Centers for Medicare & Medicaid Services

Evaluation of the Senior Risk Reduction Demonstration (SRRD) Under Medicare

Contract Number: RTOP CMS-08-023JF

First-Year Evaluation Report – FINAL

February 2012

Prepared by:

IMPAQ International, LLC
Daver Kahvecioglu, Ph.D.
Betty Fout, Ph.D.
Daniel Weinberg, Ph.D.
Nicholas Bill, M.P.P.

The RAND Corporation
Katherine Kahn, M.D.

The Lewin Group
Namrata Sen, M.H.S.A.

Submitted to:

Pauline Karikari-Martin
CMS, 7500 Security Blvd
Mail Stop WB-06-05
Baltimore, MD 21244

Submitted by:

IMPAQ International, LLC
10420 Little Patuxent Parkway, Suite 300
Columbia, MD 21044
Phone: (443)367-0088 / Fax: (443)367-0477
On the Web: www.impaqint.com

ACKNOWLEDGEMENTS

This report greatly benefited from guidance and comments by Pauline Karikari-Martin and William Clark of CMS. In addition, Peter Mueser and Jacob Benus from IMPAQ International and Carol Simon from the Lewin Group provided useful comments. We would also like to thank the participants of the third annual SRRD evaluation conference that was held at CMS on November 12, 2011.

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY

This report is an interim evaluation of the Centers for Medicare & Medicaid Services' (CMS), Senior Risk Reduction Demonstration (SRRD) Under Medicare. SRRD began on May 1, 2009 and will end on April 30, 2012. Following CMS' approval, two SRRD vendors, referred to as "Vendor A" and "Vendor B," adapted health risk reduction programs available for commercial populations to the Medicare population. In this report, we examine the impact of the first year of the demonstration on various outcomes, using data from health risk assessments (HRA) collected yearly (and at baseline) as well as Medicare claims data obtained yearly (and at baseline).

The evaluation methodology we used is the most rigorous type because it exploits the randomized control trial (RCT) demonstration design. From the population of community-dwelling, Medicare fee-for-service (FFS) beneficiaries between the ages of 67 and 74, the implementation contractor randomly assigned beneficiaries to an intervention group (IG) or an administrative control group (ACG). Each vendor then sent an HRA packet to its assigned IG beneficiaries to recruit them for the study. Once a vendor received a completed HRA from a beneficiary, the beneficiary was considered successfully recruited to the study and was termed an SRRD participant. Participants were then randomly assigned to one of three intervention arms: Arm 1 – Standard Treatment (HRA + standard tailored follow-up), Arm 2 – Enhanced Treatment (HRA + enhanced tailored follow-up), or Arm 3 – HRA Only (HRA + generic health advice). Each vendor implemented the program nationally (n=6,056 and n=5,893 for Vendor A and Vendor B, respectively) and in two local areas in partnership with the local Aging and Disability Resource Centers (n= 1,417 and n=1,482 for Vendor A and Vendor B, respectively).

SRRD can be characterized as eight distinct interventions implemented in parallel: Two vendors, each having their own unique design and implementation, two intervention arms (standard and enhanced), and two implementation settings (national and local) make up the eight variations (2x2x2=8). We constructed 27 SRRD subgroups based on systematic aggregations of these eight variations. For each of these subgroups, we estimated the impact of SRRD using Medicare administrative data on 29 outcome measures that fall into three categories: (1) Medicare Part A and Part B use and payments; (2) hospitalizations for ambulatory care sensitive conditions, developed by the Agency for Healthcare and Research (AHRQ); and (3) preventive screening measures. We also estimated the impact of SRRD using HRA data on 16 vendor-based risk measures for a number of subgroups. As described above, we cast a wide net to detect SRRD impact by looking at a number of outcome measures from various domains and at various SRRD subgroups that might have benefited from SRRD.

In this report, we examined the impact of the SRRD after one year of implementation. We found that Vendor A's national enhanced intervention (Arm 2) reduced key Medicare utilization rates and expenditures: On average, Vendor A's participants (i.e., Medicare beneficiaries who filled out an HRA) that were randomly assigned to Arm 2 (i.e., eligible to receive enhanced tailored follow-up services), cost Medicare $958 less and were 14.2 percent less likely to be

hospitalized than Vendor A's participants that were randomly assigned to Arm 3 and thus were not eligible for any tailored follow-up services. Both impacts are statistically significant at the 5 percent level. These are significant program impacts provided that the impact of the control intervention (HRA + generic health advice) is minimal. We did not find any statistically significant impact of Vendor B's national enhanced intervention. However, we did detect some impacts of Vendor B's national *standard* intervention. The impact of Vendor B's national standard intervention (Arm 1) net of its control intervention was to *increase* total and carrier Medicare expenditures by $822 and $273, respectively. However, the impact on total Medicare expenditures was less precisely estimated (significant at the 10 percent level).

Given these dramatic early impacts, we conducted further in-depth analyses, which confirmed our original estimates. The impacts were driven by participants that Vendor A stratified into high-risk and moderate-risk groups based on its proprietary HRA risk-scoring algorithm. The high-risk group represented about 7 percent of participants and experienced an impact of about $6,600 reduction in total Medicare Part A and B expenditures. The moderate-risk group constituted 28 percent of participants and experienced an impact of about $1,500 reduction in total Medicare Part A and B expenditures. The most dramatic impacts were estimated for dual eligibles (those participants who were eligible for both Medicare and Medicaid) who were in the high-risk group: on average, $14,000 reduction in total Medicare Part A and B expenditures. Even though statistically significant, this extremely large estimate may be a result of small sample size (n=75).

Given the large first-year impacts for Vendor A, the question is what did Vendor A do differently? There are several potential explanations. First, Vendor A applied a more personalized recruitment strategy and used telematching to find beneficiary telephone numbers that CMS may not have, which may have helped in recruiting hard-to-recruit but highly impactable beneficiaries such as ethnic minorities and dual eligibles. Second, Vendor A developed an effective risk-scoring algorithm and triaging approach, so that beneficiaries who had the greatest need and were most impactable received the most intensive services. Vendor A included more risk factors in the overall risk determination and developed its stratification algorithm for this demonstration rather than using existing predictive models developed on different populations. Finally, Vendor A allowed beneficiaries to opt out easily, which may have helped to focus on beneficiaries who explicitly wanted to be part of the program and thus presumably were more likely to be impacted.

This first-year report offers early learning experiences in terms of policy and the potential impact of Section 4103 of the Affordable Care Act: Medicare Coverage of Annual Wellness Visit Providing a Personalized Prevention Plan. SRRD is similar to the intervention that was instituted under Section 4103 except that it is administered outside of physicians' offices.

The results of this interim evaluation are suggestive but not conclusive. We believe that additional analysis is needed to confirm that the approaches implemented by the demonstration vendors lead to sustainable reductions in Medicare expenditures. The final analysis of SRRD will shed additional light on this important question.

II. INTRODUCTION

Summary of Results:

- The SRRD is designed to test whether risk reduction programs shown to be effective in the private sector can be replicated successfully for the Medicare population.
- The SRRD was implemented by two health management vendors, which developed tailored risk reduction programs delivered via telephone (health coaching), mail, or Internet to program participants:
  - Each vendor developed customized health risk assessments (HRAs) that address 17 CMS-indicated health risk areas (offered as a paper form or online).
  - Beneficiaries who completed an HRA became program participants and were randomly assigned into one of three treatment arms offering varying levels of health risk reduction services.
  - Vendors also used proprietary algorithms to assign participants to high-risk, moderate-risk, and low-risk categories (based on the HRA responses) and tailored interventions according to risk level.
  - Vendors based the intensity and range of services offered according to arm assignment and risk level.
- The SRRD includes national and local components; local samples intended to evaluate the impact of the demonstration in localities where Aging and Disability Resource Centers (ADRCs) deliver exemplary information, referral, and assistance services:
  - The ADRCs adapted to a proactive approach of making outbound calls to local SRRD participants (typically handle inbound calls only).
- Recruitment goals were exceeded by each vendor for first two years of the demonstration (with the exception of one local sample in Year 1); each vendor used different recruitment strategies (wave approach, in-person versus automated calls, etc.)

A. Overview of the Senior Risk Reduction Demonstration (SRRD)

Recent research suggests that well-structured risk reduction programs can achieve significant improvements in a population's health risk profile. The *Evidence Report and Evidence Based Recommendations: Health Risk Appraisals and Medicare*, commissioned by CMS and prepared by RAND, has concluded that effective risk reduction programs, beginning with administration of an HRA and including tailored behavior change follow-up interventions, can improve general health status outcomes.

The Senior Risk Reduction Demonstration (SRRD) is a demonstration in health promotion and health management designed to test whether the success of well-structured health risk reduction programs in the private sector can be replicated by appropriately adapting those

programs to the Medicare population. If successful, SRRD can help Medicare beneficiaries reduce health risks by improving their health behavior, which in turn will improve their health and produce savings or achieve cost-neutrality for Medicare. The HRAs used in the SRRD collect self-reported information from Medicare beneficiaries on 17 health risk factors, specified by CMS, that contribute to disease, including physical inactivity, obesity, smoking, depression, high blood pressure, high cholesterol, high glucose, and inappropriate use of clinical preventive services.[1] The HRA results were used to generate individual feedback reports, which were then reviewed with the beneficiaries by trained health coaches to develop plans for reducing modifiable health risk behaviors.

CMS originally selected five vendors, but after participating in the SRRD pilots three of the vendors decided not to continue with the demonstration. The remaining two vendors implemented their own distinct risk reduction programs for an eligible sample of target Medicare beneficiaries. After the second year of the demonstration, one of the vendors chose to withdraw.

Medicare beneficiaries were eligible for inclusion in the SRRD if they met the following criteria:

- Be between the ages of 67 and 74 at the start of the demonstration.
- Be a Medicare FFS beneficiary enrolled in Parts A and B.
- May be eligible for both Medicare and Medicaid.
- Medicare must be the primary payer.
- Cannot be enrolled in a Medicare Advantage Health Plan.
- Cannot be enrolled in hospice or have end-stage renal disease (ESRD).
- Cannot have resided in any institution for 100 days or more in the past 12 months.
- Initial enrollment in Medicare cannot be before age 65.

Exhibit 2.1 illustrates the SRRD's study design. Each vendor received three random samples: two local samples (L1 and L2) and a national sample (N). The local samples were intended to evaluate the impact of the demonstration in some of the localities where exemplary Aging and Disability Resource Centers (ADRC) are active. Beneficiaries were randomly assigned to an intervention group (IG) or an administrative control group (ACG) in this "first randomization." The first randomization ensured that both vendors worked with nationally representative Medicare beneficiaries and that beneficiaries assigned to vendors were comparable. Vendors sent an HRA packet to their assigned IG beneficiaries to recruit them for the study. Both vendors provided the option to complete and submit the HRA on paper or online; however,

[1] The following 17 health risk factors were used: physical inactivity/lack of exercise, poor nutrition, smoking/tobacco use, excessive alcohol consumption, high blood pressure, high blood glucose, high total cholesterol, being overweight/obese, inappropriate use of clinical preventive services, depression, high stress, lack of general well-being, burden of providing care giving, social isolation, lack of motor vehicle/home safety, falls (preventable accidents), and polypharmacy/medication issues.

most beneficiaries selected the paper mode (approximately 6 to 8 percent completed the HRA online). ACG members were never contacted. Once a vendor received a completed HRA, the beneficiary was considered successfully recruited into the study and was termed an SRRD participant. Finally, each SRRD participant was randomly assigned by the implementation contractor to one of three intervention arms in a "second randomization." Depending on the assignment, the participant received personalized feedback and/or a health coaching intervention.

As shown in Exhibit 2.1, each vendor received 20,000 total beneficiaries, which were split among the different sample types (L1, L2, and N). Another 20,000 (split among the different sample types) were assigned to the ACG, resulting in a total of 80,000 beneficiaries participating in the demonstration.

Exhibit 2.1: SRRD Study Design

Sample Type	Study Group (First Random Assignment)	Sample Size	Treatment Arm (Second Random Assignment)
N	IG	16,290	Arm 1 – Standard
			Arm 2 – Enhanced
			Arm 3 – HRA Only
	ACG	16,290	N/A
L1	IG	1,855	Arm 1 – Standard
			Arm 2 – Enhanced
			Arm 3 - HRA Only
	ACG	1,855	N/A
L2	IG	1,855	Arm 1 – Standard
			Arm 2 – Enhanced
			Arm 3 - HRA Only
	ACG	1,855	N/A
Total		**40,000 x 2 vendors = 80,000**	

The demonstration design required that each vendor administer the following three levels of intervention:

- *Arm 1 – Standard Treatment (HRA + standard tailored follow-up)*: Arm 1 participants received the standard intervention. Under the standard intervention, vendors provided the following services: an individualized feedback and follow-up report; tailored, behavioral risk-specific intervention modules delivered through the mail, via the Internet (if the participant preferred), or (optionally) through proactive telephone counseling and health coaching; high-risk programming;[2] a help line that participants could call with questions and concerns; and referrals to national or local risk reduction resources.

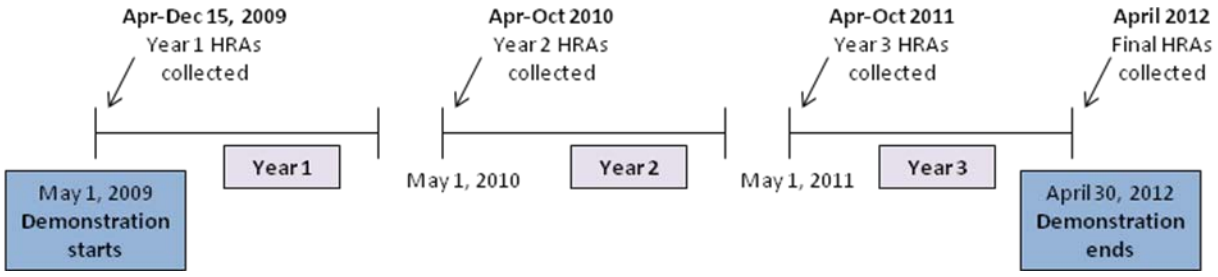
[2] Vendors could offer additional, more frequent, or more intensive interventions to participants categorized as high risk.

- *Arm 2 – Enhanced Treatment (HRA + enhanced tailored follow-up)*: Arm 2 participants received an enhanced intervention. Arm 2 provided all the components of Arm 1, but also offered more intensive interventions, including required proactive telephone counseling for a subgroup of participants. The individual vendor determined which participants were most suitable for telephone counseling. In addition, vendors could offer additional tailored behavior change modules, more frequent and/or more intensive interactions with beneficiaries, and greater access to health educators to support risk reduction efforts.
- *Arm 3 – HRA Only (HRA + generic health advice)*: Arm 3 was designed as the placebo intervention. Participants assigned to Arm 3 received an HRA, but no intervention aside from receiving a standardized, non-tailored letter describing the advantages of a healthy lifestyle. No additional follow-up interventions other than a generic health brochure were provided. The HRAs were collected and the vendors assigned risk scores to the participants. However, the participants did not receive their HRA results.

Vendors had significant discretion in the triage of participants into intervention cycles based on risk level. Within each treatment arm, participants might receive several intensities of intervention. For example, a low-risk participant in the enhanced treatment arm might receive a less intensive intervention than would a high-risk participant in the standard treatment arm. The next section provides more detail on the vendors' risk reduction programs.

Recruited beneficiaries (participants) filled out an HRA between April 1, 2009 and December 15, 2009 (the beginning of Year 1). Additional HRAs were collected at the beginning of Year 2 and Year 3, and a final HRA will be collected at the end of the demonstration (see Exhibit 2.2). Thus, four rounds of HRAs will be collected throughout the demonstration, with the Year 1 HRA considered to be the HRA taken at "baseline" (see Exhibit 2.2). For this first-year evaluation, we consider only the Year 1 and Year 2 HRAs.

Exhibit 2.2: Timing of HRA Collection [a]



[a] One vendor chose not to continue with the demonstration after Year 2. Therefore, Year 3 and final HRAs will be collected only for one vendor's participants.

B. Vendor Risk Reduction Programs

Although each vendor was required to implement certain components of the demonstration as prescribed in the demonstration design and its terms and conditions, each vendor constructed its own risk reduction program, which allowed considerable flexibility and innovation. Specifically, the vendors developed and implemented unique approaches for participant recruitment and retention; HRA questionnaire design and administration; risk stratification (prioritization) and triage; and intervention services. The approaches used by the vendors (identified as "Vendor A" and "Vendor B") are described below.

1. Recruitment and Retention

Vendor A

In Year 1, Vendor A decided to use a wave approach for recruitment. The recruitment materials were mailed in six different waves and consisted of an initial CMS packet with a cover letter and an endorsement letter from C. Everett Koop, M.D., a "Get Ready" postcard, and an introduction packet. (Reminder postcards were sent as a fourth mailing to non-respondents.) Each mailing wave was followed by up to five awareness calls to non-respondents. Vendor staff conducted an average of three calls to each non-respondent in Year 1. For Year 2, the vendor reported that far less telephonic outreach was required to reach enrollment targets. The vendor kept the same wave approach but decided to focus first on the Year 1 participants and then on new beneficiaries, which may have helped to achieve a higher re-enrollment rate.

Vendor B

Vendor B sent its recruitment materials to the entire SRRD sample (20,000) in a single wave in Year 1. The CMS packet was mailed first, followed by awareness calls (half were automated using Interactive Voice Response (IVR), the other half were made in-person through outsourcing), introductory brochures, and the HRA materials packet. Reminder postcards and another round of awareness calls followed these materials. Vendor staff conducted up to two awareness calls. The maximum number of attempts for the national sample and one of the local samples was six. For the vendor's other local sample group, vendor staff made up to eight calls to meet enrollment targets. In Year 2, the vendor decided to change its approach and sent out the recruitment materials and conducted awareness calls in four waves of 5,000 beneficiaries each. This change improved the follow-up time with beneficiaries, as well as the response rate, and made the recruitment process more fluid.

In Year 2, both vendors decided to send a CMS "bridge letter" to Year 1 beneficiaries; Vendor B also sent a CMS calendar.

2. HRA Questionnaire

Vendor A

Each vendor used previously developed HRA questionnaire tools to design the HRA for SRRD. Vendor A took a variety of already developed HRAs, matched them to the 17 risk factors outlined by CMS, and included standardized questions specific to the older adult population. The HRA takes into account all 17 CMS-indicated risk areas plus an additional four (financial barriers, transportation barriers, independence level, and life quality). Vendor A also conducted focus group and cognitive testing to ensure that the questionnaire was understandable (aimed at a sixth-grade reading level) and made it as short as possible to maximize the response rate. Vendor A has a Spanish-language version of the questionnaire available, containing the same questions as the English version. The English version of the HRA indicates at the bottom of the form that the individual may request the questionnaire in Spanish. All of the completed HRAs are entered manually, and 25 percent of the HRAs are subjected to double data entry for quality assurance purposes. No changes were made to the HRA for Year 2 of the demonstration.

Vendor B

The HRA developed by Vendor B captures all 17 of the CMS-indicated health risk factor areas. This HRA is written at a seventh-grade reading level; the Spanish version is written at a sixth-grade level (this version also has fewer questions). An optical scanner is used by the vendor to input the data from the completed HRAs. The HRA used in Year 1 was changed in Year 2 to remove questions related to family history, because of the Genetic Information Nondiscrimination Act (GINA). These questions were replaced with new questions about skipping medications.

Both vendors offer the questionnaire online, but only approximately 5–6 percent were completed online.

3. Risk Stratification (Prioritization) and Triage

Vendor A

Each vendor used its own proprietary algorithms to stratify the SRRD participant population into low-risk, moderate-risk, and high-risk levels. Vendor A developed a new algorithm for this demonstration. The development of the algorithm was guided by risk factors identified during work on a previous project, but the algorithm is unique to SRRD. The algorithm calculates weighted points for each of the CMS-defined risk categories. The points (weighted based on the "impactability" of risk behavior) are totaled for each category. The stratification algorithm placed participants with the heaviest burden of coaching-modifiable health risks into a high level of intervention, and those with less risk into moderate or low levels of intervention. For the purposes of providing an enhanced intervention to Arm 2 (enhanced arm), the algorithm enables the vendor to identify participants for additional services. Vendor A's risk stratification

algorithm aimed to allocate approximately 10 percent, 30 percent, and 60 percent of participants into the high-, medium-, and low-risk groups, respectively.

Vendor B

Vendor B used two predictive models it had already developed to assign participants into risk categories based on short- and long-term risk. Short-term risk was based on a predictive modeling tool that projects total health services use for the participant in the following 12-month period. Long-term risk was based on a predictive model that estimates avoidable healthcare costs based on demographic and health risk data collected in the HRA. Smoking and BMI data supplement the predictive model in determining the risk category for each participant, and age and gender also are used for stratification. The short-term risk model generates a "lifestyle score" between 1 and 100, representing the participant's total controllable risk. Based on this score, 85 percent of SRRD participants were assigned into the high- or moderate-risk categories, and 15 percent into the low-risk category. Low-risk participants were not eligible for any intervention services. Vendor B's triaging did not distinguish between moderate- and high-risk participants. A "state of change" model was used to help tailor HRA feedback.

4. Intervention Services

Both vendors assigned different interventions to SRRD participants based on arm assignment. According to the demonstration design, Arm 3 participants were only offered the HRA and generic health advice and did not receive any other type of intervention. Arm 2 (enhanced) participants generally received more interventions than Arm 1 (standard) participants, but within Arm 1 and Arm 2 the type and intensity of intervention services might vary based on the participant's risk category.

Vendor A

Both vendors provided individualized HRA feedback reports that were closely reviewed with SRRD participants. For Vendor A, this took place during what is referred to as a "Lifestyle Management (LM)" or "Orientation" call conducted by a health advisor. During this call, the participant chooses one or more (based on risk category) intervention program "focus areas."[3] For high- or moderate-risk participants, the call is followed by proactive calls from a health coach.[4] High- and moderate-risk participants in Arm 2 also have the option to participate in the Chronic Disease Self-Management Program (CDSMP) Webinar,[5] and high-risk Arm 2 participants are given the option of having a social worker or geriatric RN assessment for a community resource referral. The low-risk participants in Arms 1 and 2 receive self-directed health improvement guides (one for Arm 1 and two for Arm 2).

[3] Arm 1 high-risk participants can select up to six focus areas; Arm 1 moderate-risk participants can select up to three; Arm 2 high- and moderate-risk participants can select up to six focus areas.

[4] Arm 1 moderate-risk participants get an average of three calls per year. Arm 1 high-risk participants and Arm 2 high- and moderate-risk participants get an average of six calls per year.

[5] Vendor A later discontinued the CDSMP Webinar option because of limited uptake.

Vendor A developed a website for SRRD that contains a variety of tools that eligible participants (Arm 1 and Arm 2) can use to assist in their healthcare management, including a health tracker, health calendar, information on allergies and medical conditions, a "my workouts" log, family health guide, and other features.

Vendor B

Vendor B's intervention process is slightly different. After completing the HRA, all eligible Arm 1 and Arm 2 participants are automatically enrolled in an intervention program. The health coaches then place a follow-up phone call for the intervention program, and participants may elect to change the topic for enrollment.[6] Vendor B offers programming for the following topic areas: back care, blood pressure, cholesterol, nutrition, exercise, weight management, stress management, and tobacco cessation.

Enrolled participants receive coaching calls soon after enrollment. Those in Arm 1 may choose a mail-based intervention program instead of one that is phone-based (Arm 2 participants do not have the mail-based option). Arm 1 participants also have access to the full online suite; Arm 2 participants have access only to the health portal portion of the suite. Arm 1 moderate- and high-risk participants receive up to five coaching calls (or six mailings if they choose the mail-based program). Arm 2 moderate- and high-risk participants receive up to 12 coaching calls per year and at least one mailing per year based on a topic area triggered by their HRA results. Arm 2 participants can also participate in intervention programs such as Health Care Consumerism and Diabetes Management. In Vendor B's design, Arm 1 participants—regardless of risk level—cannot receive the same or a higher level of resource-intensive services (e.g., number of coaching calls) than Arm 2 participants receive.

Vendor B customized its existing online suite for the SRRD project. The suite provides a variety of tools and interactive modules personalized for participants and allows eligible participants (Arm 1) to set up health-related goals and to track progress towards meeting those goals. Among other features, the "health tools" section contains self-paced modules with interactive tools, including assessments, calculators, and quizzes that help participants develop skills for managing asthma, diabetes, nutrition, weight, cholesterol, tobacco cessation, high blood pressure, heart health, fitness, and back health.

C. Overview of Health Risk Assessments

Health risk assessments (HRAs) have been used for several decades, most widely in health education and health promotion programs in the workplace.[7] Originally, HRAs were used to collect health risk data from individuals to produce personalized epidemiological-based

[6] To be eligible for a program, the participant must be assigned to the high- or moderate-risk category.

[7] Soler, R.E. et al. (2010). A systematic review of selected interventions for worksite health promotion: the assessment of health risks with feedback. *Journal of Preventive Medicine*, 38(2S), S237–S262.

projections of mortality risk.[8] However, HRAs have evolved to become interactive, Web-based tools that provide individualized feedback and educational messages designed to motivate behavior change and risk reduction. Generally, HRAs include the following elements:[9]

1. Assessment of personal health habits and risk factors supplemented with biometric measurements of physiologic health
2. Quantitative estimation or qualitative assessment of future risk of death or adverse health outcomes
3. A mechanism for providing feedback in the form of educational messages or counseling about ways to change behavior and health habits to potentially alter risk of disease or premature death

HRAs can be important tools for raising awareness of health issues to encourage behavioral change as well as for triaging individuals into risk-appropriate interventions and tracking changes over time. However, they have limitations, which include inaccuracy of the information, recall bias, respondents' lack of understanding of health risk questions, and the need to tailor HRAs to specific literacy, cultural, and age groups. Furthermore, findings show that HRAs alone are not effective in inducing long-term behavior change. Further support is necessary, and HRAs should be considered a first step towards a comprehensive framework of behavioral change and risk reduction.[10]

As of January 2011, as specified in Section 4103 of the Affordable Care Act, Medicare covers, without cost to beneficiaries, an annual wellness visit that includes an HRA followed by a customized wellness or personal prevention plan.[11] As part of the law, the Secretary of Health and Human Services is authorized to establish publicly available guidelines for HRAs, after consultation with relevant groups and entities.[12] The Centers for Disease Control and Prevention (CDC) is currently developing and has provided interim guidance regarding HRAs and their modes of provision for Medicare beneficiaries.[13] General guidance includes the following:

- All questions in the HRA must be actionable.
- Feedback received by the patient from the provider regarding HRA results should be the result of shared decision making.
- The HRA should be written at a 5th or 6th grade literacy level and in plain language.

[8] Schoenback, V.J., Wagner, E.H., and Beery, W.L. (1987). Health risk appraisal: review of evidence for effectiveness. *Health Services Research,* 22(4), 553–580.

[9] Partnership for Prevention and Thomson Reuters (2011). General Proceedings from a Public Forum, Expert Input, and the Research Literature for the Design of Patient-Centered Health Assessments, Final Report, March 15, 2011.

[10] Partnership for Prevention and Thomson Reuters (2011).

[11] Koh, H.H. and Sebelius, K.G. (2010). Promoting prevention through the Affordable Care Act. *New England Journal of Medicine*, 363(14), 1296–99.

[12] Partnership for Prevention and Thomson Reuters (2011).

[13] Centers for Disease Control and Prevention (2011). Interim Guidance for Health Risk Assessments and Their Modes of Provision for Medicare Beneficiaries.
Available at http://prevent.org/data/files/news/healthriskassessmentscdcfinal.pdf.

- The HRA should be linguistically, age, gender, and culturally appropriate for the patient.
- The HRA should be received no less than every 2 years to ensure compliance with current science related to health promotion and disease prevention and to take advantage of advances in technology.

In addition, the CDC identified six areas that HRA and HRA delivery should address. These include (1) content and design, (2) mode of administration, (3) primary care office capacity, (4) consumer/patient perspectives, (5) data, and (6) evaluation and quality assurance.

In this current landscape, the SRRD evaluation is of particular interest. SRRD tests whether health promotion and health management programs (including HRAs) that have been developed and tested in the private sector can be tailored to and work well with Medicare beneficiaries to improve their health and reduce avoidable health services use. Many aspects, though not all, of the CDC's guidance had already been incorporated by the vendors in this demonstration. The purpose of this first-year evaluation report is to test whether the use of HRAs is effective in the Medicare population and not to assess the vendors' HRAs in light of current recommendations.

