of Health and Human Services to provide for the phased-in development, testing, evaluation, and implementation of chronic care improvement programs. Subsection (b)(5) of the legislation states that an independent evaluation shall include an assessment of the following factors for each program:

- quality improvement measures,
- beneficiary and provider satisfaction,
- health outcomes, and
- financial outcomes.

In addition to assessing legislatively required factors for each program, our evaluation seeks to answer a broader set of research questions related to (1) how well the MHSOs were able to engage their intended audiences, (2) how well the MHSOs were able to implement their planned interventions and how their programs evolved over the course of the pilot, and (3) whether the programs improved knowledge and self-management skills and led to behavioral change among participants. The evaluation includes these additional foci to better understand the factors for program success.

Section 1807(c)(1) of the Act states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves the clinical quality of care and beneficiary satisfaction and achieves targets for savings, the Secretary shall enter into agreements to expand the implementation of the program (or components) to additional geographic areas not covered under the program as conducted in Phase I.

In this chapter, we present key findings based upon the first 18 months of MHS operations, the mid-point of Phase I. Our findings are based on the experience of approximately 240,000 chronically ill Medicare beneficiaries randomized to an intervention or a comparison group in eight geographic areas in the original populations and approximately 47,000 beneficiaries in the refresh populations. To date, this is the largest randomized experiment in population-based case management ever conducted and was designed to test the scalability of such programs in Medicare FFS. Five key findings on participation, beneficiary satisfaction, clinical quality and health outcomes, and financial outcomes have important policy implications for the Centers for Medicare & Medicaid Services (CMS) and future disease management or care coordination efforts among Medicare fee-for-service (FFS) beneficiaries.

Key Finding #1: Several vulnerable sub-populations of Medicare FFS beneficiaries were less likely to agree to participate in the MHS pilot program.

In the first Report to Congress (McCall *et al.*, 2007), we observed a pattern whereby beneficiaries who agreed to participate within the first 6-months of the pilot tended to be considerably healthier and less costly in the prior year compared to those beneficiaries who never consented to participate. Over three-quarters of all original intervention beneficiaries verbally consented to participate in the MHS program during the first 18 months of the pilot; agreement rates range from 74 to 95%. We find that the participant populations continue to be healthier, less costly, and lower users of acute care services than beneficiaries who never participated during any of the first 18 months.

With the exception of McKesson, the proportion of participating beneficiaries with Medicaid enrollment is between 3 and 14 percentage points lower than for never participants. Six of the MHSOs have lower rates of Medicare beneficiaries who are under age 65, or beneficiaries with disabilities, among their participating beneficiaries.

The MHS Phase I pilot was designed to be a broad population-based FFS program. If CMS desires broadly focused care management programs, these interim findings suggest alternative recruiting and outreach strategies are needed to reach the sicker and more costly beneficiaries as well as dual Medicare/Medicaid enrollees and beneficiaries with disabilities as the current MHS recruitment strategies are not reaching these populations. These populations likely include a high proportion of beneficiaries residing in nursing homes or other institutional settings. During RTI's site visits, MHSOs reported that they found locating and engaging these populations very difficult.

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

We also examined the level of interaction between the MHSO and their participating beneficiaries. This analysis is restricted to beneficiaries who were eligible and a participant in all of the months of 7-18 of the pilot or were eligible and a participant until the time of their death. This restriction allows for a more straightforward evaluation of the distribution of months of support provided by the MHSOs and whether there is evidence of selective targeting of beneficiaries for intervention contacts based upon level of perceived need. The MHSOs received monthly management fees for all of the months these beneficiaries were alive during the 12 months of the analysis period.

Across Months 7–18 of the pilot for fully participating beneficiaries during this period, the majority of MHS beneficiaries received between 2 and 5 months of telephonic care management support. Monthly telephonic support is defined as any number of calls in the month. Green Ribbon Health, McKesson, and XLHealth provided no beneficiaries with 12 months of support. CHS and Health Dialog provided the largest percentage of beneficiaries with the most number of months of telephonic support; 45% of their participants received between 7 and 11 months of telephonic support. Healthways' distribution was bimodal; almost 40% of their beneficiaries received between 2 and 5 months of telephonic support while another almost 40% received between 7 and 11 months of telephonic support. Almost one-fifth of Aetna's fully

participating beneficiaries received no telephonic support in months 7 to 18. Fifteen percent of McKesson's and 10% of LifeMasters' fully participating beneficiaries received no telephonic support during the same year period.