D. Overview of the First Year Evaluation

1. Research Questions

The overall evaluation was designed to address four broad questions about SRRD:

1. How was the demonstration implemented?
    - Was the beneficiary population successfully assigned to treatment and control groups that are statistically identical?
    - Did the demonstration vendors implement the demonstration as designed?
    - How are the demonstration components of recruitment/retention, health risk assessment, risk stratification, triage, intervention (that is, service approaches, products, information, resources, and supports), and referrals implemented?
    - How does the implementation differ across vendors? How is it similar?
2. How successful was beneficiary participation in the demonstration?
    - What proportion of targeted beneficiaries were the vendors able to recruit? What are the characteristics of participants and how do they differ from eligible beneficiaries?
    - Similarly, what are the participation rates and characteristics of opt-outs and other subgroups defined by age, gender, geographic location, race, health risk factors, recent health services use, and expenditures?

- How do participant rates and characteristics compare across these various beneficiary groups (participants, opt-outs), across vendors, across national/local demonstration, and across local communities?

3. Was the demonstration effective for participants?
   - Does the demonstration improve health behavior and reduce health risks?
   - Does it lead to improvement in health?
   - How does it affect Medicare use and also Medicare-covered preventive services use?
   - Does it reduce Medicare expenditures?
   - Does its impact vary by its implementation?
   - How does its impact vary by subgroup?
4. Was the demonstration budget-neutral?
   - Is the SRRD budget-neutral, using CMS' pre-defined methodology approved by the Office of Management and Budget?

In this first-year report, we focus on answering key parts of research questions 1, 2, and 3. Sections D.2 through D.5 below (Site Visits, Recruitment Results, Local SRRD Implementation, and Data Management and Reporting Requirements) address the first research question regarding how the demonstration was implemented.

We address research questions 2 and 3 in the subsequent chapters of this report. In particular, we assess whether the beneficiary population was successfully assigned to treatment and control groups and whether the demonstration in its first year has generated improvements in various outcomes, including Medicare expenditure and use, and several measures from the vendors' health risk assessments. We also conducted an analysis of participation rates and attrition rates for the national and local demonstration samples and for subgroups of interest, to identify potential factors underlying successful recruiting strategies.

We conducted two sets of analyses exploiting the random assignment of participants into the three intervention groups. These analyses consider the program's effectiveness on two different samples:

- Beneficiaries who participated in the program (participants) in Year 1. CMS defines an SRRD participant as an eligible beneficiary who completes the health questionnaire and returns it to the vendor. An SRRD participant may choose not to receive actual vendor intervention beyond filling out the HRA.
- Beneficiaries who participated in the program in both Year 1 and Year 2. These are beneficiaries who completed HRAs in both Year 1 and Year 2 of the demonstration.

The first analysis is based on the Intent-to-Treat model. The second analysis examines only the subgroup of beneficiaries who completed HRAs in both years.

Finally, to address research question 4 (Was the demonstration budget-neutral?), we prepared an interim budget neutrality analysis in February 2011. The SRRD design stipulated that if per member per month (PMPM) Medicare FFS claims expenditure of the intervention group (IG) in the first demonstration year exceeded that of the administrative control group (ACG) by more than 5 percent, then the vendor must accept an increase of the fee deferral rate to 20 percent from 10 percent or terminate its participation for the last year of the demonstration. The interim budget neutrality results, for both vendors, showed that differences in PMPM for the IG and the ACG were within the acceptable 5 percent range, and no adjustment was necessary to their fees.

2. Site Visits

To address aspects of research question 1, we conducted initial onsite visits to the vendors in mid to late August 2009, during Year 1 of the SRRD implementation. The purpose was to observe program operations and conduct in-depth interviews with vendor staff, including project directors, senior managers, account managers, customer service directors, health coach supervisors, and IT programmers. Prior to conducting each visit, we prepared an interview questionnaire to collect information on the following key topic areas:

1. Organization background, staffing roles, and responsibilities
2. Recruitment and retention
3. HRA questionnaire and assessments
4. Risk stratification (prioritization) and triage
5. Intervention services
6. Local SRRD implementation
7. Data management systems and reporting requirements
8. Lessons learned from pilots and other risk reduction initiatives

After completing the initial site visits, we submitted site visit reports to CMS in July 2010 and April and May 2011. The last two reports were based on conference calls with the two vendors and with the ADRCs working with the vendors to implement the demonstration's local component.

3. Recruitment Results

The demonstration's Year 1 recruitment period extended from April 1, 2009 through December 15, 2009. The initial recruitment period had to be extended into December because the vendors had some difficulty meeting their enrollment targets, particularly for the local SRRD

samples. Exhibit 2.3 below shows that the enrollment target for one local sample was never reached, despite additional efforts that included receiving an additional pool of Medicare beneficiaries to recruit from and an extra wave of HRA packets sent to non-respondents at the end of October.

Exhibit 2.3: SRRD Year 1 Enrollment Summary[a]

Study Group	Vendor Goal	Vendor A		Vendor B	
		New Enrollees	% of Goal	New Enrollees	% of Goal
National	5,564	6,343	114%	5,970	107%
Local 1	742	742	100%	912	123%
Local 2	742	789	106%	595	80%
TOTAL	**7,048**	**7,874**	**112%**	**7,477**	**106%**

[a] The numbers in this enrollment table include a supplemental list of beneficiaries that vendors received from CMS towards the end of their recruitment period to help them meet their recruitment targets. Our analyses in subsequent sections consider only the original list of beneficiaries. A breakdown of new enrollees and re-enrollees that were used in our analyses is shown in Exhibit 3.2.

The Year 2 recruitment period lasted from April 1, 2010 through October 31, 2010. As seen in Exhibit 2.4, all enrollment targets were exceeded for Year 2. This success may be due, in part, to lessons learned from Year 1 and a shift in recruitment approaches for Year 2. As mentioned, previously, in Section II.B.1, Vendor B did not use a wave approach during Year 1. In Year 2, Vendor B decided to implement a wave approach, which improved the follow-up time with beneficiaries and made the process more fluid. In addition, both vendors sent a "bridge letter" to beneficiaries to facilitate re-enrollment. Although one local sample still posed a challenge for Vendor B, CMS provided a "refresh sample" of 400 beneficiaries, which helped Vendor B meet its targets. Vendor A also adjusted its recruitment efforts slightly, based on lessons learned from Year 1. Vendor A began by focusing on beneficiaries from Year 1 through several "reach-out" waves and then shifted to new beneficiaries. This may have helped to achieve a higher re-enrollment rate.

Exhibit 2.4: SRRD Year 2 Enrollment Summary[a]

Study Group	Vendor Goal	Vendor A				Vendor B			
		Re-Enrollees	New Enrollees	Total	% of Goal	Re-Enrollees	New Enrollees	Total	% of Goal
National	4,451	3,789	2,331	6,120	137%	3,332	1,976	5,308	119%
Local 1	594	471	212	683	115%	529	219	748	126%
Local 2	594	478	236	714	120%	358	253	611	103%
TOTAL	**5,639**	**4,738**	**2,779**	**7517**	**133%**	**4,219**	**2,448**	**6,667**	**118%**

[a] The numbers in this enrollment table include a supplemental list of beneficiaries that vendors received from CMS towards the end of their recruitment period to help them meet their recruitment targets. Our analyses in subsequent sections consider only the original list of beneficiaries vendors received. A breakdown of new enrollees and re-enrollees that were used in our analyses is shown in Exhibit 3.2.

Vendor B noted that participation rates for SRRD were lower than are generally seen in similar programs. This was attributed to the modest incentive ($10 gift card) provided to the SRRD

participants for completing the HRA. Vendor B indicated that a typical incentive includes a $150 premium reduction. Furthermore, Vendor B explained that participation is often leveraged by the ability of organizations to promote management buy-in and by organizational culture. In Year 2, both vendors changed the incentive for completing the HRA from the $10 gift card to a booklet of Forever Stamps (worth $8.80).

4. Local SRRD Implementation

In the local demonstration, participants associated with a particular vendor also had to be referred to one of the two exemplary community programs assigned to that vendor. The basic role of the ADRCs was to obtain a list of SRRD participants from the vendor and to make initial outbound and follow-up phone calls to the participants to let them know about the ADRC's services, the local programs, and the services in the participant's community that they could enroll and participate in. The ADRCs would then report the results.

The vendors and the local ADRCs began coordinating with each other prior to the SRRD implementation start date. Coordination activities included making key staff introductions, reviewing the purpose of the demonstration and proposed activities, reviewing timelines, discussing the roles and responsibilities of each organization, conferring about plans for training, and discussing how overall coordination and implementation of the SRRD would occur. The ADRCs and the vendors jointly developed memoranda of understanding.

Although the ADRCs provide information, referral, and assistance services to seniors in their communities, the design of the SRRD posed some unique challenges for these organizations, and each noted that their initial understanding of how the program would unfold evolved as contacts were made. The ADRC staff usually did not make outbound calls to their clients; rather, they were accustomed to taking inbound calls from a significantly older and frailer population in need of services. The ADRCs reported that SRRD participants, for the most part, were healthier and more active than their usual clients and were largely uninterested in the services the ADRCs offered. For this reason, the ADRCs modified their outbound call scripts to move away from a "health counseling" approach and focused more on providing information about resources available in the community, through newsletters and resource guides. Both vendors assisted their ADRCs with this process. For instance, one vendor prepared an outbound call script for the ADRC staff to use, which included some basic instructions on how to proceed if the SRRD participant was not interested in hearing about referrals (e.g., asking if the individual would like to be placed on the ADRC's mailing list). The other vendor held recurring three-way conference calls with its ADRCs to discuss how to address the lack of referral uptake, which included offering resource guides and bi-monthly newsletters instead of specific referrals. Despite these challenges, the ADRCs found that taking a proactive approach to engaging consumers in their communities was a valuable lesson learned and helped their staff to develop new skills.

5. Data Management and Reporting Requirements

Vendor A

Vendor A developed a customized database and reporting mechanism for the SRRD project, using data warehousing capabilities built upon an extensive set of tools to generate reports tailored specifically for the SRRD program. SRRD data are stored and maintained in a data repository and consist of raw files transmitted between CMS, the implementation contractor, and the vendor. Aggregate program reports and ad hoc data queries can be run through the reporting portal interface and retrieved by authorized personnel. Manually entered HRA data are audited, and a minimum of 25 percent of the paper HRA data is entered and reviewed a second time. Another system quality assurance check occurs when HRA data are transferred into the data repository, which flags records that do not match a pre-coded algorithm.

Vendor B

Vendor B also developed a central database repository for the SRRD project that maintains all demographic and HRA data on program participants, and uses XML as the standard means of transmitting structured health risk assessment data to its clients.

ADRCs

The ADRCs each received contact and basic demographic information on local program participants from the vendors, which was used to conduct outbound calls. Their data management systems were used by staff to track, through case logs, all contacts and services provided to the SRRD participants. Staff entered detailed notes that provide a narrative summary of each call, including a plan of action and follow-up. The ADRCs have access to comprehensive databases on local community health programs and on resources and services that SRRD participants can be referred to, which are updated on a regular basis.

III. DATA SOURCES

The SRRD evaluation relies on two major types of data: (1) Medicare administrative data that include the Medicare claims files, denominator file, chronic condition summary file, and model output file; and (2) vendor data that include beneficiary responses to the vendor-created HRAs.

A. Overview of Data Sources

This section provides an overview of data sources. Exhibit 3.1 lists each data source and provides summary information.

Exhibit 3.1: Overview of Data Sources

Medicare or Vendor	Data File	Description
Medicare Administrative Data[a]	Inpatient Claims Files for baseline and Year 1	The Medicare claims files are composed of several files containing detailed information on the claims of beneficiaries in fee-for-service Medicare, including date of service, diagnosis and procedure codes, and payment amount.
	Outpatient Claims Files for baseline and Year 1	
	Carrier Claims Files for baseline and Year 1	
Medicare Administrative Data	Denominator File for 2008 and 2009	The denominator file contains data on all Medicare beneficiaries, including birth and death dates, gender, race, dual eligibility, and Part D status.
Medicare Administrative Data	Chronic Condition Summary File for 2008 and 2009	The chronic condition summary file includes flags indicating whether each beneficiary had a particular chronic condition.
Medicare Administrative Data	Model Output File (MOF) for 2008 and 2009	The MOF is used for risk-adjusting Medicare Advantage (Part C) payments. This file includes various risk scores for each Medicare beneficiary regardless of whether beneficiary is in fee-for-service Medicare or Medicare Advantage. These risk scores are based on demographics and diagnosis codes in the year prior to the payment year. The risk score of interest is the community risk score.
Vendor Files	Health Risk Assessment (HRA) Data	Participants in Arms 1, 2, and 3 completed HRAs in Year 1 and/or Year 2. Variables from the HRA data were used as outcomes in the analysis. Outcomes included measures such as smoking risk, diet risk, exercise risk, self-rated health status, and overall health risk.

[a] Baseline: January 1, 2008–April 30, 2009. Year 1: May 1, 2009–April 30, 2010.

Medicare FFS claims files are routinely used to conduct research on a variety of subjects such as quality of care, access, health services use, and cost. Key advantages of the claims files are their availability and uniform measurement for all Medicare FFS beneficiaries. The baseline period

for the claims data is from January 1, 2008 to April 30, 2009. The intervention year (Year 1) is from May 1, 2009 to April 30, 2010.

In addition, we used information from the HRAs to augment our analyses by examining the impact of the SRRD on risky health behaviors, dimensions that the claims data do not capture. For example, both vendors' HRAs capture alcohol, tobacco, physical activity, and diet risk levels (among others), based on respondents' self-reported behavior patterns. While we might observe increased costs or health services use in the claims data for individuals with higher HRA risk levels, it would be difficult to discern the reason for these patterns without a very detailed and resource-intensive investigation into diagnosis and procedure codes. In addition, the HRAs may provide information on the pathways through which the SRRD operates. For example, SRRD-induced improvements in behavior (e.g., improved preventive care) might take years to translate into changes in costs and use that could be captured in the claims data. However, comparing individuals' HRA data pre- and post-intervention would allow us to investigate whether the SRRD caused changes in health risk behaviors (e.g., smoking or exercise) that are linked to improved health outcomes.

Drawbacks of the HRA data in comparison to the claims data include smaller sample size and attrition. Among the target sample (n=20,000 for each vendor), the HRAs were only administered to the beneficiaries who agreed to participate in SRRD. This consisted of 7,473 participants for Vendor A and 7,375 participants for Vendor B (see Exhibit 3.2). Moreover, in order to use outcome measures based on the HRA data, we had to have two HRAs completed by a beneficiary: one at the beginning of the first year of the demonstration (Year 1 HRA) and one at the beginning of the second year of the demonstration (Year 2 HRA). Approximately 60 percent of beneficiaries who had a Year 1 HRA also completed a Year 2 HRA (see Exhibit 3.2). If the individuals who failed to complete a second HRA are systematically different from those who completed two HRAs (an issue investigated in Section V), any results could be driven, at least in part, by the unstable sample. While small sample size makes it difficult to conduct meaningful subgroup analyses using HRA outcomes, attrition (a participant's decision not to complete a Year 2 HRA) introduces the possibility of biased results. The claims data do not have similar attrition issues and are available for the entire intervention and administrative control groups.

Exhibit 3.2: Year 1 and Year 2 Participant Counts by Treatment Arm

	Vendor A					Vendor B				
	Year 1	Year 2	Re-Enrollees[a]	Re-Enrollee Rate[b]	New Enrollees	Year 1	Year 2	Re-Enrollees[a]	Re-Enrollee Rate[b]	New Enrollees
Arm 1	2,489	1,833	1,365	54.8%	468	2,458	1,792	1,299	52.8%	493
Arm 2	2,495	1,797	1,326	53.1%	471	2,459	1,722	1,232	50.1%	490
Arm 3	2,489	2,203	1,742	70.0%	461	2,458	2,115	1,623	66.0%	492
All Arms	7,473	5,833	4,433	59.3%	1,400	7,375	5,629	4,154	56.3%	1,475

[a] Re-enrollees returned HRAs in both Year 1 and Year 2.

[b] The re-enrollee rate was calculated by dividing the number of re-enrollees by the number of Year 1 participants.

The HRAs used by Vendor A and Vendor B are similar and gather the following information:

- Presence of chronic diseases
- Self-rated health status
- Health indicators such as obesity, blood pressure, and cholesterol levels
- Physical activity level
- Diet
- Tobacco use
- Alcohol consumption
- Feelings of depression
- Stress level
- Medication use
- Activity limitations resulting from physical health
- Barriers to obtaining medical care
- Preventive care
- Perceived importance of making healthy lifestyle changes[14]
- Major barriers to making healthy lifestyle changes[15]
- Confidence in ability to make healthy lifestyle changes
- Stage of change in making healthy lifestyle changes
- Risk levels based on proprietary algorithms
    - Physical activity
    - Nutrition
    - Smoking
    - Alcohol
    - Blood pressure
    - Glucose
    - Cholesterol
    - Obesity
    - Clinical preventive services
    - Depression
    - Stress
    - General well-being
    - Caregiving burden
    - Social isolation

[14] Vendor B only. Questions to assess readiness to change are included in Vendor B's HRA tool. Reponses to these questions help to tailor HRA feedback not only to risk level, but also to state of change.
[15] Vendor B only.

- Motor vehicle and home safety
- Falls
- Polypharmacy
- Back[16]

- Overall risk stratum[17]
- Current lifestyle score[18]

B. Examination of HRA Data Collected for SRRD

Prior to conducting the impact analysis, we examined the internal validity of the data generated using the HRAs. We also compared some key elements of the HRA data to national benchmarks. Exhibits 3.3 and 3.4 present analyses conducted to determine internal validity. The exhibits display pairwise correlation coefficients between self-rated general health status and a number of variables expected to correlate with health status for Vendor A and Vendor B's Year 1 samples, respectively. The results from Vendor A's sample indicate that in all cases the correlations are statistically significant and have the hypothesized signs. Individuals with higher body mass index (BMI) tend to have lower general health status, while beneficiaries engaging in more physical activity, eating a healthier diet, making healthy lifestyle changes, and having higher educational attainment tend to have better general health status.

Data from the health risk assessment instruments display internal validity: relationships among key variables are as anticipated (e.g., self-rated health status and body mass index are statistically significantly negatively related).

The results for Vendor B are similar. BMI and stress level are negatively related to general health, while activity level, diet, and social engagement are positively correlated with health status. With one exception (frequency of contact with close relatives), all correlations were statistically significant. The fact that the correlations among key variables are as hypothesized bolsters our confidence in the HRA instruments for use as an evaluation tool.

In addition to investigating internal consistency, we compared the HRA data to nationally representative benchmarks. In particular, we examined whether data from each vendor's national sample (chosen to be nationally representative) aligned with nationally representative data on gender and race/ethnicity composition, obesity, and tobacco use. We used the Census Bureau's American Community Survey (ACS) and the CDC's Behavioral Risk Factor Surveillance System (BRFSS) as benchmarks since they are nationally representative databases. The ACS data are from 2009 and were collected from the Census Bureau's American FactFinder Website.[19]

[16] Vendor B only.
[17] Vendor A only.
[18] Vendor B only.
[19] Accessed November 11, 2010 from

The BRFSS data are for 2005 and were compiled using CDC's online cross tabulation tool.[20] We used the Year 1 HRA data for these tabulations. Exhibits 3.5 through 3.10 display the results from comparisons of each vendor's national sample and nationally representative benchmarks.

Compared to national benchmarks, individuals completing HRAs:

- Were more likely to be White
- Were more likely to be overweight or obese
- Were similar in their tobacco use for most age/race/ethnicity strata

Exhibit 3.3: Correlations between Self-Rated General Health Status and Other Covariates – Vendor A

Variable	Correlation Coefficient[a]
Body mass index	-0.1922
Moderate physical activity days per week[b]	0.2873
Vigorous physical activity days per week[b]	0.2237
Number of fruit servings per day[b]	0.1624
Number of vegetable servings per day[c]	0.1517
Status of lifestyle change – get more exercise/physical activity[d]	0.3133
Status of lifestyle change – eat healthier[d]	0.1976
Status of lifestyle change – deal with stress better[d]	0.1423
Status of lifestyle change – if you use tobacco, are you planning to quit[d]	0.1261
Educational attainment[e]	0.3517

[a] All correlations have p <.001.

[b] The respondent was asked the following question: "For the last 3 months, how many days per week did you usually do moderate/vigorous physical activity?" Responses took on values of 0–7, corresponding to the number of days per week.

[c] The respondent was asked the following question: "On average, how many servings of fruit do you eat each day?" Response options were 0 (less than 1 serving), and 1, 2, 3, and 4, corresponding to 1, 2, 3, and 4 or more servings, respectively.

[d] The respondent was asked the following question: "Are you thinking about, or have you started, making changes in the following areas?" Response options were: 1 – I have no plans to do this, "2 – I plan to do this in the next 6 months," "3 – I plan to do this in the next 30 days," and "4 – I already do this."

[e] The respondent was asked the following question: "What is the highest degree or level of school you completed?" Response options were: "1 – Never attended school or attended kindergarten only," "2 – Some elementary or middle school (grades 1–8)," "3 – Some high school (grades 9–11)," "4 – High school graduate (grade 12)," "5 – Some college or technical school (1–3 years of college)," "6 – College graduate or higher (4 or more years of college)."

http://factfinder.census.gov/servlet/DTGeoSearchByListServlet?ds_name=ACS_2009_1YR_G00_&_lang=en&_ts=308670961655

[20] Accessed November 11, 2010 from http://apps.nccd.cdc.gov/s_broker/htmsql.exe/weat/Select_Year.hsql?Analysis=freq&Title=Cross Tabulation

Exhibit 3.4: Correlations between Self-Rated General Health Status and Other Covariates – Vendor B

	Correlation Coefficient[a]
Body mass index	-0.1869
Days per week engaging in moderate-intensity physical activity	0.301
Days per week engaging in vigorous exercise	0.2931
Fruit servings per day	0.0573
Vegetable servings per day	0.0441
Participation in social activities	0.1729
Frequency of contact with close friends	0.0588
Frequency of contact with close relatives	**0.0079**
Stress level in life	-0.2995

[a] All correlation coefficients are statistically significant, with $p < .001$, except for the figure in bold, which is not significant at the 10% level.

Exhibits 3.5 and 3.6 indicate that both vendors' national samples have larger proportions of Whites and smaller proportions of most minorities when compared to the national benchmark. This is particularly true of Vendor B's sample, which is more than 90 percent White (vs. national benchmark of 78 percent) and less than 1 percent Hispanic (vs. benchmark of 7.5 percent). The gender makeup of Vendor B's sample is not statistically distinguishable from the national benchmark. Vendor A's sample is more heavily female (56.2 percent vs. 54.2 percent benchmark).

Exhibits 3.7 and 3.8 display the results of comparisons between the vendors' national samples and nationally representative benchmarks for obesity. There is no statistically significant difference in the proportion of overweight individuals between Vendor A's sample and the benchmark. However, the overall obesity rate (last row of Exhibit 3.7) is higher for Vendor A's sample than for the benchmark, and it is also higher among Whites. Overall, individuals in Vendor B's national sample are more likely to be overweight than is suggested by the nationally representative benchmark, and the same is true for Whites. In addition, the obesity rate in Vendor B's sample is higher overall and for females and Whites. Taken together, there is evidence that the vendors' samples tend to have higher obesity rates than the national benchmark, and this difference is concentrated among Whites.

Exhibit 3.5: Race/Ethnicity Comparison of HRA Data to National Benchmarks – Vendor A

Race/Ethnicity & Gender	HRA National Sample %	National Benchmark (ACS, 2009[a]) %
Male[b]	43.8 (42.6 - 45.1)[d]	45.8 (45.6 - 46.1)
Female[b]	56.2 (54.9 - 57.4)	54.2 (53.8 - 54.5)
White[c]	87.4 (86.6 - 88.3)	78.2 (78.2 - 78.3)
White male[c]	38.9 (37.7 - 40.1)	36.6 (36.6 - 36.6)
White female[c]	48.6 (47.3 - 49.8)	41.6 (41.6 - 41.6)
Hispanic or Latino[c]	3.4 (2.9 - 3.8)	7.5 (7.5 - 7.6)
Hispanic male[c]	1.4 (1.1 - 1.7)	3.4 (3.4 - 3.4)
Hispanic female[c]	2.0 (1.6 - 2.3)	4.1 (4.1 - 4.1)
Black or African American[c]	5.8 (5.2 - 6.4)	9.3 (9.2 - 9.3)
Black male[c]	2.1 (1.8 - 2.5)	3.9 (3.8 - 3.9)
Black female[c]	3.7 (3.2 - 4.1)	5.4 (5.4 - 5.4)
Asian[c]	2.1 (1.8 - 2.5)	3.7 (3.7 - 3.8)
Asian male[c]	0.8 (0.6 - 1.0)	1.7 (1.7 - 1.7)
Asian female[c]	1.4 (1.1 - 1.7)	2.0 (2.0 - 2.1)
Native Hawaiian or Other Pacific Islander[c]	0.06 (0 - 0.12)	0.08 (0.1 - 0.1)
Native Hawaiian/Pacific Islander male[c]	0.03 (0 - 0.07)	0.04 (0.0 - 0.0)
Native Hawaiian/Pacific Islander female[c]	0.03 (0 - 0.07)	0.04 (0.0 - 0.0)
American Indian or Alaska Native[c]	0.5 (0.3 - 0.6)	0.6 (0.5 - 0.6)
American Indian/Alaska Native male[c]	0.3 (0.2 - 0.4)	0.3 (0.2 - 0.3)
American Indian/Alaska Native female[c]	0.2 (0.1 - 0.3)	0.3 (0.3 - 0.3)

[a] Benchmark data are from the detailed American Community Survey tables based on 2009 data.
[b] Population for benchmark data includes individuals aged 67-74.
[c] Population for benchmark data includes individuals aged 65-74.
[d] Figures in parentheses are 5% confidence intervals for the HRA data and 10% confidence intervals for the ACS.

Exhibit 3.6: Race/Ethnicity Comparison of HRA Data to National Benchmarks – Vendor B

Race/Ethnicity & Gender	HRA National Sample %	National Benchmark (ACS, 2009[a]) %
Male[b]	44.39 (43.1 - 45.7)[d]	45.8 (45.6 - 46.1)
Female[b]	55.61 (54.3 - 56.9)	54.2 (53.8 - 54.5)
White[c]	90.15 (89.4 - 90.9)	78.2 (78.2 - 78.3)
White male[c]	40.56 (39.3 - 41.8)	36.6 (36.6 - 36.6)
White female[c]	49.59 (48.3 - 50.9)	41.6 (41.6 - 41.6)
Hispanic or Latino[c]	0.72 (0.5 - 0.9)	7.5 (7.5 - 7.6)
Hispanic male[c]	0.32 (0.2 - 0.5)	3.4 (3.4 - 3.4)
Hispanic female[c]	0.39 (0.2 - 0.6)	4.1 (4.1 - 4.1)
Black or African American[c]	6.41 (5.8 - 7.0)	9.3 (9.2 - 9.3)
Black male[c]	2.23 (1.9 - 2.6)	3.9 (3.8 - 3.9)
Black female[c]	4.18 (3.7 - 4.7)	5.4 (5.4 - 5.4)
Asian[c]	1.02 (0.8 - 1.3)	3.7 (3.7 - 3.8)
Asian male[c]	0.41 (0.3 - 0.6)	1.7 (1.7 - 1.7)
Asian female[c]	0.61 (0.4 - 0.8)	2 (2.0 - 2.1)
American Indian or Alaska Native[c]	0.19 (0.1 - 0.3)	0.6 (0.5 - 0.6)
American Indian/Alaska Native male[c]	0.14 (0 - 0.2)	0.3 (0.2 - 0.3)
American Indian/Alaska Native female[c]	0.05 (0 - 0.1)	0.3 (0.3 - 0.3)

[a] Benchmark data are from the detailed American Community Survey tables based on 2009 data.
[b] Population for benchmark data includes individuals aged 67-74.
[c] Population for benchmark data includes individuals aged 65-74.
[d] Figures in parentheses are 5% confidence intervals for the HRA data and 10% confidence intervals for the ACS.

Exhibit 3.7: Obesity Comparison of HRA Data to National Benchmarks – Vendor A[a, b]

Race/Ethnicity and Gender	% Overweight		% Obese	
	HRA 67-74	BRFSS 65-74	HRA 67-74	BRFSS 65-74
Males	46.5 (44.6 - 48.4)	49.0 (47.0 - 51.1)	27.5 (25.8 - 29.1)	24.2 (22.5 - 25.9)
Females	33.5 (32.0 - 35.1)	34.8 (33.3 - 36.3)	29.7 (28.2 - 31.2)	27.4 (26.0 - 28.8)
White	39.5 (38.2 - 40.8)	41.8 (40.6 - 43.0)	28.4 (27.2 - 29.6)	24.6 (23.6 - 25.6)
White male	46.8 (44.9 - 48.8)	49.7 (47.7 - 51.5)	27.5 (25.7 - 29.3)	24.0 (22.4 - 25.6)
White female	33.6 (31.9 - 35.3)	34.9 (33.4 - 36.3)	29.1 (27.5 - 30.7)	25.2 (23.8 - 26.5)
Hispanic	42 (35.2 - 48.7)	45.5 (37.8 - 53.2)	34.1 (27.7 - 40.6)	28.0 (21.5 - 34.5)
Hispanic or Latino male	50 (39.6 - 60.4)	53.0 (40.8 - 65.1)	27.3 (18.0 - 36.6)	24.4 (14.3 - 34.4)
Hispanic or Latino female	35.9 (27.2 - 44.6)	39.3 (30.7 - 48.0)	39.3 (30.5 - 48.2)	31.0 (22.9 - 39.1)
Black or African American	34.8 (29.8 - 39.8)	36.3 (31.2 - 41.3)	42 (36.8 - 47.1)	38.5 (33.2 - 43.7)
Black or African American male	40.6 (32.3 - 48.9)	40.9 (32.5 - 49.4)	35.3 (27.2 - 43.5)	28.6 (20.8 - 36.2)
Black or African American female	31.2 (25.0 - 37.4)	32.8 (26.8 - 38.9)	46 (39.4 - 52.7)	45.6 (38.9 - 52.3)
Asian	32.1 (24.2 - 40.0)	38.7 (20.8 - 56.5)	5.2 (1.5 - 9.0)	8.5 (1.0 - 16.3)
Asian male	39.6 (25.7 - 53.4)	42.8 (16.3 - 69.3)	6.3 (0 - 13.1)	9.4 (0.1 - 20.2)
Asian female	27.9 (18.4 - 37.4)	34.1 (15.3 - 52.8)	4.7 (0.2 - 9.1)	7.1 (0.0 - 15.9)
Total	39.3 (38.0 - 40.5)	41.4 (40.2 - 42.8)	28.7 (27.6 - 29.9)	25.9 (24.8 - 27.0)

[a] The 95% confidence intervals are shown in parentheses.

[b] The BRFSS data are available from CDC's Web-Enabled Analysis Tool (WEAT) in 5-year age increments. We combined the 65-69 and 70-74 age strata using a weighted average based on the ACS data.

Exhibit 3.8: Obesity Comparison of HRA Data to National Benchmarks – Vendor B[a, b]

Race/Ethnicity	% Overweight		% Obese	
	HRA 67-74	BRFSS 65-74	HRA 67-74	BRFSS 65-74
Males	46.3 (44.3 - 48.2)	49.0 (47.0 - 51.1)	27.3 (25.6 - 29.0)	24.2 (22.5 - 25.9)
Females	32.7 (31.1 - 34.3)	34.8 (33.3 - 36.3)	30.8 (29.2 - 32.4)	27.4 (26.0 - 28.8)
White	38.7 (37.4 - 40.0)	41.8 (40.6 - 43.0)	29.1 (27.8 - 30.3)	24.6 (23.6 - 25.6)
White male	46.4 (44.4 - 48.4)	49.7 (47.7 - 51.5)	27.9 (26.1 - 29.7)	24.0 (22.4 - 25.6)
White female	32.4 (30.7 - 34.1)	34.9 (33.4 - 36.3)	30.1 (28.4 - 31.7)	25.2 (23.8 - 26.5)
Hispanic	38.1 (23.4 - 52.8)	45.5 (37.8 - 53.2)	35.7 (21.2 - 50.2)	28.0 (21.5 - 34.5)
Hispanic male	47.4 (24.9 - 69.8)	53.0 (40.8 - 65.1)	31.6 (10.7 - 52.5)	24.4 (14.3 - 34.4)
Hispanic female	30.4 (11.6 - 49.2)	39.3 (30.7 - 48.0)	39.1 (19.2 - 59.1)	31.0 (22.9 - 39.1)
Black or African American	39.4 (34.4 - 44.3)	36.3 (31.2 - 41.3)	37.5 (32.6 - 42.4)	38.5 (33.2 - 43.7)
Black male	45 (36.5 - 53.6)	40.9 (32.5 - 49.4)	23.7 (16.4 - 30.9)	28.6 (20.8 - 36.2)
Black female	36.3 (30.3 - 42.3)	32.8 (26.8 - 38.9)	44.9 (38.7 - 51.1)	45.6 (38.9 - 52.3)
Asian	31.7 (19.9 - 43.4)	38.7 (20.8 - 56.5)	11.7 (3.5 - 19.8)	8.5 (1.0 - 16.3)
Asian male	29.2 (11.0 - 47.4)	42.8 (16.3 - 69.3)	12.5 (0 - 25.7)	9.4 (0.1 - 20.2)
Asian female	33.3 (17.9 - 48.7)	34.1 (15.3 - 52.8)	11.1 (0.8 - 21.4)	7.1 (0.0 - 15.9)
Total	38.7 (37.5 - 40.0)	41.4 (40.2 - 42.8)	29.3 (28.1 - 30.4)	25.9 (24.8 - 27.0)

[a] The 95% confidence intervals are shown in parentheses.

[b] The BRFSS data are available from CDC's Web-Enabled Analysis Tool (WEAT) in 5-year age increments. We combined the 65-69 and 70-74 age strata using a weighted average based on the ACS data.

Exhibits 3.9 and 3.10 present comparisons of tobacco use across the HRAs and the national benchmark. White males in Vendor A's sample are more likely to be current smokers compared to the national benchmark. In addition, White women are less likely to be former smokers, and the proportion of individuals who never used tobacco is higher in the Vendor A sample, with the difference concentrated among females, Whites, and White females. Compared to the benchmark, there are fewer current smokers in Vendor B's sample, and this difference is driven by the lower proportion of White smokers in the sample compared to the national benchmark. In summary, tobacco use in the vendors' samples is similar to that of a nationally representative sample for most age/race/ethnicity strata. However, there is evidence that, overall, individuals

in Vendor A's sample are less likely to have ever smoked, and individuals in Vendor B's sample are less likely to be current smokers.

Exhibit 3.9: Tobacco Use Comparison of HRA Data to National Benchmarks – Vendor A[a, b]

Race/ Ethnicity and Gender	Current Tobacco Users % of respondents (95% confidence interval)		Former Tobacco Users % of respondents (95% confidence interval)		Individuals Who Never Used Tobacco % of respondents (95% confidence interval)	
	HRA 67-74	BRFSS 65-74	HRA 67-74	BRFSS 65-74	HRA 67-74	BRFSS 65-74
Male	11.7 (10.5 - 12.9)	9.8 (8.5 - 11.0)	52.8 (51.0 - 54.7)	53.9 (51.9 - 55.9)	35.4 (33.7 - 37.2)	33.5 (31.6 - 35.5)
Female	7.5 (6.7 - 8.4)	8.8 (8.0 - 9.6)	33 (31.4 - 34.5)	34.7 (33.3 - 36.1)	59.5 (57.8 - 61.1)	53.4 (51.8 - 54.9)
White	9.5 (8.7 - 10.3)	9.1 (8.5 - 9.8)	42.9 (41.5 - 44.2)	45.9 (44.7 - 47.1)	47.7 (46.3 - 49.0)	42.2 (41.1 - 43.4)
White male	11.7 (10.4 - 13.0)	8.9 (7.9 - 10.0)	54 (52.0 - 56.0)	56.0 (54.1 - 57.9)	34.3 (32.4 - 36.2)	32.5 (30.6 - 34.2)
White female	7.7 (6.7 - 8.7)	9.3 (8.4 - 10.1)	33.9 (32.2 - 35.6)	37.4 (36.0 - 38.9)	58.4 (56.6 - 60.1)	50.4 (49.0 - 51.9)
Hispanic	8.2 (4.5 - 11.9)	7.4 (3.8 - 10.9)	27.4 (21.3 - 33.5)	30.8 (24.0 - 37.6)	64.4 (57.9 - 70.9)	59.7 (52.6 - 66.9)
Hispanic male	10.1 (3.8 - 16.4)	13.0 (4.4 - 21.5)	36 (26.0 - 45.9)	45.4 (33.9 - 57.0)	53.9 (43.6 - 64.3)	39.9 (28.8 - 50.9)
Hispanic female	6.7 (2.2 - 11.2)	3.5 (1.8 - 5.1)	21 (13.7 - 28.3)	18.9 (13.0 - 24.8)	72.3 (64.2 - 80.3)	75.9 (69.7 - 82.1)
Black/African American	10.7 (7.5 - 13.9)	10.3 (7.2 - 13.6)	42.5 (37.4 - 47.7)	37.0 (32.0 - 42.0)	46.8 (41.6 - 52.0)	46.0 (40.8 - 51.3)
Black male	14.5 (8.5 - 20.5)	13.3 (7.3 - 19.4)	50.4 (41.8 - 58.9)	49.2 (40.6 - 57.9)	35.1 (26.9 - 43.3)	30.0 (22.5 - 37.5)
Black female	8.5 (4.8 - 12.1)	8.3 (4.9 - 11.6)	37.9 (31.6 - 44.3)	28.4 (23.1 - 33.7)	53.6 (47.0 - 60.1)	57.2 (51.0 - 63.4)
Asian	3.7 (0.5 - 6.9)	4.2 (0.0 - 8.4)	17.9 (11.4 - 24.4)	28.0 (12.6 - 43.5)	78.4 (71.4 - 85.3)	66.5 (50.8 - 82.2)
Asian male	8.3 (0.5 - 16.2)	4.9 (0.0 - 11.5)	33.3 (20.0 - 46.7)	31.7 (9.0 - 54.6)	58.3 (44.4 - 72.3)	61.9 (38.5 - 85.1)
Asian female	1.2 (0 - 3.4)	3.1 (0.0 - 6.7)	9.3 (3.2 - 15.4)	22.3 (4.6 - 40.1)	89.5 (83.1 - 96.0)	73.5 (55.4 - 91.7)
Total	9.4 (8.7 - 10.1)	9.2 (8.5 - 9.9)	41.7 (40.5 - 42.9)	43.4 (42.2 - 44.6)	48.9 (47.7 - 50.2)	44.4 (43.1 - 45.6)

[a] The 95% confidence intervals are shown in parentheses.

[b] The BRFSS data are available from CDC's Web-Enabled Analysis Tool (WEAT) in 5-year age increments. We combined the 65-69 and 70-74 age strata using a weighted average based on the ACS data.

Exhibit 3.10: Tobacco Use Comparison of HRA Data to National Benchmarks – Vendor B[a, b]

Race/ Ethnicity and Gender	Current Tobacco Users % of respondents (95% confidence interval) HRA 67-74	BRFSS 65-74
Male	7.6 (6.6 - 8.6)	9.8 (8.5 - 11.0)
Female	7.7 (6.8 - 8.7)	8.8 (8.0 - 9.6)
White	7.5 (6.8 - 8.2)	9.1 (8.5 - 9.8)
White male	7.2 (6.2 - 8.3)	8.9 (7.9 - 10.0)
White female	7.7 (6.7 - 8.7)	9.3 (8.4 - 10.1)
Hispanic	12.2 (2.2 - 22.2)	7.4 (3.8 - 10.9)
Hispanic male	15.8 (0 - 32.2)	13.0 (4.4 - 21.5)
Hispanic female	9.1 (0 - 21.1)	3.5 (1.8 - 5.1)
Black/African American	11.4 (8.2 - 14.7)	10.3 (7.2 - 13.6)
Black male	14.1 (8.0 - 20.1)	13.3 (7.3 - 19.4)
Black female	10 (6.2 - 13.8)	8.3 (4.9 - 11.6)
Asian	1.7 (0 - 4.9)	4.2 (0.0 - 8.4)
Asian male	0 0	4.9 (0.0 - 11.5)
Asian female	2.8 (0 - 8.1)	3.1 (0.0 - 6.7)
Total	7.7 (7.0 - 8.4)	9.2 (8.5 - 9.9)

[a] The 95% confidence intervals are shown in parentheses.

[b] The BRFSS data are available from CDC's Web-Enabled Analysis Tool (WEAT) in 5-year age increments. We combined the 65-69 and 70-74 age strata using a weighted average based on the ACS data.

In summary, there is evidence that the vendors' national samples and nationally representative benchmarks differ:

- The samples tend to have lower proportions of minorities.
- The obesity rates are higher for the samples than for the national benchmark.
- The proportion of individuals who are former smokers is lower in Vendor A's sample than in the national benchmark.
- The proportion of current smokers in Vendor B's sample is lower than in the national benchmark.

IV. METHODOLOGY

In this section, we describe the methodology used to conduct the impact analysis. We discuss s four topics: participation analysis, outcome measures, randomization analysis, and impact analysis.

Summary:

- *Participation Analysis:* We describe our examination of the characteristics of beneficiaries who participated in the demonstration compared with the target population and opt-outs.
- *Outcome Measures:* We present and describe the outcomes used to measure the impact of the demonstration and provide baseline summary statistics.
- *Randomization Analysis:* We describe how we assessed whether the random assignment of beneficiaries successfully resulted in treatment and comparison groups that were statistically identical in terms of their baseline demographic and healthcare use characteristics.
- *Impact Analysis:* We present the regression models used to estimate differences in the outcome measures between treatment and control groups during the first year of the demonstration.

A. Participation Analysis

The participation analysis was designed to examine the characteristics of beneficiaries who participate in the SRRD. In particular, we considered the following questions:

1. How do participation rates differ across treatment arms? Do re-enrollment rates differ by arm? Do participation rates differ across vendors and local vs. national samples?
2. Do participation rates differ across beneficiary characteristics such as demographics, chronic condition burden, and baseline Medicare spending and use?
3. What are the characteristics of participating beneficiaries and opt-outs (individuals explicitly choosing not to participate in the demonstration), and how do their characteristics differ from those of the larger group from which they were recruited?

To address question 1, we generated participation rates for the various sample/vendor combinations (Vendor A national, Vendor A local, Vendor B national, and Vendor B local) by intervention arm. We conducted statistical significance tests to determine whether participation rates differ across these strata.

For question 2, we calculated participation rates for Year 1 participants (those having only a Year 1 HRA) and re-enrollees (those having both Year 1 and Year 2 HRAs), and present the rates separately by vendor for multiple strata that are policy and methodologically relevant:

- Gender
- Age
- Race/ethnicity
- Dual eligibility status
- Total Medicare expenditure
- Use of medical services covered under Medicare
- Presence of chronic conditions

Again, we used statistical inference to ascertain whether there were differences in participation rates across the strata.

Question 3 considers the characteristics of beneficiaries participating in (or opting out of) the demonstration compared to the broader group of individuals from which participants were recruited. We provide summary statistics for each of the strata listed above for the intervention group, Year 1 participants, and re-enrollees. We used statistical tests to determine whether there were significant differences between the participants and the pool of potential recruits for each of the dimensions.

B. Outcome Measures

To understand the impact of the SRRD interventions, we examined several types of outcomes:

- Medicare use (6 measures)
- Medicare expenditure (5 measures)
- Hospitalization rates for ambulatory care sensitive conditions (ACSC) (15 measures)
- Preventive screening use (3 measures)
- Selected risk measures from the vendors' HRAs (8 measures per vendor)

Exhibit 4.1 provides the list of the outcome measures and data sources. The outcomes for Medicare use, Medicare expenditure, preventive screening use, and hospitalization rates for ambulatory care sensitive conditions (ACSCs) were created from the seven Medicare claims files for Year 1 (May 1, 2009 to April 30, 2010). We also used HRA-based risk measures as developed by the vendors.

Exhibit 4.1: Outcome Variables

Category	Sub-Category	Measures	Data Source
Medicare use and expenditure history	Medicare use (any use and number of occurrences)	Any inpatient days Inpatient days Any outpatient Outpatient days[a] Carrier days[b] Emergency department visits	Medicare claims for Year 1
	Medicare expenditure (dollars)	Total Medicare payment Inpatient payment Outpatient payment Carrier payment Emergency department payment	Medicare claims for Year 1
Claims-based health outcome measures	Any hospitalizations for ambulatory care sensitive conditions	Diabetes short-term complication Perforated appendix Diabetes long term complication COPD[c] or adult asthma Hypertension Congestive heart failure Dehydration Bacterial pneumonia Urinary infection Angina Diabetes uncontrolled Lower extremity amputation Overall, Acute, and Chronic	Medicare claims for the Year 1
Health risk behavior	Preventive services use	Colorectal cancer screening Breast cancer screening Cardiovascular disease screening	Medicare claims for Year 1
HRAs	Vendor A	Self-rated health status Risk stratum Alcohol risk Diet risk Preventive services risk Physical activity risk Tobacco risk Polypharmacy risk	Vendor A HRA for participants in both Year 1 and Year 2
	Vendor B	Self-rated health status Current lifestyle risk Alcohol risk Eating risk Exams risk Exercise risk Smoking risk Self-care risk	Vendor B HRA for participants in both Year 1 and Year 2

[a] *Outpatient days* represents the number of days for which a patient had outpatient claims (e.g., two outpatient claims in one day would represent one outpatient day).
[b] *Carrier days* represents the number of days for which a patient had carrier claims (e.g., two carrier claims in one day would represent one carrier day). Carriers handle (non-institutional) Part B claims.
[c] Chronic obstructive pulmonary disease.

1. Outcome Variables from the Medicare Claims Data

Medicare use and expenditure measures captures the use and cost of services, overall and by each of several settings (inpatient, outpatient, and carrier). We also include the number and Medicare payments for emergency departments visits.

Hospitalizations for ambulatory care sensitive conditions is an indicator for whether the individual had any hospitalizations for conditions where appropriate ambulatory care could have prevented or reduced the need for admission to the hospital. While not all admissions for ACSCs are avoidable, there is evidence that appropriate ambulatory care could prevent the onset of this type of illness or condition, control an acute episodic illness or condition, or manage a chronic disease or condition. A high hospitalization rate for ACSCs is presumed to reflect problems in obtaining appropriate primary care.[21] We constructed an indicator for any hospitalizations for 12 ACSCs (out of the 15 ACSCs provided by AHRQ) using the Medicare claims data (see Exhibit 3.1).[22] We chose the 12 ACSCs that are most appropriate and relevant to this population and included overall, acute, and chronic ACSCs. The overall ACSC indicates whether there were hospitalizations for any of 11 ACSCs (excluding perforated appendix). The acute ACSC measures hospitalizations for any one of the three acute conditions: dehydration, bacterial pneumonia, and urinary infection. The chronic ACSC measures hospitalizations for any of the nine chronic conditions. Due to the small number of hospitalizations resulting from ACSCs among demonstration participants, however, we have not presented the results on the ACSC outcomes in this report (see Section VI).

Preventive screening measures includes binary indicators for beneficiaries who obtained colorectal cancer screening, breast cancer screening, and cardiovascular screening. We identified the set of preventive screenings recommended and covered by Medicare using three criteria:

1. The measure could readily be evaluated using Medicare claims.
2. Use of the preventive screening service was indicated for a large proportion of Medicare beneficiaries.
3. Use of the preventive screening service was associated in some way, directly or indirectly, with the vendor interventions aimed at good health.

Using the Healthcare Common Procedure Coding System (HCPCS), the Current Procedural Terminology (CPT) codes and the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis codes, we identified these three preventive screenings in the Medicare administrative claims data. (See Appendix A for a list of codes and exclusions.) For each preventive screening measure, we excluded beneficiaries with that condition, which is

[21] See http://www.qualityindicators.ahrq.gov/Modules/pqi_overview.aspx for more information.
[22] For the technical specifications of each ACSC indicator, see http://www.qualityindicators.ahrq.gov/modules/PQI_TechSpec.aspx

consistent with screening for undiagnosed cases of the disease.[23] For example, we calculated colorectal cancer screening rates on the population of beneficiaries without a known diagnosis of colorectal cancer, as identified in the CCW chronic condition files for 2009. Beneficiaries were considered as having been screened for colorectal cancer, for example, if any one of several services was used (fecal occult blood test annually, flexible sigmoidoscopy every 5 years, or colonoscopy every 10 years).

The acceptable periodicity varies for each of these colorectal cancer screening modalities. Our analysis involved searching the seven Medicare claims files[24] for evidence of any of these screening modalities. Since our method did not include review of 10 years of claims data, it is possible that we have underreported rates of colorectal cancer screening for patients screened with colonoscopy 10 years earlier. However, if such underreporting were present, randomization would ensure that this bias would be distributed comparably across both treatment and control groups. Indeed, we did not find screening rates to vary across treatment and control groups during the baseline period; therefore, any underreporting would not introduce bias.

Exhibits 4.2 and 4.3 show the claims outcomes summary statistics for the individuals in the ACG of Vendor A and Vendor B for Year 1. Means for the two groups are very similar. On average, individuals in the ACG of Vendors A and B spent $7,000 in total Medicare expenditures in Year 1. Approximately, 15 percent had at least one hospital stay, and the average hospital Medicare payment was $2,800 in Year 1. Approximately, 50 percent of females without breast cancer had breast cancer screening, and 61 percent of individuals without cardiovascular disease had cardiovascular screening.

[23] CMS (2009). *The Guide to Medicare Preventive Services for Physicians, Providers, Suppliers, and Other Health Care Professionals*, 3rd ed.

[24] The seven Medicare claims file types are inpatient, outpatient, skilled nursing facility, hospice, home health agency, carrier, and durable medical equipment.

Exhibit 4.2: Vendor A Year 1 ACG Means

Outcome Variables	Vendor A N=20,000
Medicare Payment ($)	
Total Medicare Payments	7,041
Inpatient Payment	2,816
Outpatient Payment	1,182
Carrier Payment	2,129
ED Payment	98
Utilization	
Any Inpatient	15.0%
Inpatient Length of Stay	1.4
Any Outpatient	69.1%
Outpatient Days	3.8
Carrier Days	15.7
ED Visits	0.4
Preventive Screening[a]	
Colorectal Cancer	20.4%
Breast Cancer	52.1%
Cardiovascular	60.7%

[a] Preventive screening variables were calculated only for beneficiaries without a diagnosis of the disease in the past year. Breast cancer screening was calculated only for females.

Exhibit 4.3: Vendor B Year 1 ACG Means

Outcome Variables	Vendor B N=20,000
Medicare Payment ($)	
Total Medicare Payments	7,085
Inpatient Payment	2,798
Outpatient Payment	1,163
Carrier Payment	2,123
ED Payment	97
Utilization	
Any Inpatient	15.1%
Inpatient Length of Stay	1.4
Any Outpatient	67.2%
Outpatient Days	3.5
Carrier Days	15.8
ED Visits	0.4
Preventive Screening[a]	
Colorectal Cancer	19.9%
Breast Cancer	51.6%
Cardiovascular	60.6%

[a] Preventive screening variables were calculated only for beneficiaries without a diagnosis of the disease in the past year. Breast cancer screening was calculated only for females.

2. Outcome Variables from the HRAs

In general, the HRAs administered by Vendor A and Vendor B contained information on the following topics:

- Self-rated health status
- Chronic conditions (e.g., arthritis, cancer, high blood pressure, kidney disease)
- Obesity
- Blood pressure
- Cholesterol
- Lifestyle (e.g., physical activity level, diet, tobacco use, and alcohol use)
- Stress and emotional well being
- Prescription medication use
- Preventive care (e.g., primary healthcare provider, flu shot, last physical examination)
- Importance of and confidence in making healthy lifestyle changes
- Several variables capturing risk level (e.g., physical activity, nutrition, tobacco)

Selecting HRA variables for use in the impact analysis was a multistep, iterative process involving evaluation and clinical expertise. Through team discussions focusing on survey design and clinical considerations, we identified initial lists (one for each vendor) of HRA variables that were most likely aligned with an effect from SRRD, if such an effect were in fact present. These variables included most of the available vendor-calculated risk scores, blood pressure, self-rated general health status, and importance of and confidence in making healthy lifestyle changes. The initial lists included more than 20 HRA outcome variables for each vendor.

Next, we further reduced the list to include only the most relevant HRA outcome variables. Given the set of 38 claims outcome variables, as well as the number of regressions in the analysis (more than 30 per outcome), a feasible first-year analysis required that we use a more parsimonious set of HRA variables. In addition, including HRA variables in the analysis was more complicated than using claims-based variables because the HRA instruments are unique across vendors and require more data preparation than the claims outcomes, which are derived from data that are homogeneous across vendors. Based on team discussions, including consideration of evidence-based clinical principles, we developed shorter lists (one for each vendor) of HRA variables recommended for inclusion.

The final step in the HRA variable selection process involved examination of the data and preparation of the measures. The specific form that the HRA-based outcomes took depended partly on the distribution of the outcomes across individuals in the data. Examples of questions considered include:

- Should risk levels (defined as high, moderate, or low in the HRA data) be dichotomous or remain trichotomous?
- Should the self-rated general health status variables continue to be five-point (Vendor A) or six-point (Vendor B) scales, or should they be dichotomized or trichotomized? If they should be dichotomized or trichotomized, what is the "cut point" between good and poor health?
- Are index variables feasible?[25]
- Are missing data values truly missing (respondent did not know the answer or did not respond to the question) or are they "not applicable" responses?

Note that the vendors used scales based on proprietary algorithms, making it difficult to compare similar topics between the vendors. We did not combine HRA variables across vendors, but instead examined them separately.

At the conclusion of this process, we produced the final lists of HRA outcome variables, which contained eight variables for Vendor A and eight for Vendor B. Exhibits 4.4 and 4.5 list the HRA

[25] We considered developing indexes for importance of and confidence in making healthy lifestyle changes. However, in addition to the complicated interpretation of these measures, we found that the denominators (individuals at risk for the lifestyle dimensions to be changed) were insufficient for constructing a useful index. The possibility of using similar indexes will be further investigated for inclusion in the final impact analysis.

outcomes used in our analysis and the values coded for each response. Each table also provides the size of the sample for each outcome and the mean value or percentage high/medium risk.

Exhibit 4.4: Vendor A HRA Outcome Variables for Arm 3

Variable	Response	Sample Size[a]	Percent[b]
Self-rated health status	Poor/fair	1,659	16.9%
	Good		40.6%
	Very good/excellent		42.5%
Triage risk stratum	High/medium risk[c]	1,742	39.2%
Excess alcohol risk level	High/medium risk[c]	1,728	7.9%
Poor nutrition risk level	High/medium risk[c]	1,690	94.8%
Inappropriate use of clinical preventive services risk level	High/medium risk[c]	1,680	20.1%
Physical activity risk level	High/medium risk[c]	1,500	35.5%
Smoking risk level	High/medium risk[c]	1,673	13.0%
Polypharmacy risk	High/medium risk[c]	1,693	45.5%

[a] The sample consists of those participants that had non-missing values for the pertinent HRA variable in both Year 1 and Year 2.
[b] Percentages are Year 2 values for each variable.
[c] Low risk is the omitted category.

Exhibit 4.5: Vendor B HRA Outcome Variables for Arm 3

Variable	Response	Sample Size[a]	Percentage[b]
Self-rated health status (past 4 weeks)	Very poor/poor/fair	1,589	16.4%
	Good		29.7%
	Very good/excellent		53.9%
Current lifestyle score	High/medium risk[c]	1,623	69.1%
Alcohol risk score	High/medium risk[c]	1,562	11.3%
Eating risk score	High/medium risk[c]	1,601	65.6%
Exams risk score	High/medium risk[c]	1,623	62.0%
Physical activity risk score	High/medium risk[c]	1,594	42.4%
Tobacco use risk score	High/medium risk[c]	1,458	8.5%
Self-care risk score	High/medium risk[c]	1,590	7.0%

[a] The sample consists of those participants that had non-missing values for the pertinent HRA variable in both Year 1 and 2.
[b] Percentages are Year 2 values for each variable.
[c] Low risk is the omitted category.

C. Assessment of Randomizing Participants into Arms

We performed statistical tests to assess whether the random assignment of beneficiaries successfully resulted in treatment and comparison groups that were statistically identical in terms of their baseline demographic and health services use characteristics. We compared the means of Arm 1 and Arm 2 against Arm 3, as well as against each other. The following comparisons were made:

- Intervention Arm 1 versus Intervention Arm 3
- Intervention Arm 2 versus Intervention Arm 3
- Intervention Arms 1 and 2 pooled versus Intervention Arm 3
- Intervention Arm 1 versus Intervention Arm 2

If the randomization was not reliably implemented, we would expect to find statistically significant differences in baseline characteristics between the intervention arms.