The percentage of fully participating beneficiaries who received any in-person support is very low (not displayed) and reflects the telephonic design of most MHS programs. Less than 10% of beneficiaries in the MHS programs run by Aetna, Healthways, CHS, and Health Dialog received any in-person support. In contrast, XLHealth provided in-person support to over one-half of their fully participating beneficiaries. This also reflects their focus upon in-person assessment centers. However, almost one-half of XLHealth's beneficiaries who received in-person support had encounter(s) in only 1 month. Between 12 and 21% of beneficiaries in MHS programs run by LifeMasters, McKesson, and Green Ribbon Health received any in-person support.

Months of telephonic support did vary by health status risk score. At the outset, the MHSOs planned to stratify their intervention (i.e., mailings only, disease management, and intensive case management) by level of perceived need by the MHSO. We do observe a general pattern of increasing average number of months of telephonic support as the level of risk score increases; however, there is limited separation between the three levels of risk score for most of the MHSOs.

Although there was no pre-determined expected number of contacts, the MHS beneficiaries are a sick and costly group of FFS beneficiaries averaging over 1 hospitalization annually in the year prior to program launch, and the MHSOs reported significant unmet clinical and psychosocial need. From one-half to 85% of the beneficiaries who fully participated during months 7 -18 received less than 6 months of contact. Given the lack of consistent monthly or bimonthly interaction with many of the MHS participants, it is unlikely that the MHSOs will be successful at changing beneficiary behavior with respect to self-management of their chronic illness. Findings from the beneficiary survey shows there has been little meaningful improvement in self-care activities. To positively affect acute care utilization, one would expect to see improvement in self-care behaviors by the mid-way point of the pilot, and savings have proved illusive to date. Further examination is warranted in how the disease management strategies were implemented and whether there is evidence of successful selective targeting of beneficiaries for intervention contacts that are associated with positive outcomes.

Key Finding #3: There was limited effect in improving beneficiary satisfaction, experience with care, self-management, and physical and mental health functioning during the first 18-months of the Phase I pilot.

The MHS pilot programs employ strategies to improve quality of care while reducing costs by empowering Medicare beneficiaries to better manage their care. They do 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 beneficiary use of medications, eating habits, and exercise, as well as interacting more effectively with their primary health care provider. The MHS programs hypothesize that lifestyle changes and better communication with providers will mitigate acute flare-ups in the chronic conditions and should reduce hospital admissions and readmissions and the use of other

costly health services such as nursing homes and visits to specialists. Experiencing better health, beneficiaries should also be more satisfied that their health care providers are effectively helping them to cope with their chronic medical conditions.

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 at baseline (Months 4 to 6 of the intervention period) and 12 months later. No further surveying of the original populations or surveying of the refresh populations will be conducted during RTI's evaluation. Thus, these reported results are final results with respect to beneficiary satisfaction.

We observe limited MHS intervention effects on 27 beneficiary survey measures across seven of the eight MHSOs²⁹. Of the 189 measures (27 x 7), 25 (13%) showed significant positive intervention effect and 4 (2%) showed negative intervention effect. The remaining 160 indicators were not statistically different between intervention and comparison beneficiaries. One of the required conditions for expansion of the programs is improvement in beneficiary satisfaction. RTI and CMS have defined satisfaction to mean that "beneficiaries were helped by their health care team to cope with their chronic conditions." For this key satisfaction measure, a positive intervention effect is observed for 2 of 7 MHSOs, Health Dialog and Aetna.

We also evaluated two measures of the care experience that are elements of each MHSO's intervention: (1) the number of helpful educational discussions, and (2) beneficiaries' communication with their health care providers. Three MHSOs showed a positive intervention effect with respect to the number of helpful discussions and two MHSOs demonstrated a positive intervention effect with respect to the quality of beneficiary communication with their health care providers. Across the key satisfaction and the two experience-of-care measures, one of the seven MHSOs, Health Dialog, had positive intervention effects for all three of the measures, while three MHSOs showed mixed results, and another three MHSOs showed no positive intervention effects.