We made these four comparisons for each of the demonstration samples, generating 65 comparisons. For each contrast, we examined whether the values of the baseline measures

were significantly different across groups. As a rule, if more than 5 percent of the baseline measures were found to be statistically different between the groups, then we would assume that the IG and ACG were not randomized reliably.

D. Impact Analysis

We used a regression-based approach to compare outcomes in the treatment group with outcomes in the comparison group during the intervention year. This enabled us to estimate the impact between treatment and control groups while controlling for baseline beneficiary characteristics. An additional benefit of the regression approach is the increase in the precision of estimates. [26]

We estimated the parameters using ordinary least squares, which models the outcome as a linear function of the predictors. We regressed the outcome of interest on an indicator for treatment versus control group status and a series of control variables using ordinary least squares (OLS) regression. Control variables included the following:

- Demographic characteristics
  - Age
  - Female (indicator variable)
  - Non-White (indicator variable)
  - Medicare-Medicaid dual eligibility (indicator variable)
- Baseline value (lagged value) of the dependent variable (see Exhibit 4.1)
- Total Medicare payments from the baseline year

For the outcomes generated from the Medicare claims data, we used only beneficiaries who completed the HRA in Year 1; these beneficiaries are termed the participants. The participants constituted 37.1 percent of the target intervention group (see Exhibit 5.1 in the next section). For outcomes obtained from the vendors' HRA data, we limited the beneficiaries to those who completed HRAs for both Year 1 and Year 2. Exhibits 4.4 and 4.5, above, show the sample sizes of the participants who completed HRAs in both Year 1 and Year 2 of the demonstration.

For each outcome and demonstration sample, we ran regressions comparing intervention arms. For claims-based outcome variables, the regression models pooled observations from Vendor A and Vendor B using interaction terms to measure impacts between intervention arms. For example, the following equation shows the regression model used to estimate the impacts of Vendor A and Vendor B participants on an outcome. In this regression, β_8 and β_9 represent the

[26] Other options for conducting the impact analysis are (1) taking the difference in means of treatment versus control groups during the intervention period or (2) using difference-in-differences. In addition to the regression-based approach, these two alternatives also produce unbiased impact estimates, though there are advantages to the regression approach. At the end of the impact analysis results section, we provide impact estimates based on all three methodologies for comparisons 26 and 32 (see Exhibit 4.6) for a subset of outcomes.

impacts of SRRD on Arm 1 and Arm 2 participants, respectively, versus Arm 3 participants in Vendor B's group. The impacts of SRRD on Vendor A's Arm 1 and Arm 2 participants versus Vendor A's Arm 3 participants are $\beta_8 + \beta_{10}$ and $\beta_9 + \beta_{11}$, respectively. Note that impact estimates are differences between Arm 1 and Arm 2 impacts and the Arm 3 impact since an Arm 3 indicator is omitted from the regression equation.

$$y_{i,t=2} = \beta_0 + \beta_1 y_{i,t=1} + \beta_2 totalpayment_{i,t=1} + \beta_3 age_{i,t=1} + \beta_4 female_i + \beta_5 nonwhite_i + \beta_6 dual_{i,t=1} + \beta_7 VendorA_i + \beta_8 arm\_1_i + \beta_9 arm\_2_i + \beta_{10} arm\_1_i * VendorA_i + \beta_{11} arm\_2_i * VendorA_i$$

Where:

$Y_{i,t=2}$ = the outcome of interest in Year 1 of the demonstration for participant i

$Y_{i,t=1}$ = the outcome of interest in the base period (January 1, 2008 to April 30, 2009)

$Totalpayment_{i,t-1}$ = total Medicare expenditures in the base period

$age_{i,t=1}$ = beneficiary's age in the base period

$female_i$ = an indicator for female versus male

$nonwhite_i$ = an indicator for being non-White

$dual_{i,t-1}$ = an indicator for dual eligibility

$VendorA_i$ = an indicator for Vendor A's participants versus Vendor B's participants

Arm_1_i = an indicator for Arm 1 participants

Arm_2_i = an indicator for Arm 2 participants

$Arm\_1 \times VendorA_i$ = an indicator for Vendor A's Arm 1 participants

$Arm\_2 \times VendorA_i$ = an indicator for Vendor A's Arm 2 participants

For the HRA outcome variables, we ran separate regressions for Vendor A and Vendor B.

Exhibit 4.6 lists each of the comparisons for each demonstration sample. Note that we combined Local 1 and Local 2 into a single local sample for each of the vendors.

Exhibit 4.6: Comparisons for Impact Analysis

Intervention Arm Comparisons		
1	Arm 1	Arm 3
2	Arm 2	Arm 3
3	Arms 1 and 2 pooled	Arm 3
4	National Arm 1	National Arm 3
5	National Arm 2	National Arm 3
6	National Arms 1 and 2 pooled	National Arm 3
7	Local Arm 1	Local Arm 3
8	Local Arm 2	Local Arm 3
9	Local Arms 1 and 2 pooled	Local Arm 3
10	Vendor A Arm 1	Vendor A Arm 3
11	Vendor B Arm 1	Vendor B Arm 3
12	Vendor A Arm 2	Vendor A Arm 3
13	Vendor B Arm 2	Vendor B Arm 3
14	Vendor A Arms 1 and 2 pooled	Vendor A Arm 3
15	Vendor B Arms 1 and 2 pooled	Vendor B Arm 3
16	National Vendor A Arm 1	National Vendor A Arm 3
17	National Vendor A Arm 2	National Vendor A Arm 3
18	National Vendor A Arms 1 and 2 pooled	National Vendor A Arm 3
19	Local Vendor A Arm 1	Local Vendor A Arm 3
20	Local Vendor A Arm 2	Local Vendor A Arm 3
21	Local Vendor A Arms 1 and 2 pooled	Local Vendor A Arm 3
22	National Vendor B Arm 1	National Vendor B Arm 3
23	National Vendor B Arm 2	National Vendor B Arm 3
24	National Vendor B Arms 1 and 2 pooled	National Vendor B Arm 3
25	Local Vendor B Arm 1	Local Vendor B Arm 3
26	Local Vendor B Arm 2	Local Vendor B Arm 3
27	Local Vendor B Arms 1 and 2 pooled	Local Vendor B Arm 3

V. PARTICIPATION ANALYSIS RESULTS

Summary of Results:

- Re-enrollment in Year 2 of the demonstration was highest among Arm 3 participants.
- Participation rates were:
  - Similar for women and men.
  - Higher among non-Hispanic Whites and non-dual eligibles
  - Similar across age groups and individuals with/without one or more inpatient days
  - Higher among individuals with cancer, osteoporosis, and arthritis
  - Higher for the middle 50% of spenders than for the upper and lower quartiles
- Compared to the intervention group from which participants were recruited, participants were less likely to be members of a minority group and to be dual eligible.
- There are two types of opt-outs: those who completed an HRA and those who did not.
  - Opt-outs who completed an HRA were more likely to have been randomized into Arms 1 or 2 and were more likely to be medium risk vs. low risk.
  - Both types of opt-outs were slightly older, more likely to be non-Hispanic Whites, and less likely to have certain chronic conditions and screenings.
  - Compared to participants, opt-outs with an HRA tended to have higher expenditures and utilization while those without an HRA had lower expenditures and utilization.

In this comprehensive evaluation of the SRRD, we sought to answer several questions related to the participation of beneficiaries in the demonstration. The questions we addressed and our main findings are summarized below; the rest of this section provides further detail.

- How do participation rates differ across treatment arms? Do re-enrollment rates differ by arm? Do participation rates differ across vendors and local vs. national samples?
  a. Year 1 participants' baseline (pre-SRRD) characteristics did not differ across treatment arms, indicating that the second randomization was successful.
  b. Re-enrollment in Year 2 of the demonstration was significantly higher among Arm 3 (comparison group) participants than among Arm 1 and Arm 2 participants. There are at least two hypotheses to explain the lower Year 2 take-up for the intervention arms (Arms 1 and 2): (1) participants found the intervention burdensome, and (2) the benefits of the intervention accrued quickly (so that wellness goals were achieved quickly) and participants considered continued enrollment to have small marginal benefit.

c. Participation rates generally did not differ by vendor or sample (national vs. local).

- Do participation rates differ across beneficiary characteristics such as demographics, chronic condition burden, baseline spending, and use of Medicare services?

    a. Participation was higher among non-Hispanic Whites and non-dual eligible beneficiaries.

    b. There were no statistically distinguishable differences in participation rates for males vs. females or participants aged 64–69 vs. participants aged 70–74.

    c. Participation rates were highest among beneficiaries with cancer, osteoporosis, and arthritis, and lowest among those with Alzheimer's disease.[27]

    d. The middle two quartiles of beneficiaries in terms of total Medicare expenditures had higher participation rates than beneficiaries with expenditures in the upper or lower quartiles.

    e. Participation rates among beneficiaries with no inpatient days did not differ from participation rates among beneficiaries with one or more inpatient days.

- What are the characteristics of participating beneficiaries and opt-outs, and how do their characteristics differ from those of the larger group from which they were recruited?

    a. Compared to the intervention group, both national and local sample Year 1 participants were:

        - More likely to be non-Hispanic Whites and to have preventive screenings
        - Less likely to be dual eligible and to have certain chronic conditions

    b. Although the characteristics of both vendors' re-enrollees were generally not statistically different from their Year 1 participants (an exception is that Vendor B's re-enrollees were more likely to be non-Hispanic Whites), the consistency of the signs of the differences between the characteristics of re-enrollees and those of Year 1 participants may indicate continuing selection issues.

    c. Compared to individuals who continued participating in the demonstration, those who opted out after completing the Year 1 HRA were more likely to have been randomized into Arms 1 or 2, more likely to be medium-risk, and less likely to be low-risk.

    d. Compared to participants, opt-outs with and those without completed Year 1 HRAs were slightly older, more likely to be non-Hispanic Whites, and less likely to have certain chronic conditions and preventive screenings.[28]

[27] One criterion for inclusion in the demonstration was that the beneficiary must be mentally capable of participating. Thus, finding low participation rates among beneficiaries with Alzheimer's disease was expected.

e. There were differences between opt-outs with and without a Year 1 HRA. In particular, compared to participants, opt-outs with an HRA tended to have higher expenditures and utilization while those without an HRA tended to have lower expenditures and utilization.

A. Participation Rates by Arm, Sample, and Vendor

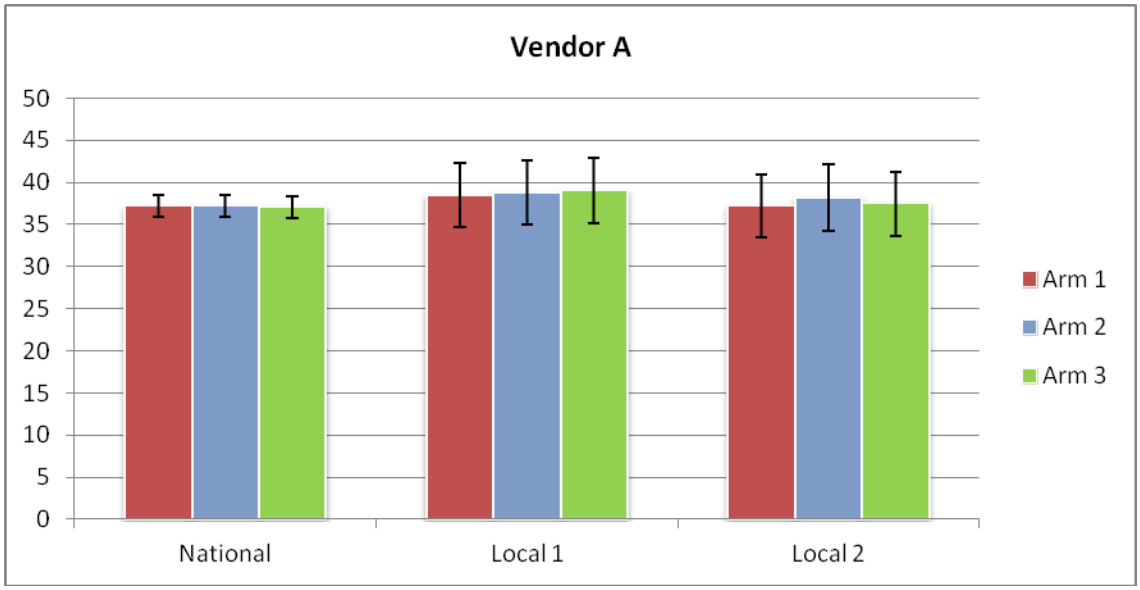
Exhibits 5.1 and 5.2 present participation rates for various samples (along the horizontal axis) for the three treatment arms (indicated by different colors and symbols on the plots) for individuals participating in Year 1 (Exhibit 5.1) and for those who both participated in both Years 1 and 2 (termed "re-enrollees"; Exhibit 5.2). The participation rate is defined as the number of individuals participating in the program, expressed as a proportion of eligible individuals (the same stratum in the IG group in the case of Year 1 participants, and Year 1 enrollees in the case of re-enrollees). The symbols in the exhibits indicate the point estimates of the participation rates, and the vertical bars show the 95 percent confidence intervals.

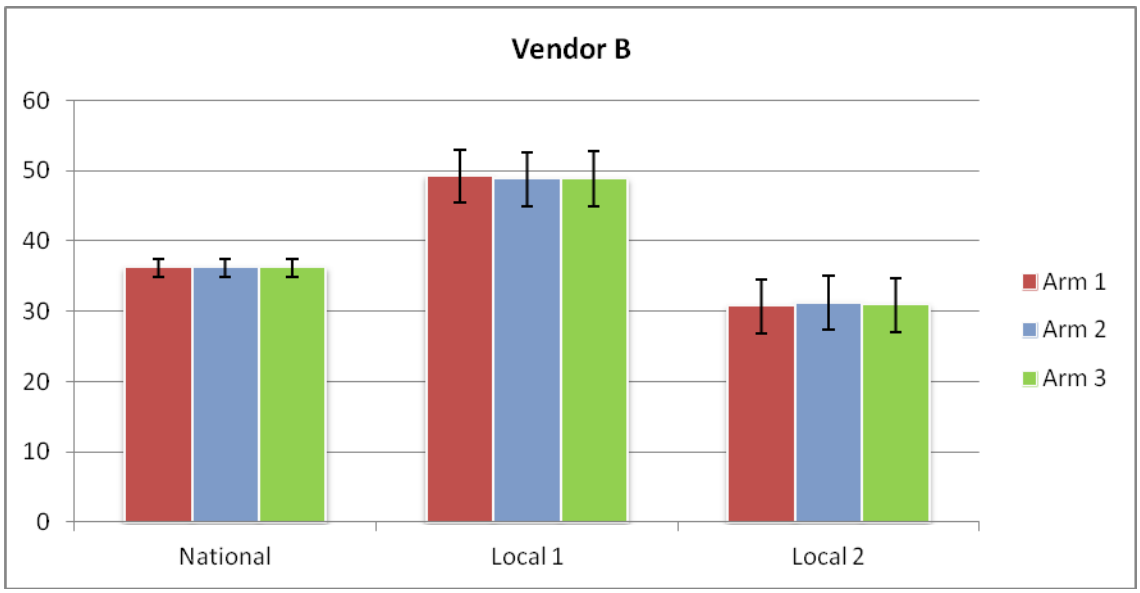
As evidenced by Exhibit 5.1, participation rates for Year 1 are statistically indistinguishable (confidence intervals overlap) across Arms 1, 2, and 3. This is expected because beneficiaries who agreed to participate were randomized into the treatment arms during the second randomization phase. In addition, participation rates did not differ across national versus local samples or by vendor. The only exception is that one of Vendor B's local samples has a statistically significantly higher participation rate than the rest of the samples.

Exhibit 5.2 shows participation rates for re-enrollees. In all cases except for one local Vendor B sample, Arm 3 participation rates exceeded those for Arm 1, and in all cases, Arm 3 participation rates exceeded Arm 2 participation rates. While the participation rates for Arm 1 were typically greater than those for Arm 2, only two of these differences were statistically significant. Arm 2 may have lower retention if (1) beneficiaries find participating in the intervention burdensome or (2) if the benefits of the intervention accrue quickly (so that wellness goals are achieved quickly) and participants consider continued enrollment to have small marginal benefit. The exhibit also shows that there were generally no differences in re-enrollee participation rates across vendors or samples.

[28] The results for chronic conditions and preventive screenings are statistically significant only for opt-outs without a Year 1 HRA. However, the results are consistent with those for the group of opt-outs with an HRA and are not statistically significant, probably because of the small sample size for this group (n=58).

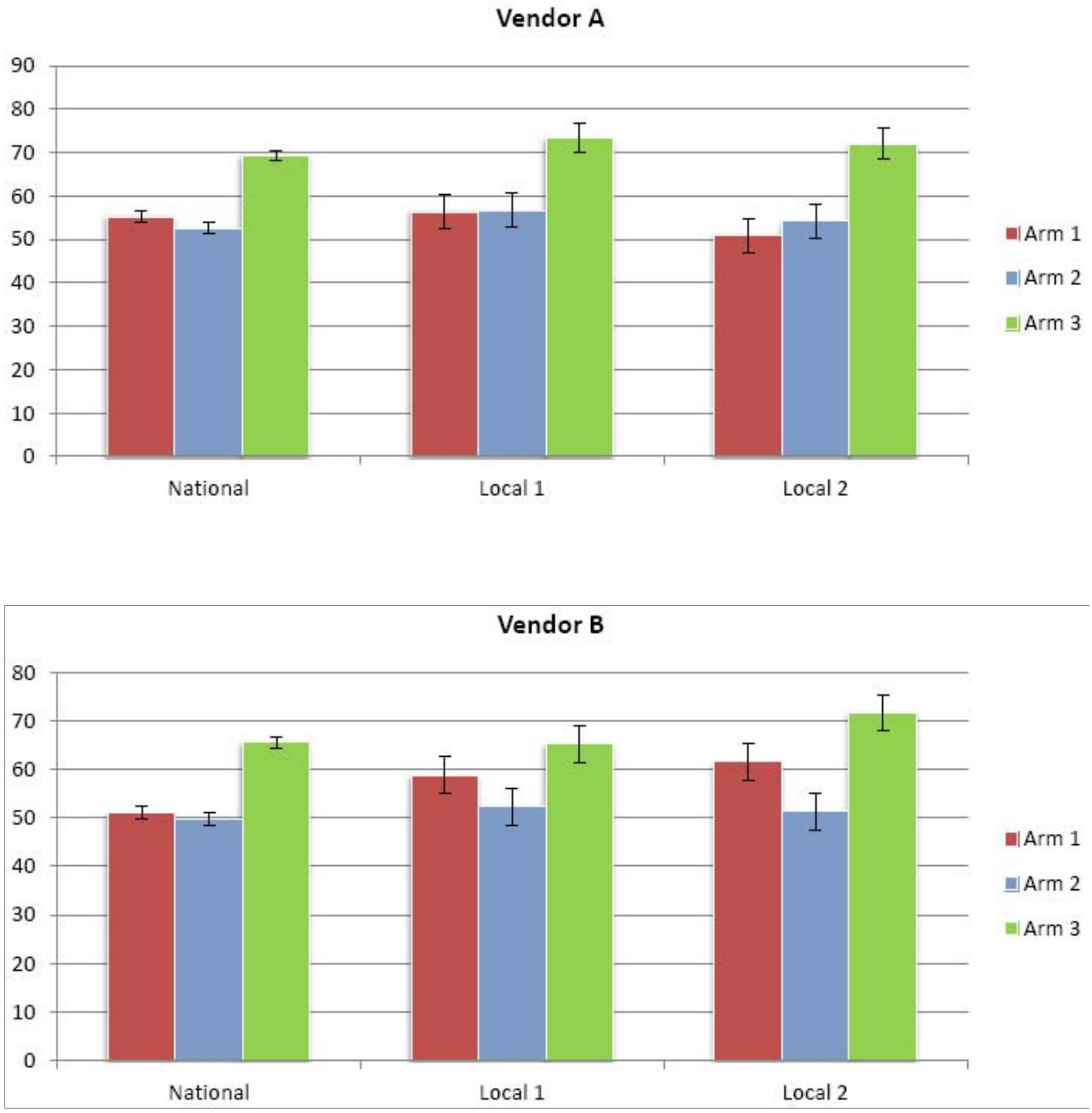
Exhibit 5.1: Year 1 Participation Rates by Vendor, National/Local Sample, and Arm





Note: The overall Year 1 participation rate across all arms, samples, and vendors was 37.1%.

Exhibit 5.2: Re-Enrollee Participation Rates by Vendor, National/Local Sample, and Arm



Note: The overall re-enrollee participation rate across all arms, samples, and vendors was 57.8%.

B. Participation Rates by Beneficiary Characteristics

Exhibits 5.3 through 5.8 display participation rates for beneficiary demographic characteristics (along the horizontal axes) by Year 1 participation (circles) and any participation (participating in Year 1 and/or Year 2; squares). The ever-enrolled (Year 1 and/or Year 2) participation rates are necessarily higher than those for Year 1 enrollment since "ever enrolled" includes Year 1 plus Year 2 participants and a number of individuals who began participating in Year 2.

Exhibits 5.3 and 5.4 display participation rates disaggregated by various demographic characteristics. For both vendors, there were no significant differences for males versus females or participants aged 65–69 versus participants aged 70–74. However, participation was significantly higher for non-Hispanic Whites than for other race/ethnicity categories, and for non-dual-eligible beneficiaries versus dual eligibles.

Exhibit 5.3: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Demographics – Vendor A

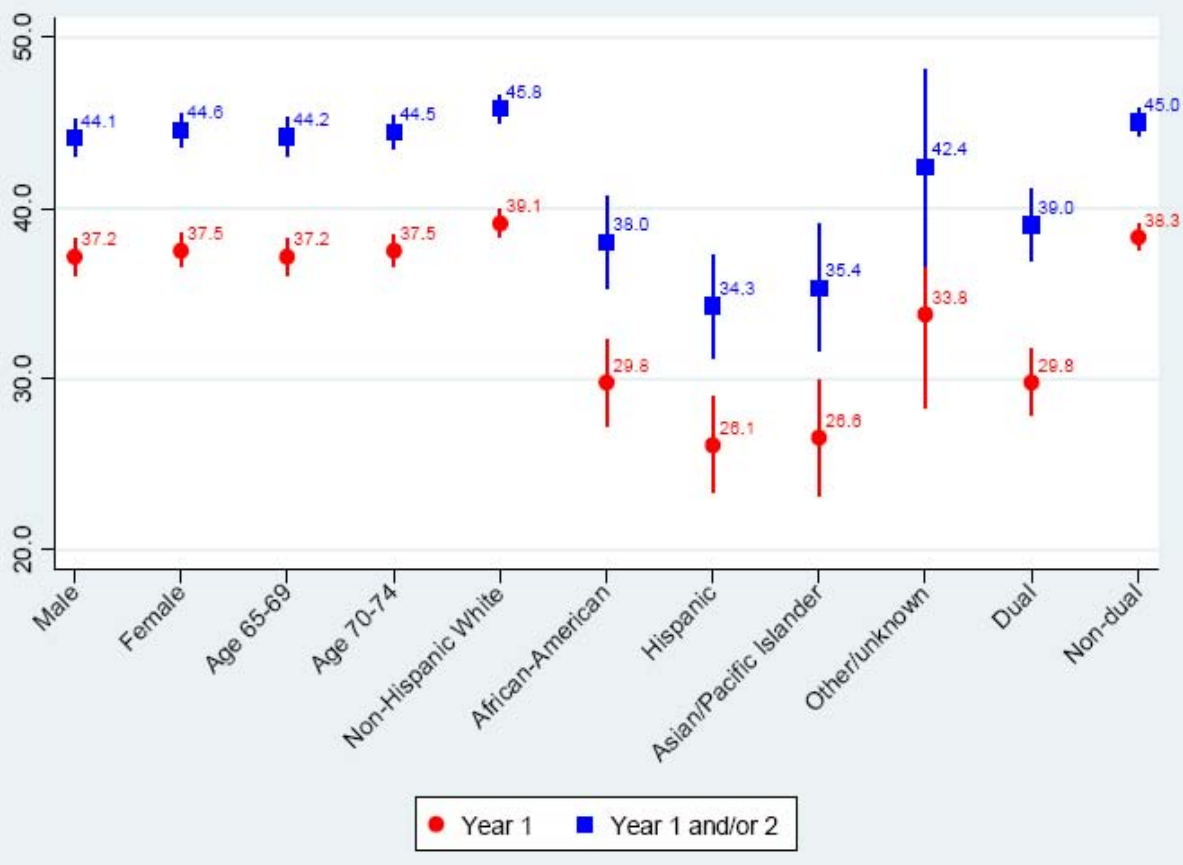
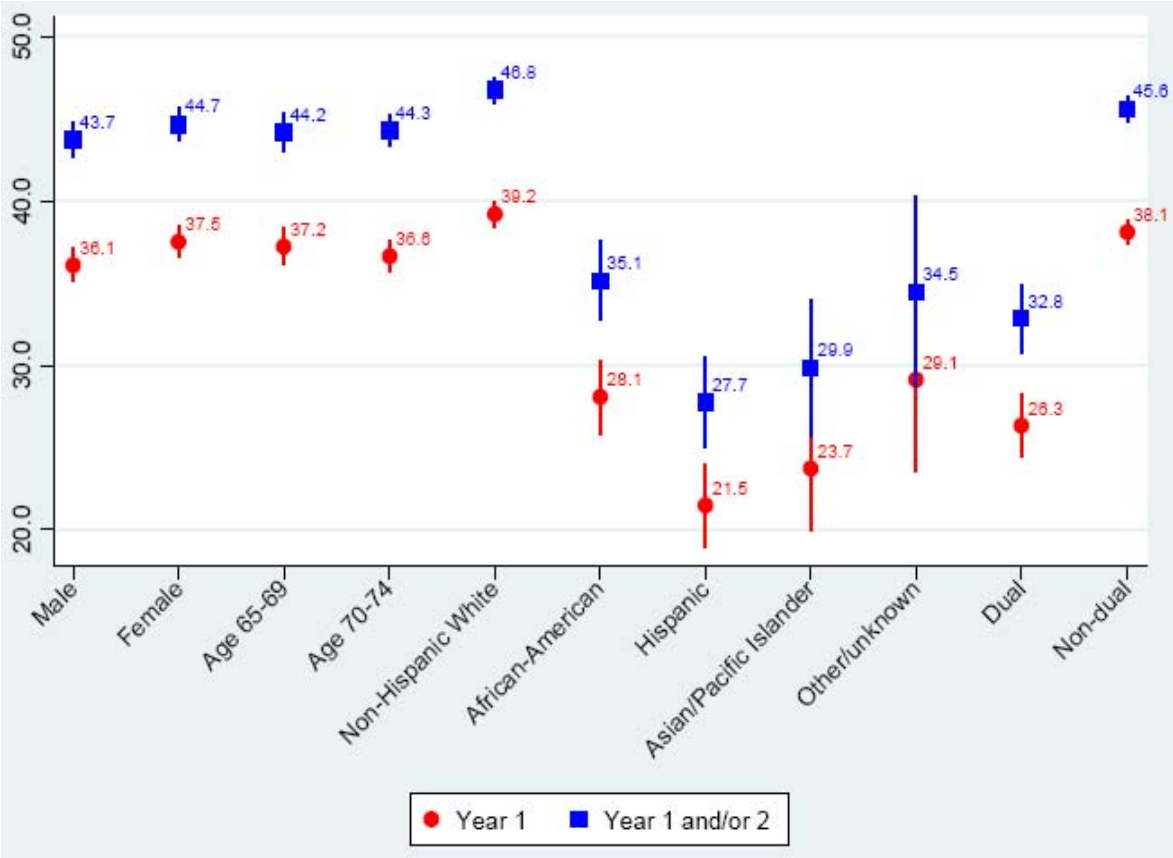
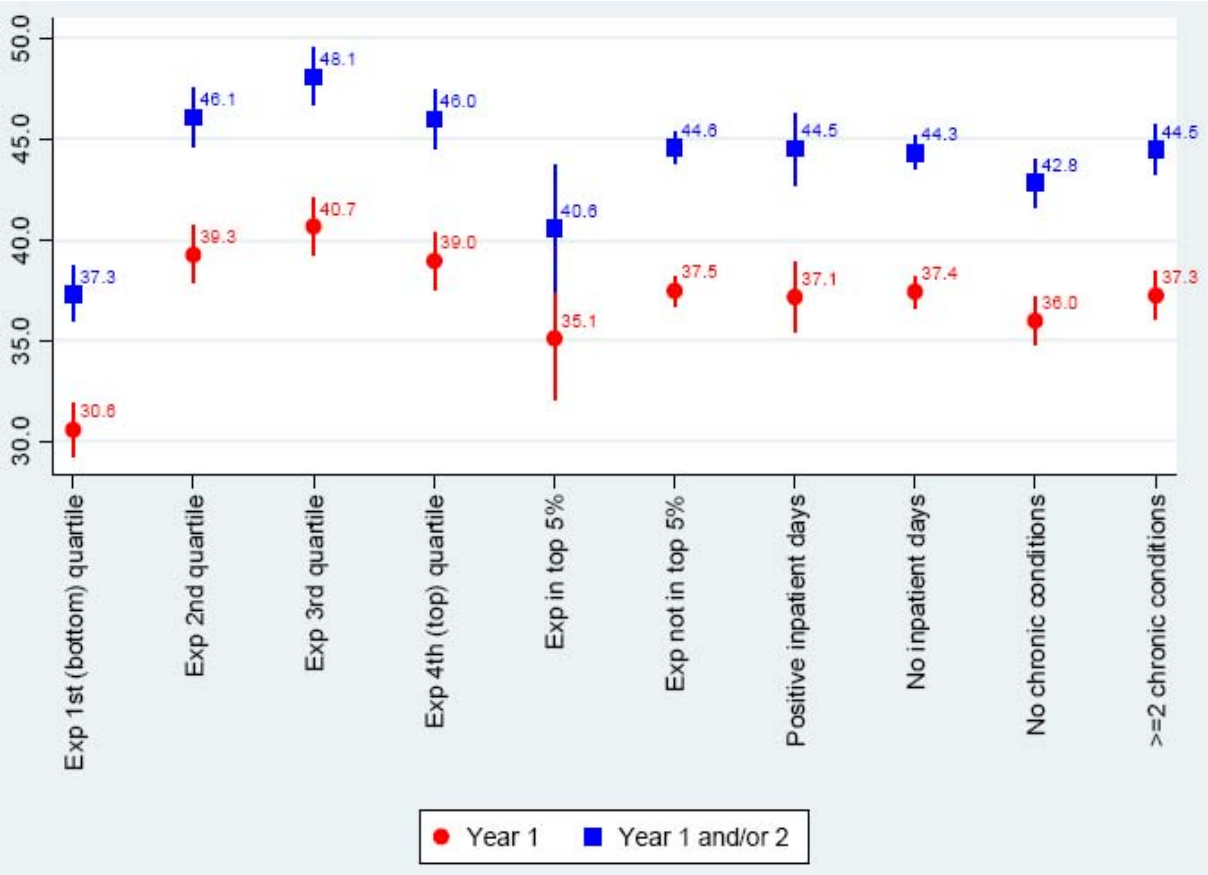