A goal of disease management is to increase compliance of appropriate self-care behaviors among the chronically ill. The survey instrument included measures to capture changes in beneficiary self-management focusing specifically upon willingness to set self-management goals, self-efficacy, and engagement in self-care activities. The MHSOs were most successful in helping beneficiaries to set goals and make plans to address their care needs. Five of the seven MHSOs showed positive intervention effects related to setting goals (Healthways, Health Dialog, Green Ribbon Health, McKesson, and XLHealth).

In spite of positive effects on setting goals, there was little meaningful improvement in self-efficacy or self-care activities. For example, both Health Dialog and McKesson showed statistically significant intervention effects in raising the proportion of beneficiaries who reported receiving help setting goals and making plans. For Health Dialog, this effect did not translate into increases in self-efficacy but did increase one of eight self-care activities. For McKesson,

²⁹ LifeMasters terminated its participation prior to RTI's fielding of the follow-up survey. Hence, no survey results are reported.

there is an observed decrease in one of the self-efficacy items and an increase in two of eight self-care activities. To positively affect acute care utilization, one would expect to see improvement in self-care behaviors by the mid-way point of the pilot.

We examined the effect of the MHS intervention on several measures of physical and mental health functional status. Given the age and the frailty of this population, we did not expect to see significant improvements in physical and mental health function; we did, however, expect to see a mitigation of the slope of the decline for the intervention group. We found only one statistically significant intervention effect in the activity of daily living (ADL) domain: the decline in the number of ADL difficulties at follow up was mitigated for the Health Dialog intervention group compared to its comparison group. Given the emphasis of many of the MHSOs on depression screening we were expecting to see some improvement in mental health; however, only one of the MHSOs, Aetna, had an impact on beneficiary mental health status with a mitigation of the decline in the PHQ score relative to the comparison group.

The Medicare Health Support (MHS) authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction, and achieves targets for savings, the program (or its components) may be expanded to additional geographic areas. Only two of the MHSOs, Health Dialog and Aetna, improved beneficiary satisfaction as measured by beneficiary assessment that their health care team helped them cope with their chronic condition.

None of the seven MHSOs included in the beneficiary survey analyses demonstrated consistent positive intervention effects across the four domains of satisfaction, care experience, self-management activities, and physical and mental health functioning. The focus of the pilot program interventions was largely on impacting beneficiary behavior to better manage their chronic illness. Yet these results show little evidence of changes in self-efficacy or self-care. We did not observe any consistent pattern of positive intervention effects by disease cohort of heart failure only, diabetes only, and heart failure and diabetes.

Key Finding #4: Seven of the MHSOs had a positive intervention effect on one or more process-of-care measures but no positive intervention effect on reduction in acute care utilization or mortality.

One of the required conditions for expansion of the MHSO programs is improvement in quality of care and the Medicare Health Support (MHS) legislation states that the evaluation shall include an assessment of quality improvement measures and health outcomes. We have defined quality improvement for this evaluation as an increase in 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, re-admissions, and ER visits, and a reduction in mortality rates. We present *interim* results for an 18-month period for the original populations and *preliminary* selected 6-month results for the refresh populations.

Across 40 quality of care measures (five measures for each of the 8 MHSOs), there was modest improvement in 16 (or 40%) measures for the original populations. Seven of the 8 MHSOs demonstrated at least one positive intervention effect. Healthways demonstrated a positive intervention effect across all five process-of-care measures and CHS across four of the

five measures. LifeMasters improved cholesterol screening rates among beneficiaries with heart failure as well as diabetes but demonstrated no other positive intervention effects. Aetna and GRH improved the rate of cholesterol screening for beneficiaries with diabetes, and McKesson improved the rates of cholesterol screening and HbA1c testing in beneficiaries in diabetes. None of these three MHSOs demonstrated a positive intervention effect related to beneficiaries with heart failure. Health Dialog demonstrated a positive intervention effect related to cholesterol screening for beneficiaries with heart failure but did not demonstrate any intervention effects related to quality of care for beneficiaries with diabetes. XLHealth did not demonstrate any positive intervention effects on quality of care.

Rates of improvement in the clinical quality of care measures were relatively modest, 2 to 4 percentage points. The MHSOs were most successful improving cholesterol screening. Examination of the underlying trends in rates of change in the comparison populations show a very clear pattern of declining rates of cholesterol screening and HbA1c testing over time suggesting a possible ceiling effect in the 70 to 80% range for a chronically ill population. The positive intervention effects observed for these two measures were primarily driven by the MHSOs mitigating the observed downward trend in the comparison groups or modestly improving the intervention groups' rates. In contrast, we observe substantial increases in rates of retinal eye examination across the comparison groups, generally a 20 percentage point improvement, over relatively low baseline rates, and, modest increases in rates of urine protein screening over baseline rates that generally were lower than those observed for cholesterol screening and HbA1c testing. Only Healthways was successful at outperforming the comparison populations' improvement in both of these measures and CHS was successful having a positive intervention effect for one of these two measures.

For both the original and refresh populations, none of the 8 MHSOs demonstrated positive intervention effects related to health outcomes. Across the 120 comparisons for the original populations, there were no statistically significant reductions in the rate of growth in hospitalizations, re-admissions, or ER visits in the intervention groups relative to the comparison groups. Nor do we observe any statistically significant reductions in rates of hospitalization or ER visits across 74 comparisons for the refresh populations. We observe no reduction in mortality rates or time to death during the first 18 months of the pilot for the original populations and the first 6 months of engagement of the refresh populations.

Key Finding #5: Fees accrued to date far exceed savings produced.

RTI, as the independent evaluator, is responsible for evaluating the financial success of MHSOs defined as budget neutrality with respect to their accrued management fees. RTI's findings through 18 months, or halfway through the pilot, are based on the experience of approximately 240,000 chronically ill Medicare beneficiaries randomized to an intervention or comparison group in the original populations and approximately 47,000 beneficiaries in the refresh populations. To date, this is the largest randomized experiment in population-based care management ever conducted.

None of the 8 MHSOs achieved gross savings rates that were statistically different from zero for their original and refresh populations. Lack of statistical success was not due to small sample sizes. RTI was able to detect savings rates well below 5% of average monthly Medicare

claims payments. The fact that 4-of-8 MHSOs' intervention groups exhibited greater increases in spending than in their randomized comparison groups reinforces our conclusion of no discernable effect of these MHSOs' interventions on Medicare claims costs. Four-of-seven MHSOs exhibited slower intervention growth rates in their refresh populations after the first six months, but none of these savings were statistically significant. Given the inconsistency of financial performance between the first 6 and next 12 months in the much larger original populations, we believe it is too early to make inferences from the "refresh" experience.

The lack of financial success was uniform across five broad disease groups. Of the 40 possible statistical tests conducted on the MHSOs' original populations, only one was significant. Using a 95% two-sided confidence interval, we should expect to observe one statistically positive or negative difference simply at random that could disappear by the end of the pilot period. Moreover, no systematic pattern of success was found by disease group that could be indicative of intervention effects at least with one disease group. It is questionable, statistically, to focus on a single positive finding when the overwhelming pattern of results is contradictory.

Controlling for beneficiary characteristics including beneficiary age, gender, type of chronic disease, comorbid risk factors, and fraction of time in the pilot had no impact on the lack of demonstrable savings. As each of these factors were strong predictors of higher costs by themselves, the fact that they had no effect on the differences in intervention and comparison PBPM growth rates implies that they were equally distributed between the two groups at randomization.

Further, no evidence of savings was found among beneficiaries who agreed to participate in the intervention. Evidence was found of engagement of less costly, healthier, beneficiaries into the participant pool. Intervention beneficiaries who never consented to participate were much more expensive in the base year and experienced higher rates of cost increases than the participant group. Nevertheless, participant cost increases, alone, were no different than for the entire comparison group. These findings were insensitive to MHSO differences in participation rates. The two MHSOs with the lowest and highest participation rates also were the least successful in slowing the growth in Medicare spending. Difficulty in engaging sicker, more costly beneficiaries raises material questions about the future success of a broad, population-based, approach to Medicare chronic disease management.