Exhibit 5.4: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Demographics – Vendor B



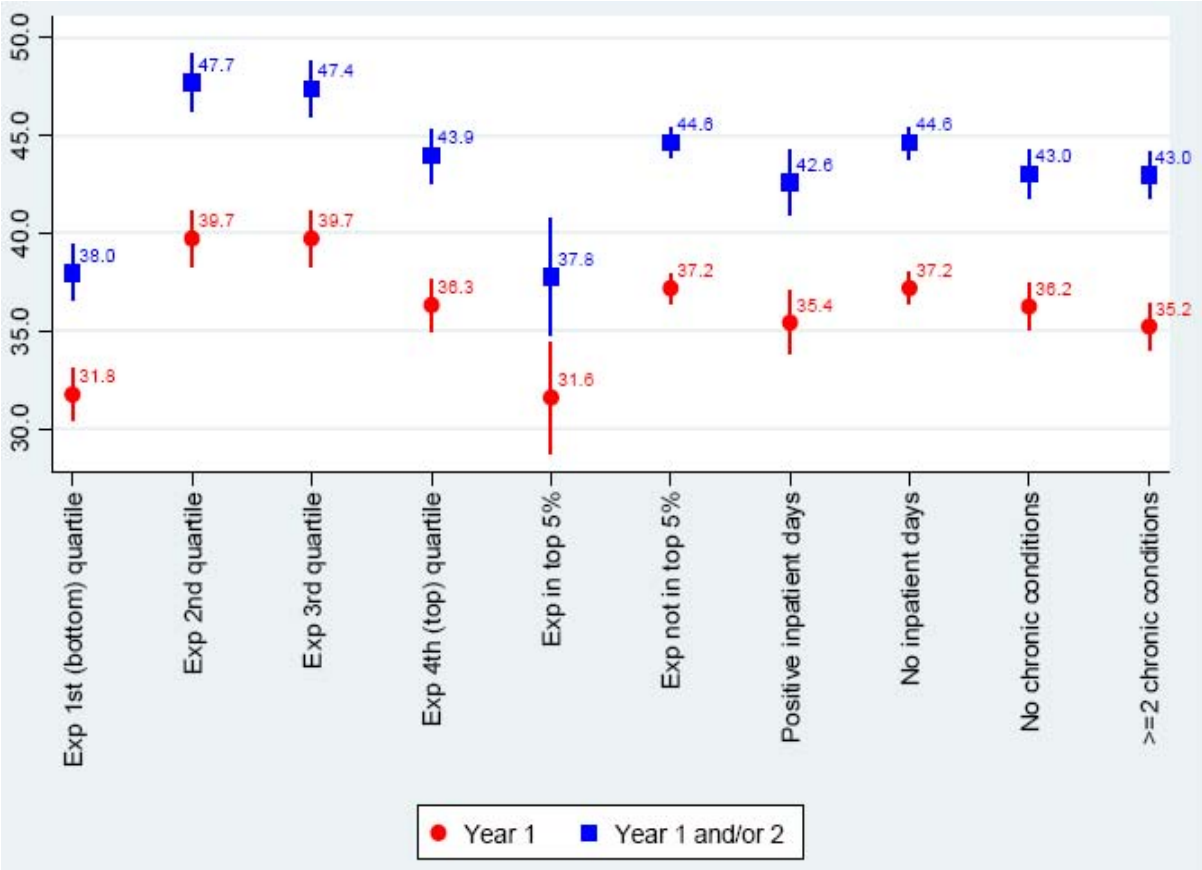
Exhibits 5.5 and 5.6 provide information on participation by expenditure, health services use, and number of chronic conditions. Both vendors' participation profiles for expenditure quartiles follow an inverted-U pattern, with lowest participation among beneficiaries in the lowest quartile, followed by those in the highest quartile, and then those in the middle 50 percent. The inverted-U pattern indicates that program uptake is highest among the population on which the intervention is likely to have the greatest effect.[29] In addition, for Vendor B, enrollees having expenditures in the top 5 percent had lower participation rates than the other 95 percent of enrollees. There were no statistically significant differences in participation rates for positive versus zero inpatient days or for two or more chronic conditions versus zero chronic conditions.

Exhibit 5.5: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Health Services Expenditure and Use – Vendor A



[29] The SRRD is likely to have the most impact among beneficiaries of average health status (i.e., those in the middle two total expenditure quartiles) rather than among the very sick (for whom no home-based intervention would be likely to prevent hospitalization) or the very healthy (for whom the hospitalization rate is very low). The study described in Wennberg, D.E., Marr, A., Lang, L., O'Malley, S., and Bennet, G. (2010), A randomized trial of a telephone care-management strategy, *New England Journal of Medicine*, 363(13), 1245-55, also employs this hypothesis.

Exhibit 5.6: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Health Services Expenditure and Use – Vendor B



Exhibits 5.7 and 5.8 examine participation rates for 11 specific chronic conditions. The exhibits indicate that, for both vendors, participation rates were highest among beneficiaries with cancer, osteoporosis, and arthritis, and lowest among those with Alzheimer's' disease. Low participation among Alzheimer's patients was expected since the demonstration requires that beneficiaries have sufficient mental capacity to engage in the SRRD.

Exhibit 5.7: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Chronic Condition – Vendor A

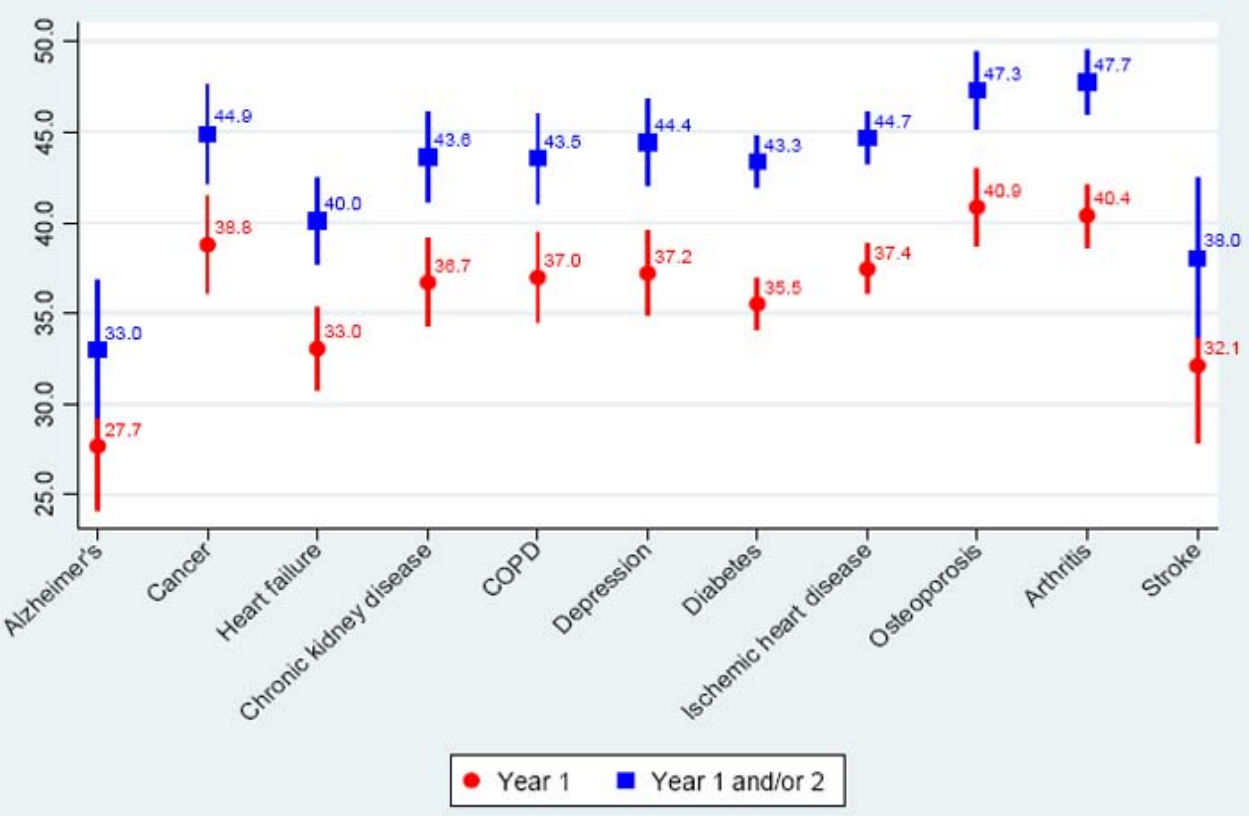
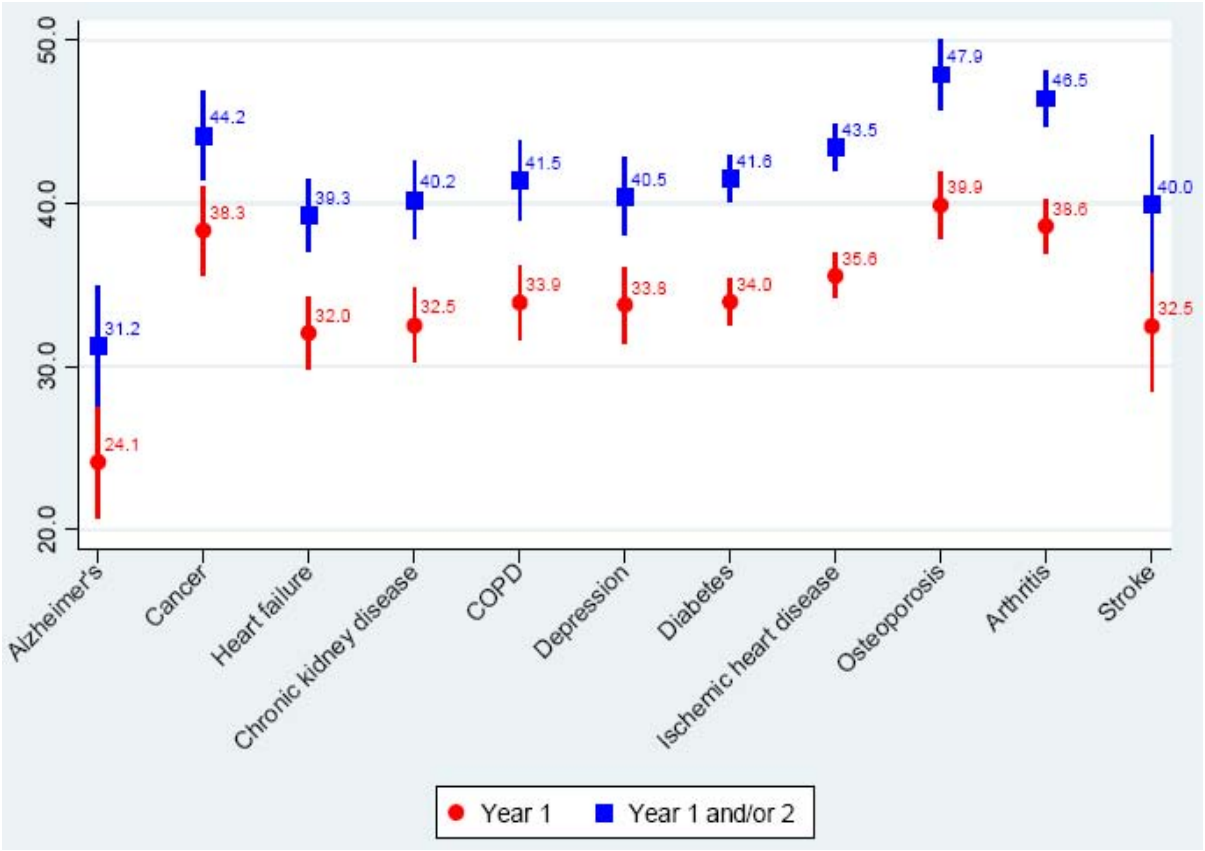


Exhibit 5.8: Year 1 and Year 1 and/or Year 2 Participation Rates by Baseline Chronic Condition – Vendor B



C. Characteristics of Participating Beneficiaries and Opt-Outs

Exhibits 5.9 through 5.10 provide information on the characteristics of beneficiaries participating in the demonstration and on tests of statistical differences between participants and the relevant eligible group. For example, column 3 compares baseline participants to the intervention group from which they were drawn, while column 5 compares re-enrollees to participants at baseline.

Exhibits 5.9 and 5.10 pool the local and national samples and present summary measures of participant characteristics. Column 3 of the exhibits indicates that both vendors recruited baseline participants who, compared to the pool of eligible beneficiaries, were more likely to be non-Hispanic Whites and to have had preventive screenings. Baseline year participants also were less likely to be dual eligible and to have some chronic conditions, including Alzheimer's disease, heart failure, and diabetes. Both vendors' baseline participants were more likely to have osteoporosis; Vendor B's participants had lower total and inpatient Medicare payments and were less likely to have chronic kidney disease than the intervention group.

Column 5 of Exhibit 5.9 shows that Vendor A's participants who re-enrolled in the demonstration were generally not statistically significantly different from the baseline enrollees. However, the sign of the differences in column 5 may indicate continuing selection issues. Column 5 of Exhibit 5.10 shows that Vendor B's re-enrollees were more likely to be non-Hispanic Whites than were the baseline enrollees. Even though other differences are not statistically significant, the consistency of the signs with those in column 3 may suggest ongoing selection issues among re-enrollees.

Exhibit 5.9: Participant Characteristics at Baseline – Vendor A, Local and National

Baseline beneficiary characteristics[a]	IG	Year 1 pooled arms[b]	Diff.[c]	Re-enrollees	Diff.[c]
	(1)	(2)	(3)=(2)-(1)	(4)	(5)=(4)-(2)
	N=20,000	N=7,473		N=4,433	
Demographic characteristics					
Female	55.6%	55.8%	0.3	56.0%	0.2
Age (as of 5/1/09)	70.2	70.2	0.0	70.2	0.0
Race (RTI race code)					
Non-Hispanic White	83.8%	87.7%	3.9**	88.7%	1.0
African American	6.5%	5.2%	-1.3**	4.6%	-0.5
Hispanic	5.0%	3.5%	-1.5**	3.2%	-0.3
Asian/Pacific Islander	3.3%	2.3%	-1.0**	2.2%	-0.2
Medicaid/Medicare dual eligibility	11.2%	8.9%	-2.3**	8.5%	-0.4
Medicare use					
Total Medicare payment	$7,619	$7,573	-46	$7,107	-467
Inpatient payment	$2,893	$2,812	-81	$2,532	-280
Physician payment	$2,447	$2,551	105	$2,453	-99
Outpatient payment	$1,387	$1,394	7	$1,375	-19
Chronic conditions[d]					
Alzheimer's/related disorders/senile dementia	3.2%	2.4%	-0.8**	1.9%	-0.5
Cancer	6.8%	7.0%	0.3	7.4%	0.4
Heart failure	8.9%	7.9%	-1.0**	7.5%	-0.4
Chronic kidney disease	8.2%	8.0%	-0.1	7.2%	-0.8
COPD	8.1%	8.0%	-0.1	7.0%	-1.0
Depression	9.0%	9.0%	0.0	7.9%	-1.1*
Diabetes	25.3%	24.0%	-1.3*	23.2%	-0.8
Ischemic heart disease	25.9%	26.0%	0.1	25.5%	-0.5
Osteoporosis	11.6%	12.7%	1.1*	13.6%	0.8
Rheumatoid/osteoarthritis	17.7%	19.1%	1.4**	19.1%	0.0
Stroke/TIA	2.4%	2.1%	-0.3	1.7%	-0.4
Preventive screening use					
Colorectal CA screening	22.9%	25.6%	2.7**	26.4%	0.8
Breast cancer screening	58.4%	67.2%	8.8**	70.2%	3.0*
Cardiovascular screening	68.1%	73.3%	5.2**	75.1%	1.8

[a] Characteristics are reported for the baseline period from January 1, 2008 to April 30, 2009, unless indicated otherwise.

[b] Year 1 participants filled out an HRA at the beginning of the first year of the demonstration.

[c] * indicates significance at the 5 percent level; ** indicates significance at the 1 percent level.

[d] Chronic conditions were measured at mid-year 2009.

Exhibit 5.10: Participant Characteristics at Baseline – Vendor B, Local and National

Baseline beneficiary characteristics[a]	IG	Year 1 pooled arms[b]	Diff.[c]	Re-enrollees pooled arms	Diff.[c]
	(1)	(2)	(3)=(2)-(1)	(4)	(5)=(4)-(2)
	N=20,000	N=7,375		N=4,154	
Demographic characteristics					
Female	54.6%	55.5%	1.0	56.6%	1.1
Age (as of 5/1/09)	70.3	70.2	0.0	70.3	0.1
Race (RTI race code)					
Non-Hispanic White	83.0%	88.3%	5.3**	90.1%	1.8**
African American	7.9%	6.0%	-1.9**	5.2%	-0.8
Hispanic	5.3%	3.1%	-2.2**	2.3%	-0.8*
Asian/Pacific Islander	2.4%	1.6%	-0.9**	1.4%	-0.2
Medicaid/Medicare dual eligibility	10.5%	7.5%	-3.0**	6.8%	-0.7
Medicare use					
Total Medicare payment	$8,225	$7,486	-739.5**	$7,184	-302
Inpatient payment	$3,248	$2,777	-470.3**	$2,669	-109
Physician payment	$2,665	$2,678	13	$2,617	-61
Outpatient payment	$1,320	$1,352	32	$1,287	-66
Chronic conditions[d]					
Alzheimer's/related disorders/senile dementia	3.2%	2.1%	-1.1**	1.7%	-0.4
Cancer	6.9%	7.2%	0.3	6.9%	-0.3
Heart failure	9.7%	8.4%	-1.3**	7.7%	-0.7
Chronic kidney disease	8.8%	7.8%	-1.0**	7.1%	-0.7
COPD	8.7%	8.0%	-0.7	7.2%	-0.8
Depression	8.8%	8.0%	-0.7	8.0%	0.0
Diabetes	25.8%	23.7%	-2.0**	22.8%	-0.9
Ischemic heart disease	27.1%	26.2%	-1.0	26.4%	0.2
Osteoporosis	11.8%	12.8%	1.0*	13.2%	0.4
Rheumatoid/osteoarthritis	18.4%	19.2%	0.9	19.1%	-0.1
Stroke/TIA	2.7%	2.4%	-0.3	2.1%	-0.3
Preventive screening use					
Colorectal CA screening	23.2%	26.0%	2.8**	26.7%	0.7
Breast cancer screening	57.4%	67.8%	10.3**	70.6%	2.9*
Cardiovascular screening	67.8%	72.5%	4.7**	74.2%	1.8

[a] Characteristics are reported for the baseline period from January 1, 2008 to April 30, 2009, unless indicated otherwise.

[b] Year 1 participants filled out an HRA at the beginning of the first year of the demonstration.

[c] * indicates significance at the 5 percent level; ** indicates significance at the 1 percent level.

[d] Chronic conditions were measured at mid-year 2009.

Exhibit 5.11 displays, for Vendor A, baseline characteristics for individuals who participated in Year 1 (column 1), those who opted out after participating in Year 1 (column 2), and those who opted out before participating (column 4).[30] Columns 3 and 5 provide statistical comparisons between opt-outs and participants.

Compared to participants, opt-outs were more likely to have been randomized into treatment Arms 1 or 2 and much less likely to have been randomized into Arm 3. This is consistent with the participation rate results presented earlier, that individuals in Arm 3 were more likely to re-enroll than were individuals in Arms 1 or 2. Also, opt-outs were more likely to be categorized as medium risk and less likely to be categorized as low risk, compared to the broader group of participants.

It can also be seen that opt-outs were slightly older than participants and were more likely to be non-Hispanic Whites. Moreover, relative to participants, opt-outs without an HRA were less likely to have a number of chronic conditions (depression, heart disease, arthritis, and stroke) and preventive screenings (colorectal, breast cancer, and cardiovascular). Although a similar pattern exists for opt-outs with an HRA, the differences are not statistically significant. The likely reason is the small sample size (n=58) for opt-outs with an HRA.

In addition to differences between opt-outs and the broader participant group, there are also differences between opt-outs who completed a Year 1 HRA and those who did not. Those without an HRA opted out based on very minimal interaction with the SRRD, while those with an HRA opted out sometime during the first year of the demonstration, but after filling out the HRA. Compared to participants, opt-outs with an HRA tended to have higher Medicare expenditures and health services utilization, while those without an HRA tended to have lower expenditures and utilization.

[30] This information was not available for Vendor B.

Exhibit 5.11: Participant Characteristics at Baseline – Opt-Outs, Vendor A, Local and National

Baseline beneficiary characteristics[a]	Year 1 pooled arms[b]	Opt-outs with a Year 1 HRA		Opt-outs without a Year 1 HRA	
	(1)	(2)	Difference[c]	(4)	Difference[c]
	N=7,473	N=58		N=3,889	
General information					
Arm 1	33.3%	46.6%	13.2**	-	-
Arm 2	33.4%	48.3%	14.9**	-	-
Arm 3	33.3%	5.2%	-28.1**	-	-
High risk	6.8%	10.3%	3.5	-	-
Medium risk	28.1%	43.1%	15.0**	-	-
Low risk	65.0%	46.6%	-18.5**	-	-
Demographic characteristics					
Female	55.8%	53.4%	-2.4	56.4%	0.6
Age (as of 5/1/09)	70.2	70.7	0.5*	70.4	0.2**
Race (RTI race code)					
Non-Hispanic White	87.7%	94.8%	7.1*	89.8%	2.1**
African American	5.2%	1.7%	-3.4	3.9%	-1.3**
Hispanic	3.5%	0.0%	-3.5	2.6%	-0.9**
Asian/Pacific Islander	2.3%	3.4%	1.1	2.4%	0.0
Medicaid/Medicare dual eligibility	8.9%	8.6%	-0.3	7.6%	-1.3**
Medicare expenditures					
Total Medicare payment	$7,573	$11,777	4,204*	$6,770	-804**
Total Medicare payment in 1st (bottom) quartile	20.6%	10.3%	-10.2*	26.4%	5.9**
Total Medicare payment in 2nd quartile	26.8%	29.3%	2.5	26.7%	-0.1
Total Medicare payment in 3rd quartile	27.4%	24.1%	-3.2	24.8%	-2.6**
Total Medicare payment in 4th (top) quartile	25.2%	36.2%	11.0*	22.1%	-3.2**
Inpatient payment	$2,812	$5,315	2,503*	$2,506	-306
Physician payment	$2,551	$3,054	503	$2,282	-270**
Outpatient payment	$1,395	$2,239	845	$1,263	-131
Medicare use					
Inpatient length of stay	1.3	3.2	1.9**	1.2	-0.1
Carrier days	21.0	27.3	6.3**	18.8	-2.2**
Outpatient days	5.2	6.3	1.1	4.6	-0.6**
Health status[d] and chronic conditions[e]					
Self-rated health status (1=poor, 5=excellent)	3.3	3.1	-0.1		
Alzheimer's/related disorders/senile dementia	2.4%	5.2%	2.8	3.9%	1.6**
Cancer	7.0%	10.3%	3.3	6.8%	-0.2
Heart failure	7.9%	6.9%	-1.0	8.1%	0.2
Chronic kidney disease	8.0%	8.6%	0.6	7.4%	-0.6

Baseline beneficiary	Year 1 pooled arms[b]	Opt-outs with a Year 1 HRA		Opt-outs without a Year 1 HRA	
	(1)	(2)	Difference[c]	(4)	Difference[c]
	N=7,473	N=58		N=3,889	
COPD	8.0%	8.6%	0.6	7.4%	-0.6
Depression	9.0%	8.6%	-0.3	7.8%	-1.1**
Diabetes	24.0%	31.0%	7.0	23.0%	-1.0
Ischemic heart disease	26.0%	25.9%	-0.1	24.2%	-1.8**
Osteoporosis	12.7%	12.1%	-0.7	12.0%	-0.7
Rheumatoid/osteoarthritis	19.1%	24.1%	5.1	16.8%	-2.3**
Stroke/TIA	2.1%	5.2%	3.1	2.6%	0.5*
Preventive screening use					
Colorectal CA screening	25.6%	16.1%	-9.5	23.4%	-2.2**
Breast cancer screening	67.2%	64.5%	-2.7	58.3%	-8.9**
Cardiovascular screening	73.3%	70.7%	-2.5	68.0%	-5.3**

[a] Characteristics are reported for the baseline period from January 1, 2008 to April 30, 2009, unless indicated otherwise.
[b] Year 1 participants filled out an HRA at the beginning of the first year of the demonstration.
[c] *p <.10, **p <.05.
[d] Health status is based on the HRA item "In general, how would you describe your health?" Possible responses include: Poor (1), Fair (2), Good (3), Very Good (4), Excellent (5).
[e] Chronic conditions were measured at mid-year 2009.

VI. IMPACT ANALYSIS RESULTS

In this section, we describe our findings from the impact analysis. We estimated SRRD impacts in the first 12 months of the demonstration (May 1, 2009–April 30, 2010) on a number of key Medicare expenditure outcome measures, including behavioral risk measures constructed by the vendors. Impacts were estimated by comparing the outcome measures for Arm 1 – Standard Treatment (HRA + standard tailored follow-up) and Arm 2 – Enhanced Treatment (HRA + enhanced tailored follow-up) to the outcome measures for Arm 3 – HRA Only (HRA + generic health advice). Prior to the estimation of impacts, we assessed the randomization into these three arms because successful randomization is necessary for unbiased impact estimates.