Savings one-half way through the Phase I pilot period have offset 12-26% of estimated accrued fees for 4 of 8 MHSOs³⁰. The remaining 4 MHSOs show no evidence of savings. Halfway through the pilot, Medicare's return on investment, at best, is minus 74% and, at worse, greater than minus 100%. The 4 MHSOs with modest savings would have to increase their rate of savings by roughly 3-to-7-fold while the second group of 4 "non-savers" would have to reduce Medicare spending on services more than the entire average monthly fee they have accrued. This seems unlikely given performance over the first 18 months.

³⁰ RTI did not factor in a reduction in the negotiated fee requested by one MHSO after the 18-month period. The MHSO requested that its fee be reduced to \$0. Incorporating such a reduction would have resulted in this MHSO appearing more successful on the budget neutrality criterion than they actually were at the mid-point of the pilot.

Conclusion

The Medicare Health Support authorizing legislation states that if the results of the independent evaluation indicate that a program (or the components of such a program) improves clinical quality of care and beneficiary satisfaction, and achieves targets for savings, the program (or its components) may be expanded to additional geographic areas. None of the MHS pilot programs at the mid-point of the pilot have yet to meet the three statutory requirements to improve clinical quality of care and beneficiary satisfaction and achieve budget neutrality with respect to their fees.

Among their original populations, 7 of the MHS pilot programs modestly improved rates of receipt of guideline-concordant care but none reduced rates of acute care hospitalization, readmission, or ER visits. None reduced the rate of mortality. Two of the MHSOs improved beneficiary satisfaction. None of the MHSOs achieved budget neutrality within the first 18 months of program operations within their original populations. The 2 MHSOs that improved beneficiary satisfaction each had a positive modest intervention effect on one of five process-of-care measures. Neither of these 2 MHSOs lowered acute care hospitalizations or ER visits nor did they achieve budget neutrality. One MHSO had no intervention effect on beneficiary satisfaction, quality of care, or budget neutrality. The other 5 MHSOs modestly improved rates of receipt of guideline-concordant care but none lowered acute care hospitalizations, readmissions, or ER visits or achieved budget neutrality.

Among their refresh populations, and for only the first 6 months of intervention experience, none of the 7 MHSOs that accepted a refresh population improved health outcomes. We view these results as **preliminary**. We also observe no statistically significant reductions in rates of hospitalization or ER visits. Nor do we observe reduction in mortality rates during the first 6 months of engagement of the refresh populations. None of the 7 MHSOs experienced statistically lower per beneficiary per month (PBPM) growth in their intervention versus comparison group payments needed to achieve budget neutrality in this preliminary analysis.

Given the limited gains regarding quality of care and savings to offset accrued monthly management fees, it will be difficult to justify these private disease management models on cost effectiveness grounds—at least for chronically ill Medicare FFS beneficiaries. With 16 statistical successes out of 40 possible improvements in evidence-based process-of-care measures, the cost per successful improvement is approximately \$16 million, based on CMS' estimate of \$250 million in accrued MHS fees through 18 months for the 160,000 original population intervention beneficiaries. The cost would be \$6.4 million per percentage point improvement. Accounting for the 25 (of 189) improved indicators of beneficiary satisfaction, care experience, and self-management does not materially alter our conclusion. Nor is there any obvious correlation between MHSOs that partially offset their fees and their quality of care improvements.

The findings presented in this second Report to Congress are based upon the first 18 months of MHS operations for the original populations, the mid-point of Phase I, and 6 months of MHS operations for the refresh populations. We include the experiences of both the original and refresh populations in this Report to Congress to capture the impact of the early evolution of the MHS programs on acute care utilization and savings. The third Report to Congress will contain the evaluation of the full 3-year Phase I implementation experience of the original

populations and 2-year experience of the refresh populations and will report on provider satisfaction with the MHS Phase I pilot and the MHSOs' effect on quality of care, health outcomes, and Medicare program savings.