Summary of Results:

- The results of our tests for statistically significant differences in baseline measures between study groups were consistent with a successful randomization.
- In Vendor A's national program, Arm 2 participants had $958 less total Medicare expenditures, were 14.2 percent less likely to be hospitalized, and had 0.3 fewer hospital days than Arm 3 participants.
- The impacts for Vendor A were driven by participants who Vendor A triaged into high-risk and moderate-risk intervention groups. For high-risk Arm 2 participants in Vendor A's national sample, total Medicare payments decreased by $6,634 compared with payments for Arm 3 participants. For the high-risk group, the number of emergency department visits was also lower for national Arm 2 participants versus Arm 3 participants.
- We did not detect any statistically significant effect of Vendor B's national Arm 2 program compared to its national Arm 3 program. However, for Vendor B's national Arm 1 intervention, we detected increases of $822 and $273 in total Medicare expenditures and carrier payments, respectively, compared to its Arm 3 sample.
- For both vendors, there were no statistically significant impacts on the use of preventive screenings in their national Arm 2 programs as compared to their Arm 3 programs. However, Vendor A's local participants in Arms 1 and 2 pooled were 6.1 percentage points more likely to use cardiovascular screening than were their Arm 3 counterparts.
- We did not detect any impact on number of hospitalizations for ambulatory care sensitive conditions (ACSC). However, there were very few participants who had hospitalizations for any ambulatory care conditions, making it difficult to detect impacts.
- Of the 16 HRA-based measures (eight per vendor) we examined, there were impacts on only two (one per vendor) that were statistically significant at the 5 percent level. As compared to their Arm 3 counterparts, Vendor A's national Arm 2 participants were 5.3 percentage points less likely to have high/medium polypharmacy risk. Vendor B's national Arm 2 participants were 3.9 percentage points less likely to have high/medium exercise risk than their Arm 3 counterparts.

- As a robustness test, we examined two alternative estimators: difference in means (DIM) and difference-in-differences (DID).
  - The regression methodology produced more precise estimates than did the DIM and DID methodologies.
  - The regression results were more conservative than the DIM and DID results because the regressions produced estimates that were smaller in magnitude than those of the other two methods.

A. Assessment of Randomizing Participants into Arms

Using t-tests of means, we tested for any differences in baseline characteristics between the comparison groups.[31] In a successful randomization, baseline characteristics should not differ between comparison groups. Exhibits 6.1 and 6.2 show the mean differences in claims-based and HRA variables between Arm 2 and Arm 3 participants for Vendor A and Vendor B national samples, respectively.[32] The validity of our main impact estimates (where we compare outcomes across arms) depends on the success of the randomization. Our pre-specified statistical significance level is 5 percent. Note that, just by chance, we would expect 5 out 100 tests to be significant at the 5 percent level. Given that we examined about 20 characteristics, one statistically significant difference is attributable to chance.

We concluded that there were no statistically significant differences in baseline measures between comparison groups. We interpreted this as being consistent with a successful randomization.

[31] The baseline period for claims-based measures is January 1, 2008 to April 30, 2009 (the 16 months since the start of the demonstration). The baseline period for HRA variables is the Year 1 HRA responses, which were collected at the beginning of the demonstration.

[32] We also assessed the randomization into IG and ACG. Randomization produced balanced IG and ACG groups for both vendors. For the results, see Exhibits B.1 and B.2 in Appendix B.

Exhibit 6.1: Vendor A Beneficiary Baseline Characteristics: National Sample Arm 2 versus Arm 3

Baseline Characteristics	National Arm 2	National Arm 3	Difference: [a] Arm 2 – Arm 3	P-Value
Sample Size	2,019	2,016		
Medicare Expenditures				
Total Medicare Payments	$7,966	$7,433	$533	0.325
Inpatient Payment	$3,040	$2,663	$376	0.279
Outpatient Payment	$1,374	$1,362	$12	0.924
Carrier Payment	$2,634	$2,600	$34	0.821
ED Payment	$114	$106	$8	0.465
Medicare Use				
Any Inpatient	16.3%	16.4%	-0.07%	0.949
Inpatient Length of Stay	1.48	1.19	0.29*	0.095
Any Outpatient	75.9%	76.2%	-0.26%	0.846
Outpatient Days	4.8	4.8	0.0	0.953
Carrier Days	21.4	20.9	0.5	0.431
ED Visits	0.47	0.43	0.04	0.342
Demographics				
Female	54.7%	56.2%	-1.42%	0.364
Age	70.9	70.8	0.04	0.512
Race: Non-Hispanic White	87.0%	86.7%	0.27%	0.802
Race: African American	6.1%	6.0%	0.09%	0.904
Race: Hispanic	3.4%	4.4%	-0.95%	0.120
Race: Asian/Pacific Islander	2.1%	1.7%	0.39%	0.364
Race: Other/Unknown	1.4%	1.2%	0.20%	0.580
Dual Eligibility Status	9.0%	9.0%	-0.06%	0.944
Health Status				
Multiple Chronic Conditions	34.0%	35.6%	-1.64%	0.275
HCC Risk Score (2009)[b]	0.85	0.85	0.00	0.937
HRA Variables[c]				
Self-Reported Health Status = Very Good/Excellent	39.7%	39.7%	0.01%	0.996
Self-Reported Health Status = Good	39.9%	39.7%	0.21%	0.892
Self-Reported Health Status = Poor/Fair	20.4%	20.6%	-0.22%	0.864
Self-Reported Health Status (1=very poor; 5=very good)	3.2	3.2	-0.01	0.737
% High Risk (Based on HRA responses)	7.6%	7.3%	0.34%	0.685
% Medium Risk (Based on HRA responses)	27.4%	28.6%	-1.18%	0.403
% Low Risk (Based on HRA responses)	65.0%	64.1%	0.85%	0.574

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[b] Due to missing values, the sample size for the HCC Risk Score is 2,015 for Arm 3.

[c] Due to missing values, the sample sizes for the HRA variables are 1,950 for Arm 2 and 1,963 for Arm 3.

Exhibit 6.2: Vendor B Beneficiary Baseline Characteristics: National Sample Arm 2 versus Arm 3

Baseline Characteristics	National Arm 2	National Arm 3	Difference: [a] Arm 2 – Arm 3	P-Value
Sample Size	1,964	1,965	-	-
Medicare Expenditures				
Total Medicare Payments	$7,310	$7,801	-490	0.339
Inpatient Payment	$2,685	$3,071	-386	0.238
Outpatient Payment	$1,351	$1,327	24	0.850
Carrier Payment	$2,645	$2,674	-29	0.847
ED Payment	$97	$92	6	0.516
Medicare Use				
Any Inpatient	17.0%	18.7%	-1.8%	0.147
Inpatient Length of Stay	1.2	1.3	-0.2	0.305
Any Outpatient	77.5%	76.8%	0.7%	0.627
Outpatient Days	4.9	4.9	0.0	0.888
Carrier Days	20.65	21.07	-0.41	0.480
ED Visits	0.4	0.4	0.0	0.463
Demographics				
Female	54.0%	55.3%	-1.2%	0.433
Age	70.8	70.9	0.0	0.756
Race: Non-Hispanic White	86.5%	87.9%	-1.4%	0.195
Race: African American	7.4%	6.2%	1.2%	0.126
Race: Hispanic	3.1%	3.3%	-0.2%	0.717
Race: Asian/Pacific Islander	1.7%	1.6%	0.1%	0.799
Race: Other/Unknown	1.3%	1.1%	0.2%	0.560
Dual Eligibility Status	8.2%	7.9%	0.4%	0.678
Health Status				
Multiple Chronic Conditions[b]	34.9%	35.0%	-0.1%	0.929
HCC Risk Score (2009)	0.85	0.84	0.00	0.858
HRA Variables[c]				
Self-Reported Health Status = Very Good/Excellent	50.9%	52.4%	-1.4%	0.369
Self-Reported Health Status = Good	30.3%	29.9%	0.4%	0.769
Self-Reported Health Status = Very Poor/Poor/Fair	18.7%	17.7%	1.0%	0.415
Self-Reported Health Status (1=Very Poor; 6=Excellent)	4.5	4.5	0.0	0.617

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[b] Due to missing values, the sample size for the HCC Risk Score is 1,963 for Arm 2 and 1,963 for Arm 3.

[c] Due to missing values, the sample sizes for HRA variables are 1,949 for Arm 2 and 1,947 for Arm 3.

We also tested the success of the randomization using regression models. We tested the impact of each of the baseline characteristics on the probability of being in each arm for the vendors. We did not find statistically significant impacts, confirming that, based on the characteristics we examined, beneficiaries were randomly assigned effectively.

B. Impact of SRRD

As described above, we analyzed the impact of the SRRD on the outcomes listed in Exhibit 4.1 using a regression-based approach to implement each of the comparisons listed in Exhibit 4.6. These outcomes include Medicare expenditures and use, use of preventive screenings covered by Medicare, hospitalizations for ambulatory care sensitive conditions, and HRA-based measures. This approach enabled us to cast a wide net in examining how SRRD impacted beneficiaries who received various levels of intervention. In this section, we present the most important findings.

Based on the design of the demonstration, any existing impacts would occur among the participants since they were eligible to receive intervention services. Impacts would be diluted among the IG since only a little over of a third of this group participated. The impact on the rest of IG (non-participants) can be assumed to be none because they did not receive any intervention services. We therefore examined demonstration impacts only among the participants.

Our analysis of the impact of the SRRD on Arm 1 and Arm 2 participants who completed at least the Year 1 HRA involved the comparisons shown earlier in Exhibit 4.6. For most of these comparisons, Medicare use and expenditure outcomes for treatment versus control group beneficiaries were not statistically significantly different. This pattern held for all comparisons with the exception of the comparison between Vendor A's national Arm 2 and Arm 3 (comparison 17 in Exhibit 4.6). For this reason, we focus in this section on the results for the impact of the national Arm 2 intervention compared with the national Arm 3 intervention for both vendors. In Exhibits 6.3 and 6.4, we present these impacts for Vendor A and vendor B, respectively. In Appendix C, we report the impact estimates for all comparisons.

Impacts on Medicare expenditures and use: As shown in Exhibit 6.3, Medicare expenditures for participants in Arm 2 of Vendor A's national sample were $958 less (β= -958, p= 0.047) than for participants in Arm 3 of Vendor A's national sample, or 14.2 percent less in Medicare expenditures based on the regression-adjusted mean Medicare expenditure for Arm 3. The probability of having any inpatient hospitalizations was 2.1 percentage points lower (β= -2.1%, p= 0.048) for Vendor A's Arm 2 national sample compared to the participants in Vendor A's Arm 3 national sample, or 14.4 percent less in national Arm 2 Medicare expenditures than in national Arm 3 expenditures based on the regression-adjusted mean for Arm 3. Payments for inpatient hospitalizations and inpatient days were lower for participants in Arm 2 compared to those in Arm 3 of Vendor A's national sample (β= -491, p= 0.144 for inpatient payments and β= -0.3, p=0.095 for inpatient days), though payments for inpatient hospitalizations were not statistically significant. Payments for any outpatient, outpatient services, and outpatient days

were lower for participants in Arm 2 compared to those in Arm 3 of Vendor A's national sample (β= -2.2, p= 0.083 for any outpatient; β= -280, p= 0.003 for outpatient payments; and β= -0.2, p=0.062 for outpatient days). The number of carrier days and carrier payments also were lower for participants in Vendor A's Arm 2 compared to those in Arm 3 (β= -200, p= 0.099 for carrier payment and β= -0.6, p= 0.106 for carrier days). We did not find any significant differences in emergency department payments or emergency department visits between Vendor A's Arm 2 and Arm 3 national sample.

As shown in Exhibit 6.4, we did not detect any statistically significant effect of Vendor B's national Arm 2 program compared to its national Arm 3 program. However, we detected an impact of an increase in total Medicare expenditures (β= +822, p= 0.092) and carrier payments (β= +273, p= 0.026) for participants in Vendor A national Arm 1 sample versus Arm 3 sample (see Exhibit C.23 in Appendix C).

Impacts on use of preventive **screenings:** For either vendor, there were no statistically significant impacts on the use of preventive screenings of their national Arm 2 programs as compared to their Arm 3 programs. The only intervention that caused an impact was Vendor A's local intervention when both Arms 1 and 2 were pooled. Vendor A's local participants in Arms 1 and 2 pooled were 6.1 percentage points more likely (β= 6.1%, p= 0.036) to use cardiovascular screening than were their Arm 3 counterparts (see Exhibit C.22 in Appendix C).

Impacts on hospitalizations for ACSCs**:** The outcome measure "any hospitalizations" for the 12 ACSCs and three composite ACSCs were examined using regression analyses. However, we found that the number of hospitalizations in our sample for ACSCs was too small to detect any meaningful impacts of the demonstration. For example, less than 15 percent of Vendor A's national participants had any hospitalizations. Of these patients, less than 20 percent had a hospitalization for any ACSC.

Exhibit 6.3: Impact of SRRD on Vendor A's Participants in National Arm 2 versus Arm 3[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample size[c]	-	-	2,019	2,016	-
Medicare Payment					
Total Medicare Payments	-958**	0.047	5,807	6,765	-14.2%
Inpatient Payment	-491	0.144	2,090	2,581	-19.0%
Outpatient Payment	-280***	0.003	915	1,196	-23.4%
Carrier Payment	-200*	0.099	2,037	2,237	-8.9%
ED Payment	2	0.831	87	85	1.9%
Medicare Use					
Any Inpatient	-2.1%**	0.048	12.7%	14.8%	-14.4%
Inpatient Days	-0.3*	0.095	0.9	1.2	-24.6%
Any Outpatient	-2.2%*	0.083	70.5%	72.6%	-3.0%
Outpatient Days	-0.2*	0.062	3.6	3.8	-6.5%
Carrier Days	-0.6	0.106	16.3	16.9	-3.5%
ED Visits	0.0	0.944	0.3	0.3	0.6%
Preventive Screening[d]					
Colorectal Cancer	1.7%	0.211	24.3%	22.7%	7.3%
Breast Cancer	-0.5%	0.800	60.3%	60.7%	-0.8%
Cardiovascular	0.4%	0.790	66.0%	65.5%	0.7%

[a] The regression included only those who completed the Year 1 HRA.

[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.

[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

Exhibit 6.4: Impact of SRRD on Vendor B's Participants in National Arm 2 versus Arm 3[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample Size[c]	-	-	1,964	1,965	-
Medicare Payment					
Total Medicare Payments	519	0.288	7,034	6,516	8.0%
Inpatient Payment	405	0.235	2,835	2,430	16.7%
Outpatient Payment	34	0.724	1,177	1,143	3.0%
Carrier Payment	54	0.658	2,248	2,194	2.5%
ED Payment	-0.5	0.944	84.8	85.3	-0.6%
Medicare Use					
Any Inpatient	-0.2%	0.885	14.3%	14.5%	-1.1%
Inpatient Length of Stay	0.2	0.286	1.3	1.1	17.1%
Any Outpatient	1.4%	0.260	73.5%	72.0%	2.0%
Outpatient Days	-0.1	0.450	3.7	3.8	-2.7%
Carrier Days	0.4	0.325	17.1	16.7	2.2%
ED Visits	0.0	0.812	0.3	0.3	1.9%
Preventive Screening[d]					
Colorectal Cancer	0.6%	0.676	23.4%	22.9%	2.5%
Breast Cancer	0.0%	0.987	60.7%	60.6%	0.0%
Cardiovascular	1.0%	0.552	66.4%	65.4%	1.5%

[a] The regressions included only those who completed the Year 1 HRA.
[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.
[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

Impacts on HRA-based measures: In addition to using the claims data to assess the impact of the SRRD on participants, we also examined impacts on HRA-based risk measures (see Exhibit 6.5). The probability of reporting a poor or fair health status was 2.2 percentage points lower (β= -2.2%, p = 0.061) for Arm 2 participants than for Arm 3 participants in the national sample. Arm 2 participants in Vendor A's national sample had lower polypharmacy risk compared to Arm 3 participants (β= -5.3%, p= 0.003). For Vendor B, we found a lower probability of being high or medium risk for exercise (β= -3.9%, p= 0.030) for Arm 2 participants versus Arm 3 participants in the national sample. However, there is no indication of an impact on any of the other HRA variables. Overall, the analyses demonstrated no significant patterns of change in the HRA variables that could be attributed to the SRRD intervention for the first year of the intervention.

We checked the robustness of these results by recoding variables such as self-rated health status into finer categories: instead of poor/fair, good, and very good/excellent, we used five

categories—poor, fair, good, very good, and excellent. Defining health status more precisely might have enabled us to detect incremental changes in health status. However, recoding did not change our results. Since some participants dropped out in Year 2, we present sample sizes by HRA measures in Appendix B, Exhibit B.3.

Exhibit 6.5: Impact of SRRD on HRA Responses for Vendor A and Vendor B National Participants

HRA Variables[a]	Sample Size[b] (1)	Coefficient[c] (2)	P-value (3)	Regression- Adjusted Means Arm 2 (4)	Arm 3 (5)	Percent Impact (6)=100x(2)/(5)
Vendor A National Sample						
Health status = Poor/Fair	3,407	-2.2%*	0.061	14.7%	16.9%	-12.84
Health status = Good	3,407	1.0%	0.575	42.2%	41.2%	2.46
Health status = Very good/Excellent	3,407	1.0%	0.509	43.0%	42.0%	2.48
Risk stratum = High/medium	3,574	-2.8%	0.103	34.3%	37.1%	-7.49
Alcohol risk = High/medium	3,546	-0.2%	0.814	7.5%	7.7%	-2.18
Diet risk = High/medium	3,457	-0.9%	0.272	94.5%	95.4%	-0.93
Preventive services risk = High/medium	3,449	0.0%	0.987	18.9%	18.9%	-0.09
Physical activity = High/medium risk	3,092	1.4%	0.423	35.8%	34.4%	4.12
Tobacco risk = High/medium	3,431	-0.8%	0.373	12.7%	13.5%	-5.81
Polypharmacy risk = High/medium	3,479	-5.3%***	0.003	41.1%	46.3%	-11.37
Vendor B National Sample						
Self-Reported Health Status = Very Poor/Poor/Fair	3,196	-0.5%	0.704	16.6%	17.1%	-2.98
Self-Reported Health Status = Good	3,196	-0.5%	0.799	29.6%	30.1%	-1.55
Self-Reported Health Status = Very Good/Excellent	3,196	0.8%	0.619	53.6%	52.8%	1.60
Current lifestyle = High/medium	3,267	-2.8%	0.118	65.3%	68.1%	-4.11
Alcohol risk = High/medium	3,211	-0.3%	0.704	10.8%	11.1%	-3.00
Eating risk = High/medium	3,146	-0.7%	0.710	64.2%	64.9%	-1.08
Exams risk = High/medium	2,953	-2.6%	0.141	61.5%	64.1%	-4.08
Exercise risk = High/medium	3,225	-3.9%**	0.030	38.8%	42.7%	-9.16
Smoking risk = High/medium	3,207	0.0%	0.971	8.5%	8.4%	0.28
Self-care risk = High/medium	3,267	-1.1%	0.301	6.2%	7.3%	-14.75

[a] Dependent variable is binary: for risk variables, 1 = high/medium risk; 0 = low risk.

[b] Participants who completed the Year 1 and Year 2 HRAs.

[c] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

C. Robustness Checks

Strong positive findings for Vendor A's national Arm 2 program were unexpected given the "soft touch" nature of the intervention, which did not include any in-person contact. Furthermore, the follow-up period of 12 months is relatively short for an intervention that aims to change participant behavior to ultimately influence health services use and expenditures.[33] In general, it is hypothesized that behavioral interventions such as SRRD increase health services use in the first year. Due to the unexpected nature of the results for Vendor A's national Arm 2 group, additional analyses were conducted to explore the robustness of the results.

1. Outliers

Many of the key Medicare outcomes, such as total Medicare expenditures, inpatient expenditures, and number of hospital days, have very large values for a relatively small group of beneficiaries. We hypothesized that a few outliers were driving the results even though there is no systematic large SRRD impact. However, we ruled out this possibility because the impact estimate on one of the key outcomes— hospitalization rate—is large and statistically significant and is consistent with the estimated impacts on other outcomes. Hospitalization rate is based on a binary (0/1) variable (whether the beneficiary had any inpatient stay in the following 12 months) and thus cannot have any outliers. A large portion of Medicare total expenditure savings comes from the savings in inpatient expenditures, which is explained by lower hospitalization rates and fewer hospital days.

2. Examination of Alternative Impact Estimators

All of the foregoing results were produced using the regression approach described in the impact analysis methodology section. In addition to the regression approach, two alternative estimators are available: the difference in means (DIM) and difference-in-differences (DID) methodologies.

Since SRRD participants were randomized into Arms 1, 2, and 3, and since our randomization assessment revealed that there were no statistically significant differences in beneficiary characteristics across the arms, regression adjustment for beneficiary characteristics is not essential to obtain unbiased impact estimates. Rather, the DIM in outcome measures between Arms 1 or 2 on the one hand, and Arm 3 on the other, is also an unbiased estimator of the demonstration's impact. The DIM for outcome *Y* is given in the following equation (where *t=2* refers to the intervention [post] period and $\bar{Y}$ denotes the mean of *Y)*:

$$\text{DIM impact of SRRD on } Y = \bar{Y}_{t=2,Arm2} - \bar{Y}_{t=2,Arm3}$$

[33] On average, the duration between the receipt of the HRA and the end of the first year of the demonstration was 10 months.

This is the difference in the mean of *Y* during the intervention period for Arm 2 versus Arm 3.

Another option is DID. This methodology builds on the DIM estimate by computing the following equation (*t=1* refers to the baseline [pre] period):

$$\text{DID impact of SRRD on } Y = \left(\bar{Y}_{t=2,Arm2} - \bar{Y}_{t=2,Arm3}\right) - \left(\bar{Y}_{t=1,Arm2} - \bar{Y}_{t=1,Arm3}\right)$$

The expression in the first set of parentheses is the DIM impact estimate. From this, the DID estimate subtracts the second expression, which is the Arm 2 minus Arm 3 difference in *Y* during the baseline period. Subtracting the Arm 2–Arm 3 difference during the baseline period from that during the intervention period removes any time-invariant differences in outcomes between the treatment (Arm 2) and control (Arm 3) groups that would exist regardless of the time period (pre vs. post). This is important in cases where there could be differences in the treatment and control groups that do not change over time (e.g., genetic predisposition to illness).

As mentioned previously, the regression approach controls for beneficiary characteristics including age, gender, race, dual eligibility, baseline value of the dependent variable, and baseline total Medicare payments. The fact that these differences across arms were not statistically significant indicates that the variations are due to chance and are not systematic; this is expected given the randomized control trial design. However, as Peduzzi et al. (2002, p. 26) explain,[34] statistically indistinguishable imbalances across treatment groups may "exert a strong influence on the observed result of the trial." Indeed, although not statistically significant, there do appear to be some differences in beneficiary characteristics across treatment arms. For example, Vendor A's Arm 3 national sample participants had, on average, baseline total Medicare expenditures that were $533 greater than those of Arm 2 national sample participants. One recommended approach is to present both adjusted (e.g., regression, DID) and unadjusted (simple difference) results, as we do below for a subset of the outcomes. In addition to controlling for differences across treatment and control groups, statistical control for covariates reduces the residual variance and thus reduces the standard errors of the estimated impact parameters. This increased precision enables detection of statistically significant results that may have been missed.

In Exhibits 6.6 through 6.10, we present pre- and post-period means for Arms 1 and 2, as well as impact estimates based on DIM, DID, and regression methodologies. As suggested by the preceding discussion, regression provides more precise estimates than do DIM and DID. The standard errors produced by the regression methodology are always smaller than those generated by DID; this is true for all 11 outcomes examined below and for both vendors. Similarly, regression standard errors are smaller than DIM standard errors for all but 2 of the 22 regressions presented below (the exceptions are Vendor B total Medicare payments and Vendor B inpatient payments).

[34] Peduzzi, P., Henderson, W., Hartigan, P. and Lavori, P. (2002). Analysis of randomized controlled trials. *Epidemiologic Reviews* 24(1), 26-38.

There is also evidence that differences in the characteristics of participants in Arms 2 and 3 influence the impact estimates. The impacts from the regression methodology are always smaller in magnitude than are those from DID; this is true for all 11 outcomes and for both vendors. Similarly, the regression estimates are smaller in magnitude than are the DIM estimates for 20 of the 22 regression presented below (the exceptions are Vendor B total Medicare payments and Vendor B inpatient payments). The fact that regression coefficients are closer to zero than are the DIM and DID estimates indicates that regression adjustment is the more conservative methodology and would decrease the likelihood of finding a spurious statistically significant impact.

In summary, the regression approach is useful because (1) it controls for imbalances across treatment and control groups that, while statistically insignificant, can affect the impact estimates; and (2) it reduces the residual variance and thus provides more precise impact estimates. This is illustrated below. Moreover, regression is a more conservative approach since it produces impact parameter estimates that are smaller in magnitude than do the other two methodologies.

Exhibit 6.6: DIM, DID, and Regression Impact Estimates – National Arm 2, Total Medicare Payments

Total Medicare Payments ($)		Pre		Post		Estimators[a,b]		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	7310.25	7800.59	6562.23	6198.70	363.53	853.87	518.77
	SE	365.77	359.53	364.42	317.78	483.50	704.86	488.69
	P-value					0.452	0.230	0.288
Vendor A	Mean or beta	7965.60	7433.03	6576.71	7332.56	-755.85	-1288.43	-958.40
	SE	396.34	367.61	417.05	396.96	575.78	789.80	482.21
	P-value					0.189	0.10*	0.047**

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[b] DIM stands for Difference-in-Means; DID stands for Difference-in-Differences.

Exhibit 6.7: Simple Difference, DID, and Regression Impact Estimates – National Arm 2, Inpatient Measures

Any inpatient use		Pre		Post		Estimators[a,b]		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	0.170	0.187	0.143	0.148	-0.005	0.010	-0.002
	SE	0.008	0.009	0.008	0.008	0.011	0.020	0.011
	P-value					0.655	0.440	0.885
Vendor A	Mean or beta	0.163	0.164	0.134	0.154	-0.020	-0.020	-0.021
	SE	0.008	0.008	0.008	0.008	0.011	0.020	0.011
	P-value					0.077*	0.240	0.048**

Inpatient days		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	1.152	1.306	1.216	1.063	0.153	0.310	0.192
	SE	0.104	0.108	0.159	0.115	0.196	0.250	0.179
	P-value					0.434	0.210	0.286
Vendor A	Mean or beta	1.479	1.192	1.177	1.387	-0.210	-0.500	-0.296
	SE	0.134	0.109	0.151	0.132	0.201	0.260	0.177
	P-value					0.295	0.06*	0.095*

Inpatient payment ($)		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	2685.33	3071.42	2620.74	2287.41	333.33	719.42	404.71
	SE	238.32	224.26	244.96	200.56	316.57	455.30	340.56
	P-value					0.292	0.110	0.235
Vendor A	Mean or beta	3039.59	2663.13	2579.35	2976.89	-397.55	-774.01	-491.42
	SE	265.87	224.28	299.09	267.01	400.97	530.84	336.04
	P-value					0.322	0.140	0.144

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[b] DIM stands for Difference-in-Means; DID stands for Difference-in-Differences.

Exhibit 6.8: DIM, DID, and Regression Impact Estimates – National Arm 2, Physician Measures

Physician visits		Pre		Post		Estimators[a,b]		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	20.653	21.065	16.575	16.424	0.151	0.560	0.363
	SE	0.408	0.419	0.356	0.343	0.494	0.770	0.369
	P-value					0.759	0.460	0.325
Vendor A	Mean or beta	21.365	20.887	16.810	17.145	-0.335	-0.810	-0.589
	SE	0.439	0.418	0.360	0.355	0.506	0.790	0.364
	P-value					0.508	0.300	0.106

Physician payment ($)		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	2645.33	2673.96	2126.89	2079.71	47.18	75.82	54.26
	SE	103.88	105.44	89.73	92.66	128.99	196.34	122.67
	P-value					0.715	0.700	0.658
Vendor A	Mean or beta	2633.73	2599.90	2169.43	2352.33	-182.90	-216.73	-199.70
	SE	107.48	103.76	93.70	104.26	140.17	204.86	121.04
	P-value					0.192	0.290	0.099*

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[b] DIM stands for Difference-in-Means; DID stands for Difference-in-Differences.

Exhibit 6.9: DIM, DID, and Regression Impact Estimates – National Arm 2, Outpatient Measures

Any outpatient visits		Pre		Post		Estimators[a,b]		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	0.775	0.768	0.739	0.724	0.015	0.010	0.014
	SE	0.009	0.010	0.010	0.010	0.014	0.020	0.013
	P-value					0.284	0.660	0.260
Vendor A	Mean or beta	0.759	0.762	0.696	0.719	-0.023	-0.020	-0.022
	SE	0.010	0.009	0.010	0.010	0.014	0.020	0.012
	P-value					0.111	0.300	0.083*

Outpatient visits		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	4.886	4.918	3.738	3.855	-0.118	-0.090	-0.102
	SE	0.155	0.157	0.113	0.129	0.172	0.280	0.135
	P-value					0.493	0.760	0.450
Vendor A	Mean or beta	4.811	4.824	3.545	3.803	-0.258	-0.240	-0.248
	SE	0.149	0.176	0.116	0.135	0.179	0.290	0.133
	P-value					0.149	0.400	0.062*

Outpatient payment ($)		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	1350.76	1327.23	1141.45	1096.76	44.69	21.15	33.99
	SE	89.45	86.66	78.47	75.01	108.56	165.21	96.23
	P-value					0.681	0.900	0.724
Vendor A	Mean or beta	1374.01	1361.95	1009.59	1278.52	-268.92	-280.98	-280.27
	SE	81.77	97.09	64.09	91.98	112.08	169.32	94.95
	P-value					0.016***	0.10*	0.003***

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[b] DIM stands for Difference-in-Means; DID stands for Difference-in-Differences.

Exhibit 6.10: DIM, DID, and Regression Impact Estimates – National Arm 2, Emergency Department Measures

Emergency department visits		Pre		Post		Estimators[a,b]		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	0.426	0.397	0.326	0.307	0.018	-0.010	0.007
	SE	0.032	0.021	0.029	0.017	0.034	0.050	0.028
	P-value					0.584	0.850	0.812
Vendor A	Mean or beta	0.474	0.434	0.373	0.352	0.022	-0.020	0.002
	SE	0.034	0.026	0.027	0.023	0.035	0.060	0.027
	P-value					0.536	0.730	0.944

Emergency department Payment ($)		Pre		Post		Estimators		
		Arm 2	Arm 3	Arm 2	Arm 3	DIM	DID	Regression
Vendor B	Mean or beta	97.41	91.69	79.43	77.99	1.44	-4.28	-0.53
	SE	7.06	5.27	6.67	4.95	8.30	12.11	7.57
	P-value					0.863	0.720	0.944
Vendor A	Mean or beta	114.27	105.82	93.59	87.86	5.73	-2.72	1.59
	SE	9.14	7.07	7.54	6.33	9.85	15.18	7.47
	P-value					0.561	0.860	0.831

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[b] DIM stands for Difference-in-Means; DID stands for Difference-in-Differences.

D. Subgroups Driving the Impacts

We estimated the impacts on key Vendor A subgroups to determine whether any of these were driving the large estimated impacts for Vendor A's national Arm 2. The subgroups we used for this purpose are Vendor A's high-risk, medium-risk, and low-risk participants. As explained previously in Section II.B, Vendor A scored the HRA responses and determined risk levels for each targeted health risk behavior for each participant in any arm. The vendor then summed these risk levels using weights based on impactability of risk behavior. The stratification algorithm placed participants with the heaviest burden of coaching-modifiable health risks into a high level of intervention, and those with less risk into moderate or low levels of intervention. The percentage of participants falling into high-risk, moderate-risk, and low-risk strata were 7 percent, 28 percent, and 65 percent, respectively. If this stratification was designed and implemented correctly, one may expect to find largest impacts on the high-risk participants and smallest impacts on the low-risk participants, while impacts for moderate-risk participants will be somewhere in between. In Exhibits 6.11, 6.12, and 6.13 below, we report estimated impacts for high-, moderate, and low-risk participants, respectively. Given that risk stratification is made using baseline characteristics only (Year 1 HRA responses) and that participants were triaged into risk subgroups using the same (proprietary) formula for each treatment arm, we would expect the second randomization to produce comparable Arm 2 and Arm 3 groups for each risk stratum. See Exhibit B.4 in Appendix B for an assessment of balance between high-risk Arm 2 and high-risk Arm 3. Baseline characteristics do not differ from each other at any standard level

of statistical significance, indicating that randomization was successful even for the risk subgroups within each treatment arm.

The results indicate that the Arm 2 impacts are driven by the high-risk and moderate-risk participants (35 percent of all demonstration participants). For high-risk Arm 2 participants in the national sample, the impact on total Medicare payments is a reduction of $6,634, which is seven times the impact for all Arm 2 participants in the sample. It appears that the major source of the dramatic savings is lower inpatient expenditures resulting from lower hospital admission rates and shorter hospital stays. On average, the hospital admission rate is 7.3 percentage points lower, and the number of hospital days is 2.7 days lower for high-risk Arm 2 participants than for high-risk Arm 3 participants. For moderate-risk Arm 2 participants in the national sample, the impacts are not as large, but they are still larger than the impact on all Arm 2 participants. The impact on total Medicare expenditures is a reduction of $1,517.[35] The hospital admission rate was 5.5 percentage points lower for moderate-risk Arm 2 participants than for moderate-risk Arm 3 participants. We did not detect any consistent impacts on low-risk participants.[36]

[35] This is statistically significant at the 10% level.

[36] The impact on outpatient expenditures was -$234 and is statistically significant at the 5% level. However, none of the other estimates were consistent with that estimate.

Exhibit 6.11: Impact of SRRD on Vendor A
National Participants Triaged into High Risk[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample Size[c]	-	-	154	147	-
Medicare Payment					
Total Medicare Payments	-6,634***	0.000	6,536	13,170	-50.4%
Inpatient Payment	-4,344***	0.000	2,544	6,888	-63.1%
Outpatient Payment	-990***	0.004	968	1,958	-50.6%
Carrier Payment	-880**	0.047	1,989	2,870	-30.7%
ED Payment	-32.03	0.241	88.76	120.80	-26.5%
Medicare Use					
Any Inpatient	-7.3%*	0.064	16.6%	23.9%	-30.5%
Inpatient Length of Stay	-2.7***	0.000	1.1	3.8	-70.1%
Any Outpatient	0.4%	0.929	70.6%	70.2%	0.6%
Outpatient Days	-0.4	0.449	3.7	4.1	-9.0%
Carrier Days	-3.6***	0.007	16.4	20.0	-18.1%
ED Visits	-0.2*	0.061	0.4	0.5	-34.7%
Preventive Screening[d]					
Colorectal Cancer	2.5%	0.607	24.9%	22.4%	11.1%
Breast Cancer	-3.1%	0.634	53.4%	56.5%	-5.6%
Cardiovascular	3.9%	0.548	68.0%	64.1%	6.1%

[a] The regression includes only those who completed the Year 1 HRA.
[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.
[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

Exhibit 6.12: Impact of SRRD on Vendor A National Participants Triaged into Medium Risk[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample Size[c]	-	-	553	576	-
Medicare Payment					
Total Medicare Payments	-1,517*	0.095	7,737	9,254	-16.4%
Inpatient Payment	-681	0.283	3,339	4,020	-16.9%
Outpatient Payment	-195	0.276	1,117	1,312	-14.9%
Carrier Payment	-526**	0.021	2,358	2,885	-18.2%
ED Payment	11	0.448	118.87	108.16	9.9%
Medicare Use					
Any Inpatient	-5.5%***	0.007	15.5%	21.0%	-26.3%
Inpatient Length of Stay	-0.5	0.136	1.4	1.9	-26.8%
Any Outpatient	-8.2%***	0.000	69.1%	77.3%	-10.6%
Outpatient Days	-0.7***	0.004	3.6	4.3	-16.9%
Carrier Days	-1.8***	0.009	17.2	19.0	-9.4%
ED Visits	0.0	0.528	0.5	0.4	7.7%
Preventive Screening [d]					
Colorectal Cancer	1.3%	0.614	23.7%	22.4%	5.6%
Breast Cancer	-1.2%	0.730	58.9%	60.0%	-1.9%
Cardiovascular	-1.7%	0.572	63.4%	65.1%	-2.7%

[a] The regression includes only those who completed the Year 1 HRA.
[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.
[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

Exhibit 6.13: Impact of SRRD on Vendor A
National Participants Triaged into Low Risk[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample Size[c]	-	-	1,312	1,293	-
Medicare Payment					
Total Medicare Payments	-32	0.957	5,898	5,930	-0.5%
Inpatient Payment	51	0.903	2,173	2,122	2.4%
Outpatient Payment	-234**	0.047	951	1,185	-19.8%
Carrier Payment	29	0.850	2,097	2,069	1.4%
ED Payment	2	0.824	79	77	2.7%
Medicare Use					
Any Inpatient	0.0%	0.999	12.1%	12.1%	0.0%
Inpatient Length of Stay	0.1	0.757	1.0	1.0	7.1%
Any Outpatient	0.2%	0.874	70.0%	69.8%	0.4%
Outpatient Days	0.0	0.899	3.5	3.6	-0.6%
Carrier Days	0.3	0.507	16.6	16.3	1.8%
ED Visits	0.0	0.720	0.3	0.3	3.9%
Preventive Screening [d]					
Colorectal Cancer	1.7%	0.296	24.1%	22.4%	7.7%
Breast Cancer	0.0%	0.990	63.2%	63.2%	0.0%
Cardiovascular	0.9%	0.632	67.4%	66.5%	1.4%

[a] The regression includes only those who completed the Year 1 HRA.
[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.
[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

Among the special populations of interest is dual eligibles—Medicare beneficiaries who are also eligible for Medicaid. We further divided the Arm 2 participants into those who are also dual eligible. As presented in Exhibit 6.14 below, estimated impacts on total Medicare expenditures and inpatient expenditures are even more dramatic for this group. Despite the small sample size, these estimated impacts were statistically significant at the 5 percent level for total Medicare expenditures (β= -14052, p= 0.009), inpatient expenditures (β= -9258, p= 0.020), number of carrier days (β= -5.9, p= 0.050), and number of ED visits (β= -0.8, p= 0.018).

Exhibit 6.14: Impact of SRRD on Vendor A National Dual Eligible Participants Triaged into High Risk[a]

			Regression-Adjusted Means		
	Coefficient[b]	P-value	Arm 2	Arm 3	Percent Impact
	(1)	(2)	(3)	(4)	(5)=100x(1)/(4)
Sample Size[c]	-	-	38	37	-
Medicare Payment					
Total Medicare Payments	-14,052***	0.009	5,710	19,762	-71.1%
Inpatient Payment	-9,298**	0.020	1,043	10,342	-89.9%
Outpatient Payment	-979	0.250	1,526	2,506	-39.1%
Carrier Payment	-1,194	0.149	1,905	3,100	-38.5%
ED Payment	-182*	0.060	153	336	-54.3%
Medicare Use					
Any Inpatient	-9.3%	0.303	17.1%	26.4%	-35.3%
Inpatient Length of Stay	-4.0*	0.066	0.6	4.6	-87.2%
Any Outpatient	10.2%	0.233	80.1%	69.8%	14.6%
Outpatient Days	-0.9	0.457	4.7	5.6	-15.8%
Carrier Days	-5.9**	0.050	16.5	22.4	-26.3%
ED Visits	-0.8**	0.018	0.4	1.3	-66.1%
Preventive Screening [d]					
Colorectal Cancer	12.1%	0.165	27.6%	15.6%	77.5%
Breast Cancer	5.0%	0.697	44.0%	39.0%	12.7%
Diabetes	4.3%	0.543	15.3%	11.1%	38.6%
Cardiovascular	9.1%	0.531	65.1%	56.0%	16.3%

[a] The regression includes only those who completed at the Year 1 HRA.
[b] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.
[c] The actual regression included national sample participants in the three treatment arms for both vendors, and interaction terms were used to estimate the impacts of Arm 2.
[d] Sample sizes for preventive screening differ due to exclusions (e.g., breast cancer screening was calculated only for females without a previous year diagnosis of breast cancer).

E. Characteristics of the Subgroups Driving the Results

After finding that high-risk and moderate-risk participants were driving the impacts estimated for Vendor A's national Arm 2, we identified the baseline characteristics of these participants. Exhibit 6.15 presents a number of baseline demographic, Medicare expenditure, Medicare use, health status, and preventive use characteristics for high-, medium-, and low-risk Arm 2 participants as well as for all Arm 2 participants. We also show the differences in baseline means between the risk strata and all Arm 2 participants and indicate whether the differences are statistically significant. For these calculations, we combined participants with Year 1 HRAs (that is, first-year participants) in Vendor A's local and national sample, pooled across all three arms. Note that Vendor A (as well as Vendor B) did not have access to Medicare claims data and constructed the risk strata purely based on HRA responses.

There is a very clear pattern in the three risk strata: the high-risk stratum includes the largest proportion of African Americans, Hispanics, and dual eligibles. The representation of dual eligibles, in particular, was very uneven: 25.2 percent of the high-risk stratum was dual eligible compared to only 12.6 percent and 5.6 percent of the moderate-risk and low-risk strata, respectively.

The high-risk stratum included beneficiaries with the highest baseline utilization of Medicare covered health services and expenditures and yet the lowest utilization of Medicare covered preventive screenings. Participants in the high-risk stratum also seem to be less healthy than others as indicated by the lowest self-rated health assessments and the highest incidences of almost all of the 11 chronic conditions (except osteoporosis) that we identified from the 2009 chronic condition summary file.

Exhibit 6.15: Baseline Characteristics of Vendor A's High-, Medium-, and Low-Risk Participants with a Year 1 HRA

Baseline beneficiary characteristics[a]	All participants with a Year 1 HRA	High-risk participants with a Year 1 HRA		Medium-risk participants with a Year 1 HRA		Low-risk participants with a Year 1 HRA	
	Mean or %	Mean or %	Diff.	Mean or %	Diff.	Mean or %	Diff.
	(1)	(2)	(3)= (2)-(1)	(4)	(5)= (4)-(1)	(6)	(7)= (6)-(1)
	N=7,473	N=511		N=2,100		N=4,861	
Demographic characteristics							
Female (%)	55.8	58.7	2.9	59.3	3.5**	54.0	-1.8**
Age (as of 5/1/09)	70.2	70.1	-0.1	70.1	-0.1	70.2	0.0
Race (RTI race code) (%)							
Non-Hispanic White	87.7	82.4	-5.3**	87.1	-0.6	88.5	0.8
African American	5.2	8.0	2.9**	6.2	1.0*	4.4	-0.7*
Hispanic	3.5	5.7	2.2**	4.1	0.6	3.0	-0.5
Asian/Pacific Islander	2.3	2.2	-0.2	1.6	-0.7**	2.7	0.3
Medicaid/Medicare dual eligibility	8.9	25.2	16.3**	12.6	3.7**	5.6	-3.3**
Medicare expenditures							
Total Medicare payment	7,573	11,882	4,309**	9,229	1,656**	6,406	-1,167**
Total Medicare payment in 1st (bottom) quartile	20.6	19.0	-1.6	20.1	-0.5	21.0	0.4
Total Medicare payment in 2nd quartile	26.8	13.9	-12.9**	23.6	-3.3**	29.6	2.8**
Total Medicare payment in 3rd quartile	27.4	24.7	-2.7	27.1	-0.3	27.8	0.4
Total Medicare payment in 4th (top) quartile	25.2	42.5	17.2**	29.2	4.0**	21.7	-3.5**
Inpatient payment	2,812	4,500	1,689**	3,724	912**	2,241	-571**
Physician payment	2,551	3,681	1,129**	2,769	218*	2,339	-213**
Outpatient payment	1,394	1,861	467**	1,607	213**	1,253	-141*
Medicare use							
Number of inpatient days	1.3	2.6	1.3**	1.8	0.5**	1.0	-0.3**
Number of days with physician visit	21.0	26.0	5.0**	22.1	1.1**	19.9	-1.0**
Number of days with outpatient visit	5.2	7.0	1.8**	6.0	0.8**	4.7	-0.5**
Health status[b] and chronic conditions (%)[c]							
Self-rated health status (1=poor, 5=excellent)	3.3	2.3	-1.0**	2.9	-0.3**	3.5	0.2**
Alzheimer's/related disorders/senile dementia	2.4	3.3	1.0	3.0	0.7*	2.0	-0.4
Cancer	7.0	7.8	0.8	7.6	0.6	6.7	-0.3
Heart failure	7.9	18.0	10.1**	9.9	2.0**	6.0	-1.9**
Chronic kidney disease	8.0	15.1	7.1**	10.3	2.3**	6.3	-1.7**

Baseline beneficiary characteristics[a]	All participants with a Year 1 HRA	High-risk participants with a Year 1 HRA		Medium-risk participants with a Year 1 HRA		Low-risk participants with a Year 1 HRA	
	Mean or %	Mean or %	Diff.	Mean or %	Diff.	Mean or %	Diff.
	(1)	(2)	(3)= (2)-(1)	(4)	(5)= (4)-(1)	(6)	(7)= (6)-(1)
	N=7,473	N=511		N=2,100		N=4,861	
COPD	8.0	22.1	14.1**	10.9	2.9**	5.3	-2.7**
Depression	9.0	18.4	9.4**	12.5	3.6**	6.4	-2.5**
Diabetes	24.0	40.5	16.5**	30.9	6.9**	19.3	-4.7**
Ischemic heart disease	26.0	36.6	10.6**	28.2	2.2**	23.9	-2.1**
Osteoporosis	12.7	10.6	-2.2	11.2	-1.5*	13.6	0.9
Stroke/TIA	2.1	4.7	2.6**	2.5	0.4	1.6	-0.5*
Rheumatoid/ osteoarthritis	19.1	23.9	4.8**	22.4	3.3**	17.1	-1.9**
Preventive screening use (%)							
Colorectal CA screening	25.6	23.4	-2.2	23.2	-2.4**	26.8	1.2
Breast cancer screening	67.2	49.1	-18.1**	60.1	-7.2**	72.7	5.4**
Cardiovascular screening	73.3	57.1	-16.2**	65.8	-7.4**	77.6	4.4**

[a] Characteristics are reported for the baseline period from January 1, 2008 to April 30, 2009, unless indicated otherwise. *indicates p <.10, ** indicates p <.05

[b] Health status is based on Vendor A's HRA item "In general, how would you describe your health?" Possible responses include: Poor (1), Fair (2), Good (3), Very Good (4), Excellent (5)

[c] Chronic conditions were measured at mid-year 2009.

VII. CONCLUSIONS

SRRD is a CMS demonstration in which wellness, prevention, and health promotion programs that have been developed and tested in the private sector were tailored to Medicare beneficiaries to improve their health and reduce avoidable health services utilization. The demonstration was implemented throughout the United States as a randomized controlled trial in a real-world setting. At the core of the demonstration is the use of a health risk assessment (HRA), through which participants' health risks were measured. The results of the HRAs were then followed by risk stratification and triaging of follow-up services to participants. Two vendors developed and implemented the program both nationally, with about 6,000 participants each, and also in two local areas working with the local Aging and Disability Resource Centers, with about 1,500 participants each. Participants were randomly assigned to one of three study groups: Intervention Arm 1 – Standard Treatment (HRA + standard tailored follow-up), Intervention Arm 2 – Enhanced Treatment (HRA + enhanced tailored follow-up), and Arm 3 – HRA Only (HRA + generic health advice).

In this report, we examined the impact of the SRRD after one year of implementation. We found that Vendor A's national enhanced intervention (Arm 2) reduced key Medicare utilization rates and expenditures: the national enhanced intervention participants, on average, cost Medicare $958 less and were 14.2 percent less likely to be hospitalized than Vendor A's national randomly selected controls. These are significant program impacts provided that the impact of the control intervention (HRA + generic health advice) is minimal. We did not find any statistically significant impact of Vendor B's national enhanced intervention. The impact of Vendor B's national standard intervention (Arm 1) net of its control intervention was to increase total and carrier Medicare expenditures by $822 and $273, respectively. However, the impact on total Medicare expenditures was only significant at the 10 percent level.

Given these dramatic early impacts, we conducted further in-depth analyses, which confirmed our original estimates. Impacts were driven by participants that Vendor A stratified into high-risk and moderate-risk groups based on its proprietary HRA risk-scoring algorithm. The high-risk group represented about 7 percent of participants and experienced an impact of about $6,600 reduction in total Medicare Part A and B expenditures. The moderate-risk group constituted 28 percent of participants and experienced an impact of about $1,500 reduction in total Medicare Part A and B expenditures. The most dramatic impacts were estimated for dual eligibles (those participants who were eligible for both Medicare and Medicaid) that were in the high-risk group: On average, $14,000 reduction in total Medicare Part A and B expenditures. Even though statistically significant, this extremely large estimate may be a result of small sample size (n=75). The impact on dual eligible beneficiaries should examined in more detail in the final evaluation report.

We hypothesize that the impacts for Vendor A may be due to the use of (1) a more personalized recruitment strategy and telematching to find beneficiary telephone numbers CMS may not have, which may help in recruiting hard-to-recruit but highly impactable

beneficiaries, such as ethnic minorities and dual eligibles; and (2) an effective risk-scoring algorithm and triaging, so that beneficiaries who have the greatest need and are most impactable receive the most intensive services (Vendor A included more risk factors in the overall risk determination and developed its stratification algorithm for this particular demonstration rather than using existing predictive models developed on different populations); and (3) allowing beneficiaries to opt out easily, which may have helped to focus on beneficiaries who explicitly wanted to be part of the program and thus presumably were more likely to be impacted.

The results of this interim evaluation are suggestive but not conclusive. We believe that additional analysis is needed to confirm that the approaches implemented by the demonstration vendors lead to sustainable reductions in Medicare expenditures. More detailed data elements, such as admission reasons, diagnosis codes, procedure codes, and place of service, should be considered in a more qualitative and more clinical analysis, where the circumstances of each high-risk participant is examined and compared to those of high-risk control group participants. The final analysis of SRRD will shed additional light on the effectiveness of SRRD.

APPENDIX A– PREVENTIVE SCREENING CODES

Colorectal Cancer Screening		
Description	Beneficiaries with one or more of the following: ▪ Fecal occult blood test ▪ Flexible sigmoidoscopy ▪ Colonoscopy ▪ Barium enema	Exclusion: Patients with colorectal cancer from CCW chronic care condition file for 2009 Claims Files: All 7 files CC_Screen = 1 Any of the qualifying codes is present CC_Screen = 0 Otherwise
Qualifying Codes		**Code Description**
HCPCS	G0104	Colorectal cancer screening; flexible sigmoidoscopy
HCPCS	G0105	Colorectal cancer screening; colonoscopy on individual at high risk
HCPCS	G0106	Colorectal cancer screening; alternative to G0104, screening sigmoidoscopy, barium enema
HCPCS	G0107	Colorectal cancer screening; fecal-occult blood test, 1-3 simultaneous determinations
CPT	82270	Blood, occult, by peroxidase activity (e.g., guaiac), qualitative; feces, consecutive collected specimens with single determination, for colorectal neoplasm screening (i.e., patient was provided 3 cards or single triple card for consecutive collection)
HCPCS	G0120	Colorectal cancer screening; alternative to G0105, screening colonoscopy, barium enema
HCPCS	G0121	Colorectal cancer screening; colonoscopy on individual not meeting criteria for high risk
HCPCS	G0122	Colorectal cancer screening; barium enema (not covered)
HCPCS	G0328	Colorectal cancer screening; fecal occult blood test, immunoassay,1-3 simultaneous
CPT	82274	Diagnostic fecal occult blood test (FOBT)
HCPCS	G0394	FOBT
CPT	45330-45335, 45337-45342, 45345	Flexible sigmoidoscopy
ICD-9 Procedure	45.24	Double contrast barium enema
CPT	44388-44394, 44397, 45355, 45378-45387, 45391, 45392	Colonoscopy
ICD-9 Procedure	45.22, 45.23, 45.25, 45.42, 45.43	Colonoscopy
References: CMS (2009). *The Guide to Medicare Preventive Services for Physicians, Providers, Suppliers, and Other Health Care Professionals,* 3rd ed. http://www.ncqa.org/Portals/0/PublicComment/HEDIS2010/NCQA_HEDIS_2010_Public_Comment_COL.pdf		

Breast Cancer Screening		
Description	Female beneficiaries with one or more mammograms	Exclusion: Males and patients with breast cancer in CCW chronic conditions file for 2009 Claims Files: All 7 files BC_Screen = 1 Any of the qualifying codes is present BC_Screen = 0 Otherwise
Qualifying Codes		**Code Description**
CPT	77052	Computer-aided detection (computer algorithm analysis of digital image data for lesion detection) with further physician review for interpretation, with or without digitization of film radiographic images; screening mammography (list separately in addition to code for primary procedure) (Use 77052 in conjunction with 77057)
CPT	77055	Mammography; unilateral (Use 77055 in conjunction with 77051 for computer-aided detection applied to a diagnostic mammogram)
CPT	77056	Mammogram; bilateral (Use 77056 in conjunction with 77051 for computer-aided detection applied to a diagnostic mammogram)
CPT	77057	Screening mammography, bilateral (2-view film study of each breast) (Use 77057 in conjunction with 77052 for computer-aided detection applied to a screening mammogram) (For electrical impedance breast scan, use 76499)
HCPCS	G0202	Screening mammography, producing direct digital image, bilateral, all views
ICD-9 Procedure	87.36	Xerography of breast
ICD-9 Procedure	87.37	Other mammography
ICD-9 Diagnosis	V76.11	Screening mammogram for high-risk patient
ICD-9 Diagnosis	V76.12	Other screening mammogram
Revenue Center	0403	Other imaging services-screening mammography (effective 1/1/91)
References: CMS (2009). *The Guide to Medicare Preventive Services for Physicians, Providers, Suppliers, and Other Health Care Professionals,* 3rd ed. http://library.ahima.org/xpedio/groups/public/documents/ahima/bok1_033059.hcsp?dDocName=bok1_033059		

Cardiovascular Screening		
Description	Beneficiaries screened with: ▪ Total cholesterol test ▪ Cholesterol test for HDLs ▪ Triglycerides test	Exclusion: beneficiaries with cardiovascular disease (CVD includes ischemic heart disease, heart failure, acute myocardial infarction) from CCW chronic condition file for 2009 Claims Files: All 7 files CV_Screen = 1 Any of the qualifying codes is present CV_Screen = 0 Otherwise
Qualifying Codes		**Code Description**
CPT	80061	Lipid panel
CPT	82465	Cholesterol, serum, total
CPT	83718	Lipoprotein (HDL cholesterol)
CPT	84478	Triglycerides
References: CMS (2009). *The Guide to Medicare Preventive Services for Physicians, Providers, Suppliers, and Other Health Care Professionals,* 3rd ed.		

APPENDIX B– MISCELLANEOUS TABLES

Exhibit B.1: Vendor A Beneficiary Baseline Characteristics: National Sample IG versus ACG[a]

Baseline Characteristics	National IG	National ACG	Difference: IG–ACG	P-Value
Sample Size	16,290	16,290	-	-
Medicare Expenditures				
Total Medicare Payments	$7,633	$7,673	-$40	0.841
Inpatient Payment	$2,904	$2,902	$2	0.984
Outpatient Payment	$1,279	$1,328	-$49	0.296
Carrier Payment	$2,543	$2,531	$12	0.838
ED Payment	$104	$107	-$4	0.287
Medicare Use				
Any Inpatient	17.2%	16.8%	0.5%	0.275
Inpatient Days	1.4	1.5	0.0	0.688
Any Outpatient	72.0%	72.1%	-0.1%	0.843
Outpatient Days	4.4	4.5	-0.1	0.324
Carrier Days	19.9	19.9	0.0	0.996
ED Visits	0.5	0.5	0.0	0.715
Demographics				
Female	55.6%	55.5%	0.1%	0.894
Age	70.9	70.9	0.0	0.974
Race: Non-Hispanic White	83.5%	83.4%	0.1%	0.800
Race: African American	7.5%	7.4%	0.1%	0.849
Race: Hispanic	5.0%	5.3%	-0.3%	0.240
Race: Asian/Pacific Islander	2.7%	2.5%	0.2%	0.263
Race: Other/Unknown	1.3%	1.3%	-0.1%	0.660
Dual Eligibility Status	11.0%	10.6%	0.4%	0.260
Health Status				
Multiple Chronic Conditions	35.4%	35.3%	0.1%	0.853
HCC Risk Score (2009)[b]	0.86	0.86	0.00	0.697

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[b] Due to missing values, the sample sizes for the HCC Risk Score are 16,283 for the IG and 16,280 for the ACG.