In conducting its final analyses of the full 36-month MHS Phase I pilot and in support of the third Report to Congress, RTI will repeat the key analyses conducted and included in this Report to Congress. In addition, RTI will expand its core set of analyses to more fully examine four evaluation issues: (1) the impact of the evolution of the MHS programs on outcomes; (2) the impact of interventional targeting on outcomes; (3) the natural progression of costliness of Medicare FFS beneficiaries with the MHS targeted clinical conditions and high risk scores; and (4) the impact of the MHS programs on beneficiaries who were less likely to agree to participate in MHS. The primary purposes of these additional analyses are to more fully examine the interim findings and to help inform the future design of programs in the Medicare FFS population.

During the first 18 months of the Phase I pilot, we observed evolution of the MHS programs. While some of evolution reflected gained knowledge about the level of co-morbidity among the MHS FFS beneficiaries, key substantive changes were made in most of the MHS programs to address cognitive/psychological and social support services, and end-of-life care planning. Substantive changes were also made in the engagement and the initial stages of intervention of the refresh populations. RTI will examine the full 3-year experience of the original populations and 2-year experience of the refresh populations, thus capturing the dynamic nature of the implemented programs.

Although there was no pre-determined expected number of contacts, the MHS beneficiaries are a sick and costly group of FFS beneficiaries and the MHSOs reported significant unmet clinical and psychosocial need. The majority of fully eligible and participating MHS beneficiaries during months 7-18 of the pilot received between 2 and 5 months of telephonic support³¹. RTI will more fully examine how the care management strategies were implemented and whether there is evidence of selective targeting of beneficiaries for intervention contacts. We will construct alternative measures of degree of intervention and identify characteristics of intervention beneficiaries that had a high degree of intervention versus low degree of intervention. These measures will also be incorporated into multivariate modeling of acute care utilization and financial savings to explore whether there are levels of intervention associated with positive outcomes.

Because MHS pilot beneficiaries were selected based on projected costliness due to chronic illness, they have much higher than average PBPM costs. RTI will explore in greater detail the regression-to-the mean phenomenon and implications of this phenomenon on explaining intervention effects in MHS pilot programs. We will study how strong is the regression-to-the mean effect in higher vs. lower cost subgroups, the impact of the regression-to-the mean effect on variation in PBPM costs and the statistical confidence in estimated savings, and whether regression-to-the mean effects differ between participating and non-participating beneficiaries. Further, RTI will examine the cost trajectory over the 3-year pilot period of subpopulations of MHS beneficiaries defined by baseline costliness, clinical comorbidities, and

³¹ Monthly support is defined as any number of calls in the month.

prior utilization of selected health care services and the implications of these findings on the design of Medicare programs.

Lastly, the MHS Phase I pilot was designed to be a broad population-based FFS program. If CMS desires broadly focused care management programs, these interim findings suggest alternative recruiting and outreach strategies are needed to reach the sicker and more costly beneficiaries as the current MHS recruitment strategies are not reaching these populations to the degree they are reaching other FFS beneficiaries. However, we have not examined whether the MHS interventions are having a positive impact on participating beneficiaries with characteristics similar to those beneficiaries that were less likely to agree to participate, (i.e., disabilities, Medicare/Medicaid dual enrollees, men, racial minorities, and those residing in institutional settings). RTI will explore the degree of intervention within these subpopulations and the impact of the MHS interventions on acute care utilization and financial savings.

REFERENCES

Bernard, S.L., and J. Uhrig (February 2002). "Subgroup Analysis: Implementation of Medicare Fee-for-Service Survey." Final report for the 2001 Survey. Prepared for the Centers for Medicare and Medicaid Services.

Kazis, L.F., Lee, A., Spiro, A., et al. Measurement comparisons of the Medical Outcomes Study and Veterans SF-36 Health Survey. Health Care Financing Review 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.

Toobert, D.J., S.E., Hampson, R.E. Glasgow: The summary of diabetes self-care activities measure: Results from 7 studies and a revised scale. Diabetes Care 23(7):943-50, 2000.