Exhibit B.2: Vendor B Beneficiary Baseline Characteristics: National Sample IG versus ACG

Baseline Characteristics	National IG	National ACG	Difference: IG–ACG [a]	P-Value
Sample Size	16,290	16,290	-	-
Medicare Expenditures				
Total Medicare Payments	$7,919	$8,257	-$339	0.116
Inpatient Payment	$3,040	$3,309	-269*	0.056
Outpatient Payment	$1,346	$1,329	$17	0.703
Carrier Payment	$2,597	$2,627	-$31	0.595
ED Payment	$110	$107	$2	0.541
Medicare Use				
Any Inpatient	17.3%	18.5%	-1.2%***	0.006
Inpatient Length of Stay	1.6	1.6	-0.1	0.538
Any Outpatient	72.5%	72.5%	-0.1%	0.911
Outpatient Days	4.6	4.6	0.0	0.786
Carrier Days	20.1	20.2	-0.2	0.485
ED Visits	0.5	0.5	0.0	0.532
Demographics				
Female	55.1%	54.5%	0.7%	0.234
Age	70.9	70.9	0.0	0.315
Race: Non-Hispanic White	83.0%	82.9%	0.1%	0.735
Race: African American	7.8%	8.1%	-0.2%	0.413
Race: Hispanic	5.2%	5.2%	0.0%	0.881
Race: Asian/Pacific Islander	2.8%	2.6%	0.2%	0.304
Race: Other/Unknown	1.2%	1.2%	0.0%	0.721
Dual Eligibility Status	11.2%	10.9%	0.3%	0.427
Health Status				
Multiple Chronic Conditions	36.5%	36.4%	0.1%	0.863
HCC Risk Score (2009)[b]	0.87	0.88	-0.01	0.191

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[b] Due to missing values, the sample sizes for the HCC Risk Score are 16,279 for the IG and 16,281 for the ACG.

Exhibit B.3: Number of High-Risk Arm 2 Participants Not Re-Enrolling ("Drop-Outs") by Selected HRA Responses

HRA Response	Number of High-Risk Participants Taking Year 1 HRA (1)	Number of Participants Taking Year 2 HRA (2)	Number of Year 1 High-Risk "Drop-Outs" (3)=(1)-(2)	Year 1 High-Risk "Drop-Out" Rate (4)=(3)/(1)
Vendor A National Sample				
Self-Reported Health Status = Poor/Fair[a]	397	150	247	62%
Self-Reported Health Status = Good[a]	778	420	358	46%
Self-Reported Health Status = Very Good/Excellent[a]	775	434	341	44%
Risk Stratum	154	59	95	62%
Alcohol Risk	43	23	20	47%
Diet Risk	395	173	222	56%
Preventive Services Risk	31	15	16	52%
Physical Activity Risk	420	151	269	64%
Tobacco Risk	183	91	92	50%
Polypharmacy Risk	80	39	41	51%
Vendor B National Sample				
Self-Reported Health Status = Very Poor/Poor/Fair[a]	365	158	207	57%
Self-Reported Health Status = Good[a]	591	278	313	53%
Self_Reported Health Status = Very Good/Excellent[a]	993	519	474	48%
Alcohol Risk	73	35	38	52%
Eating Risk	393	181	212	54%
Exams Risk	64	30	34	53%
Exercise Risk	462	190	272	59%
Smoking Risk	36	15	21	58%
Self-care Risk	43	19	24	56%

[a] For the self-reported health status variables, the values are the number of participants reporting the listed response (unrelated to risk levels).

Exhibit B.4: Vendor A Beneficiary Baseline Characteristics: High-Risk National Sample Arm 2 versus Arm 3

Baseline Characteristics	National Arm 2	National Arm 3	Difference: [a] Arm 2 – Arm 3	P-Value
Sample Size	154	147		
Medicare Expenditures				
Total Medicare Payments	$10,909	$12,105	-1,196	0.589
Inpatient Payment	$4,846	$4,946	-100	0.944
Outpatient Payment	$1,552	$1,420	132	0.688
Carrier Payment	$3,191	$3,731	-541	0.394
ED Payment	$202	$241	-38	0.517
Medicare Use				
Any Inpatient	29.2%	27.9%	1.3%	0.799
Inpatient Length of Stay	2.9	2.9	0.0	0.995
Any Outpatient	83.1%	76.9%	6.2%	0.175
Outpatient Days	6.4	5.3	1.1	0.224
Carrier Days	26.6	24.8	1.8	0.524
ED Visits	0.9	1.0	-0.1	0.772
Demographics				
Female	56.5%	61.9%	-5.4%	0.340
Age	70.6	70.9	-0.2	0.349
Race: Non-Hispanic White	81.8%	82.3%	-0.5%	0.911
Race: African American	7.8%	8.2%	-0.4%	0.905
Race: Hispanic	5.2%	6.8%	-1.6%	0.556
Race: Asian/Pacific Islander	1.9%	1.4%	0.6%	0.690
Race: Other/Unknown	3.2%	1.4%	1.9%	0.278
Dual Eligibility Status	24.0%	25.9%	-1.8%	0.715
Health Status				
Multiple Chronic Conditions	57.1%	56.5%	0.7%	0.905
HCC Risk Score (2009)	1.24	1.35	-0.11	0.394
HRA Variables [b]				
Self-Reported Health Status = Very Good/Excellent	7.4%	5.6%	1.8%	0.526
Self-Reported Health Status = Good	31.8%	29.4%	2.4%	0.659
Self-Reported Health Status = Poor/Fair	60.8%	65.0%	-4.2%	0.456
Self-Reported Health Status (1=very poor; 5=very good)	2.3	2.2	0.1	0.501

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

[b] Due to missing values, the sample sizes for the HRA variables are 148 for Arm 2 and 142 for Arm.

APPENDIX C - IMPACT RESULTS

Exhibit C.1: Sample Sizes by Comparison Groups

Exhibit Number	Treatment Group	Sample Size	Control Group	Sample Size
C.2	Arm 1	4,947	Arm 3	4,947
C.3	Arm 2	4,954	Arm 3	4,947
C.4	Arm 12	4,947	Arm 3	4,947
C.5	National Arm 1	3,985	National Arm 3	3,981
C.6	National Arm 2	3,985	National Arm 3	3,981
C.7	National Arms 1 and 2 Pooled	7,968	National Arm 3	3,981
C.8	Local Arm 1	962	Local Arm 3	966
C.9	Local Arm 2	971	Local Arm 3	966
C.10	Local Arms 1 and 2 Pooled	1,933	Local Arm 3	966
C.11	Vendor A Arm 1	2,489	Vendor A Arm 3	2,489
C.12	Vendor B Arm 1	2,458	Vendor B Arm 3	2,458
C.13	Vendor A Arm 2	2,495	Vendor A Arm 3	2,489
C.14	Vendor B Arm 2	2,459	Vendor B Arm 3	2,458
C.15	Vendor A Arms 1 and 2 Pooled	4,984	Vendor A Arm 3	2,489
C.16	Vendor B Arms 1 and 2 Pooled	4,917	Vendor B Arm 3	2,458
C.17	National Vendor A Arm 1	2,021	National Vendor A Arm 3	2,016
C.18	National Vendor A Arm 2	2,019	National Vendor A Arm 3	2,016
C.19	National Vendor A Arms 1 and 2 Pooled	4,040	National Vendor A Arm 3	2,016
C.20	Local Vendor A Arm 1	468	Local Vendor A Arm 3	473
C.21	Local Vendor A Arm 2	476	Local Vendor A Arm 3	473
C.22	Local Vendor A Arms 1 and 2 Pooled	944	Local Vendor A Arm 3	473
C.23	National Vendor B Arm 1	1,964	National Vendor B Arm 3	1,965
C.24	National Vendor B Arm 2	1,964	National Vendor B Arm 3	1,965
C.25	National Vendor B Arms 1 and 2 Pooled	3,928	National Vendor B Arm 3	1,965
C.26	Local Vendor B Arm 1	494	Local Vendor B Arm 3	493
C.27	Local Vendor B Arm 2	495	Local Vendor B Arm 3	493
C.28	Local Vendor B Arms 1 and 2 Pooled	989	Local Vendor B Arm 3	493

Exhibit C.2: Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	18.9	0.951	6,733.0	6,714.1
	Inpatient Payment	-172.4	0.411	2,444.8	2,617.1
	Outpatient Payment	27.6	0.678	1,202.8	1,175.2
	Carrier Payment	59.8	0.436	2,241.6	2,181.8
	ED Payment	6.4	0.184	90.1	83.7
Utilization	Any Inpatient (%)	-0.9	0.186	13.8	14.7
	Inpatient Days	-0.1	0.444	1.1	1.2
	Any Outpatient (%)	0.9	0.243	74.1	73.2
	Outpatient Days	0.0	0.961	3.9	3.9
	Carrier Days	-0.1	0.683	16.7	16.8
	ED Visits	0.0	0.148	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	0.3	0.695	23.0	22.6
	Breast Cancer (%)	0.0	0.981	61.0	61.0
	Cardiovascular (%)	1.5	0.144	66.4	64.8

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.3: Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-69.3	0.821	6,674.2	6,743.5
	Inpatient Payment	-24.2	0.908	2,543.6	2,567.8
	Outpatient Payment	-56.2	0.397	1,147.0	1,203.1
	Carrier Payment	-56.3	0.463	2,164.2	2,220.5
	ED Payment	0.8	0.870	86.4	85.6
Utilization	Any Inpatient (%)	-1.2*	0.074	13.6	14.8
	Inpatient Days	-0.0	0.930	1.2	1.2
	Any Outpatient (%)	-0.2	0.762	73.3	73.6
	Outpatient Days	-0.2**	0.034	3.8	4.0
	Carrier Days	-0.1	0.535	16.7	16.8
	ED Visits	0.0	0.851	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.8	0.370	23.2	22.5
	Breast Cancer (%)	0.0	0.985	61.0	61.0
	Cardiovascular (%)	1.0	0.310	66.1	65.0

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.4: Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-25.3	0.924	6,712.0	6,737.2
	Inpatient Payment	-98.2	0.588	2,527.0	2,625.2
	Outpatient Payment	-14.3	0.803	1,179.6	1,194.0
	Carrier Payment	1.7	0.980	2,202.3	2,200.6
	ED Payment	3.6	0.390	87.0	83.5
Utilization	Any Inpatient (%)	-1.1*	0.073	14.0	15.1
	Inpatient Days	-0.0	0.623	1.1	1.2
	Any Outpatient (%)	0.3	0.618	73.6	73.3
	Outpatient Days	-0.1	0.233	3.9	4.0
	Carrier Days	-0.1	0.553	16.7	16.8
	ED Visits	0.0	0.345	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.5	0.457	22.9	22.4
	Breast Cancer (%)	0.0	0.981	61.0	61.0
	Cardiovascular (%)	1.3	0.153	65.8	64.5

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.5: National Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	48.6	0.887	6,752.8	6,704.2
	Inpatient Payment	-128.5	0.582	2,474.0	2,602.5
	Outpatient Payment	7.0	0.924	1,189.1	1,182.1
	Carrier Payment	63.1	0.461	2,243.8	2,180.7
	ED Payment	3.3	0.541	88.0	84.8
Utilization	Any Inpatient (%)	-0.6	0.460	14.0	14.6
	Inpatient Days	-0.1	0.540	1.1	1.2
	Any Outpatient (%)	1.4	0.107	74.4	73.0
	Outpatient Days	0.1	0.443	4.0	3.9
	Carrier Days	0.0	0.924	16.8	16.8
	ED Visits	0.0	0.230	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	0.2	0.803	22.9	22.7
	Breast Cancer (%)	0.1	0.932	61.1	60.9
	Cardiovascular (%)	1.6	0.171	66.4	64.8

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.6: National Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-222.4	0.515	6,572.2	6,794.6
	Inpatient Payment	-46.2	0.843	2,528.9	2,575.1
	Outpatient Payment	-124.3*	0.093	1,101.6	1,225.9
	Carrier Payment	-73.0	0.394	2,153.1	2,226.1
	ED Payment	1.0	0.853	86.5	85.5
Utilization	Any Inpatient (%)	-1.2	0.131	13.6	14.7
	Inpatient Days	-0.1	0.673	1.1	1.2
	Any Outpatient (%)	-0.4	0.655	73.2	73.6
	Outpatient Days	-0.2*	0.080	3.8	4.0
	Carrier Days	-0.1	0.646	16.7	16.8
	ED Visits	0.0	0.751	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	1.1	0.242	23.5	22.4
	Breast Cancer (%)	-0.2	0.869	60.8	61.1
	Cardiovascular (%)	0.7	0.557	65.8	65.1

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.7: National Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-86.9	0.769	6,691.4	6,778.3
	Inpatient Payment	-87.4	0.666	2,530.6	2,618.0
	Outpatient Payment	-58.6	0.360	1,164.9	1,223.5
	Carrier Payment	-4.9	0.947	2,200.1	2,205.0
	ED Payment	2.1	0.645	86.6	84.4
Utilization	Any Inpatient (%)	-0.9	0.194	14.1	14.9
	Inpatient Days	-0.1	0.550	1.1	1.2
	Any Outpatient (%)	0.5	0.501	73.7	73.1
	Outpatient Days	-0.0	0.571	3.9	4.0
	Carrier Days	-0.0	0.834	16.8	16.8
	ED Visits	0.0	0.381	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.7	0.412	23.0	22.3
	Breast Cancer (%)	-0.0	0.965	61.0	61.0
	Cardiovascular (%)	1.1	0.259	65.7	64.6

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.8: Local Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-103.7	0.881	6,610.3	6,714.1
	Inpatient Payment	-353.7	0.456	2,263.4	2,617.1
	Outpatient Payment	112.7	0.453	1,287.9	1,175.2
	Carrier Payment	45.8	0.792	2,227.6	2,181.8
	ED Payment	19.2*	0.078	102.9	83.7
Utilization	Any Inpatient (%)	-2.3	0.136	12.4	14.7
	Inpatient Days	-0.1	0.623	1.1	1.2
	Any Outpatient (%)	-1.1	0.536	72.1	73.2
	Outpatient Days	-0.3	0.149	3.6	3.9
	Carrier Days	-0.6	0.263	16.2	16.8
	ED Visits	0.0	0.401	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	0.7	0.703	23.4	22.6
	Breast Cancer (%)	-0.4	0.893	60.6	61.0
	Cardiovascular (%)	1.1	0.619	66.0	64.9

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.9: Local Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	558.0	0.420	7,301.6	6,743.6
	Inpatient Payment	66.0	0.889	2,633.8	2,567.8
	Outpatient Payment	223.2	0.137	1,426.4	1,203.2
	Carrier Payment	12.3	0.944	2,232.8	2,220.5
	ED Payment	-0.1	0.995	85.5	85.6
Utilization	Any Inpatient (%)	-1.5	0.327	13.3	14.8
	Inpatient Days	0.2	0.510	1.3	1.2
	Any Outpatient (%)	0.4	0.831	73.9	73.6
	Outpatient Days	-0.3	0.213	3.8	4.0
	Carrier Days	-0.2	0.636	16.6	16.8
	ED Visits	-0.0	0.830	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	-0.6	0.733	21.8	22.5
	Breast Cancer (%)	1.0	0.702	62.0	61.0
	Cardiovascular (%)	2.5	0.286	67.5	65.0

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.10: Local Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	228.7	0.703	6,965.9	6,737.2
	Inpatient Payment	-142.9	0.728	2,482.3	2,625.2
	Outpatient Payment	168.2	0.196	1,362.2	1,194.0
	Carrier Payment	29.0	0.847	2,229.6	2,200.6
	ED Payment	9.5	0.312	93.0	83.5
Utilization	Any Inpatient (%)	-1.9	0.154	13.2	15.1
	Inpatient Days	0.0	0.921	1.2	1.2
	Any Outpatient (%)	-0.4	0.816	72.9	73.3
	Outpatient Days	-0.3	0.121	3.7	4.0
	Carrier Days	-0.4	0.358	16.4	16.8
	ED Visits	0.0	0.720	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.0	0.983	22.4	22.4
	Breast Cancer (%)	0.3	0.885	61.3	61.0
	Cardiovascular (%)	1.8	0.367	66.3	64.5

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.11: Vendor A Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient [a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-630.7	0.144	6,083.3	6,714.1
	Inpatient Payment	-628.6**	0.033	1,988.5	2,617.1
	Outpatient Payment	9.0	0.923	1,184.3	1,175.2
	Carrier Payment	-117.3	0.279	2,064.6	2,181.8
	ED Payment	4.9	0.469	88.6	83.7
Utilization	Any Inpatient (%)	-1.0	0.312	13.7	14.7
	Inpatient Days	-0.3*	0.077	0.9	1.2
	Any Outpatient (%)	0.8	0.456	74.0	73.2
	Outpatient Days	0.1	0.329	4.1	3.9
	Carrier Days	-0.4	0.214	16.4	16.8
	ED Visits	0.0	0.284	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	1.4	0.241	24.0	22.6
	Breast Cancer (%)	-2.2	0.184	58.8	61.0
	Cardiovascular (%)	2.9**	0.045	67.8	64.8

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.12: Vendor B Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	676.8	0.119	7,171.7	6,494.9
	Inpatient Payment	289.7	0.330	2,752.9	2,463.2
	Outpatient Payment	46.4	0.622	1,215.3	1,169.0
	Carrier Payment	239.1**	0.028	2,361.2	2,122.1
	ED Payment	7.8	0.249	91.1	83.2
Utilization	Any Inpatient (%)	-0.8	0.391	13.8	14.6
	Inpatient Days	0.1	0.487	1.2	1.1
	Any Outpatient (%)	1.0	0.364	74.2	73.1
	Outpatient Days	-0.1	0.361	3.9	4.0
	Carrier Days	0.2	0.502	16.9	16.7
	ED Visits	0.0	0.329	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	-0.7	0.536	22.3	23.0
	Breast Cancer (%)	2.3	0.164	62.5	60.2
	Cardiovascular (%)	0.1	0.954	65.4	65.3

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.13: Vendor A Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-758.9*	0.078	5,984.7	6,743.7
	Inpatient Payment	-398.4	0.177	2,169.5	2,567.9
	Outpatient Payment	-229.2**	0.014	974.0	1,203.2
	Carrier Payment	-161.0	0.137	2,059.6	2,220.6
	ED Payment	3.7	0.582	89.3	85.6
Utilization	Any Inpatient (%)	-2.0**	0.039	12.8	14.8
	Inpatient Days	-0.2	0.186	1.0	1.2
	Any Outpatient (%)	-1.7	0.124	71.9	73.6
	Outpatient Days	-0.2*	0.055	3.8	4.0
	Carrier Days	-0.5	0.107	16.3	16.8
	ED Visits	0.0	0.696	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	1.3	0.272	23.8	22.5
	Breast Cancer (%)	0.2	0.901	61.2	61.0
	Cardiovascular (%)	1.5	0.314	66.5	65.0

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.14: Vendor B Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	629.3	0.147	7,139.8	6,510.4
	Inpatient Payment	354.9	0.232	2,796.2	2,441.3
	Outpatient Payment	119.2	0.205	1,263.9	1,144.6
	Carrier Payment	49.6	0.649	2,234.8	2,185.2
	ED Payment	-2.2	0.746	84.4	86.6
Utilization	Any Inpatient (%)	-0.4	0.643	14.1	14.5
	Inpatient Days	0.2	0.227	1.3	1.1
	Any Outpatient (%)	1.2	0.263	74.3	73.1
	Outpatient Days	-0.1	0.285	3.9	4.0
	Carrier Days	0.2	0.459	16.9	16.7
	ED Visits	-0.0	0.899	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.2	0.871	22.9	22.7
	Breast Cancer (%)	-0.1	0.935	60.9	61.0
	Cardiovascular (%)	0.6	0.677	65.8	65.1

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.15: Vendor A Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-694.9*	0.063	6,042.5	6,737.4
	Inpatient Payment	-513.4**	0.045	2,111.9	2,625.3
	Outpatient Payment	-110.2	0.174	1,083.8	1,194.0
	Carrier Payment	-139.1	0.138	2,061.5	2,200.7
	ED Payment	4.3	0.462	87.8	83.5
Utilization	Any Inpatient (%)	-1.5*	0.076	13.6	15.1
	Inpatient Days	-0.2*	0.074	1.0	1.2
	Any Outpatient (%)	-0.4	0.646	72.8	73.3
	Outpatient Days	-0.1	0.585	3.9	4.0
	Carrier Days	-0.5*	0.099	16.4	16.8
	ED Visits	0.0	0.399	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	1.3	0.190	23.7	22.4
	Breast Cancer (%)	-1.0	0.484	59.9	60.9
	Cardiovascular (%)	2.2*	0.083	66.7	64.5

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.16: Vendor B Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1s and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	653.1*	0.082	6,938.0	6,284.9
	Inpatient Payment	322.3	0.210	2,667.1	2,344.8
	Outpatient Payment	82.8	0.310	1,212.0	1,129.2
	Carrier Payment	144.3	0.126	2,249.8	2,105.5
	ED Payment	2.8	0.632	86.8	84.0
Utilization	Any Inpatient (%)	-0.6	0.446	14.2	14.8
	Inpatient Days	0.1	0.272	1.2	1.1
	Any Outpatient (%)	1.1	0.242	73.9	72.7
	Outpatient Days	-0.1	0.253	3.9	4.0
	Carrier Days	0.2	0.415	16.8	16.6
	ED Visits	0.0	0.624	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	-0.3	0.792	22.7	22.9
	Breast Cancer (%)	1.1	0.440	61.4	60.2
	Cardiovascular (%)	0.3	0.786	65.5	65.1

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.17: National Vendor A Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-694.9	0.149	5,963.2	6,658.0
	Inpatient Payment	-620.2*	0.065	1,980.1	2,600.4
	Outpatient Payment	-23.8	0.802	1,136.1	1,159.9
	Carrier Payment	-139.5	0.249	2,043.3	2,182.8
	ED Payment	1.1	0.881	85.7	84.5
Utilization	Any Inpatient (%)	-0.7	0.539	13.9	14.6
	Inpatient Days	-0.3*	0.074	0.9	1.2
	Any Outpatient (%)	1.8	0.149	74.2	72.4
	Outpatient Days	0.2	0.187	4.0	3.9
	Carrier Days	-0.4	0.280	16.4	16.8
	ED Visits	0.0	0.552	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	1.4	0.277	24.4	23.0
	Breast Cancer (%)	-2.7	0.139	58.4	61.1
	Cardiovascular (%)	2.2	0.182	67.4	65.3

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.18: National Vendor A Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 2	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-958.4**	0.047	5,794.4	6,752.8
	Inpatient Payment	-491.4	0.144	2,084.0	2,575.4
	Outpatient Payment	-280.3***	0.003	923.8	1,204.1
	Carrier Payment	-199.7*	0.099	2,029.5	2,229.2
	ED Payment	1.6	0.831	87.1	85.5
Utilization	Any Inpatient (%)	-2.1**	0.048	12.7	14.8
	Inpatient Days	-0.3*	0.095	0.9	1.2
	Any Outpatient (%)	-2.2*	0.083	70.9	73.0
	Outpatient Days	-0.2*	0.062	3.7	3.9
	Carrier Days	-0.6	0.106	16.2	16.8
	ED Visits	0.0	0.944	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	1.7	0.211	24.4	22.7
	Breast Cancer (%)	-0.5	0.800	60.8	61.2
	Cardiovascular (%)	0.4	0.790	66.0	65.6

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.19: National Vendor A Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-826.5**	0.048	5,906.7	6,733.2
	Inpatient Payment	-555.9*	0.056	2,060.0	2,615.9
	Outpatient Payment	-152.0*	0.065	1,049.4	1,201.4
	Carrier Payment	-169.5	0.106	2,037.8	2,207.4
	ED Payment	1.4	0.834	85.7	84.4
Utilization	Any Inpatient (%)	-1.4	0.134	13.6	15.0
	Inpatient Days	-0.3**	0.046	0.9	1.2
	Any Outpatient (%)	-0.2	0.869	72.4	72.5
	Outpatient Days	-0.0	0.752	3.9	3.9
	Carrier Days	-0.5	0.119	16.3	16.8
	ED Visits	0.0	0.701	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	1.5	0.177	24.2	22.6
	Breast Cancer (%)	-1.6	0.314	59.6	61.2
	Cardiovascular (%)	1.3	0.356	66.3	65.0

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.20: Local Vendor A Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-294.7	0.760	6,590.8	6,885.5
	Inpatient Payment	-641.4	0.288	2,014.9	2,656.4
	Outpatient Payment	159.1	0.569	1,396.9	1,237.8
	Carrier Payment	-21.0	0.931	2,143.1	2,164.1
	ED Payment	23.8	0.123	105.3	81.5
Utilization	Any Inpatient (%)	-2.3	0.292	12.6	15.0
	Inpatient Days	-0.1	0.746	1.0	1.1
	Any Outpatient (%)	-3.3	0.165	73.1	76.4
	Outpatient Days	-0.1	0.768	4.2	4.3
	Carrier Days	-0.4	0.530	16.5	17.0
	ED Visits	0.1	0.169	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	1.1	0.668	22.4	21.3
	Breast Cancer (%)	0.5	0.899	60.4	59.9
	Cardiovascular (%)	6.1*	0.068	69.3	63.2

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.21: Local Vendor A Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 2	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	40.5	0.966	6,708.3	6,667.9
	Inpatient Payment	-53.8	0.929	2,468.7	2,522.5
	Outpatient Payment	-18.2	0.948	1,185.9	1,204.1
	Carrier Payment	-8.5	0.972	2,166.2	2,174.8
	ED Payment	11.1	0.472	98.0	87.0
Utilization	Any Inpatient (%)	-1.4	0.531	13.3	14.7
	Inpatient Days	0.1	0.576	1.1	1.0
	Any Outpatient (%)	0.2	0.948	76.0	75.8
	Outpatient Days	-0.2	0.494	4.0	4.3
	Carrier Days	-0.3	0.707	16.6	16.9
	ED Visits	0.0	0.504	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	-0.2	0.947	21.6	21.8
	Breast Cancer (%)	3.0	0.411	62.5	59.5
	Cardiovascular (%)	6.0*	0.071	68.8	62.8

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.22: Local Vendor A Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	-125.6	0.880	6,591.4	6,717.0
	Inpatient Payment	-345.3	0.508	2,300.2	2,645.5
	Outpatient Payment	69.9	0.772	1,237.4	1,167.5
	Carrier Payment	-14.8	0.944	2,147.5	2,162.3
	ED Payment	17.4	0.193	98.1	80.7
Utilization	Any Inpatient (%)	-1.8	0.332	13.6	15.5
	Inpatient Days	0.0	0.889	1.1	1.0
	Any Outpatient (%)	-1.6	0.447	74.7	76.2
	Outpatient Days	-0.2	0.571	4.2	4.4
	Carrier Days	-0.3	0.562	16.7	17.1
	ED Visits	0.1	0.239	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	0.5	0.835	22.0	21.5
	Breast Cancer (%)	1.8	0.577	61.3	59.6
	Cardiovascular (%)	6.1**	0.036	68.4	62.3

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.23: National Vendor B Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	822.2*	0.092	7,225.9	6,403.7
	Inpatient Payment	383.6	0.260	2,815.7	2,432.1
	Outpatient Payment	40.6	0.673	1,189.7	1,149.1
	Carrier Payment	273.2**	0.026	2,386.8	2,113.6
	ED Payment	5.7	0.455	89.4	83.8
Utilization	Any Inpatient (%)	-0.5	0.673	14.1	14.6
	Inpatient Days	0.2	0.322	1.3	1.1
	Any Outpatient (%)	1.0	0.421	73.6	72.5
	Outpatient Days	-0.0	0.858	3.9	3.9
	Carrier Days	0.5	0.209	17.1	16.6
	ED Visits	0.0	0.279	0.4	0.3
Preventive Screening	Colorectal Cancer (%)	-1.0	0.446	22.4	23.4
	Breast Cancer (%)	3.0	0.102	63.2	60.1
	Cardiovascular (%)	1.0	0.539	66.5	65.5

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.24: National Vendor B Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 2	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	518.8	0.288	7,023.4	6,504.6
	Inpatient Payment	404.7	0.235	2,829.6	2,424.9
	Outpatient Payment	34.0	0.724	1,185.3	1,151.3
	Carrier Payment	54.3	0.658	2,240.8	2,186.5
	ED Payment	-0.5	0.944	85.3	85.8
Utilization	Any Inpatient (%)	-0.2	0.885	14.3	14.5
	Inpatient Days	0.2	0.286	1.3	1.1
	Any Outpatient (%)	1.4	0.260	73.8	72.4
	Outpatient Days	-0.1	0.450	3.8	3.9
	Carrier Days	0.4	0.325	17.0	16.7
	ED Visits	0.0	0.812	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.6	0.676	23.4	22.9
	Breast Cancer (%)	0.0	0.987	61.2	61.2
	Cardiovascular (%)	1.0	0.552	66.5	65.5

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.25: National Vendor B Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	670.5	0.113	6,901.2	6,230.7
	Inpatient Payment	394.2	0.181	2,691.2	2,297.0
	Outpatient Payment	37.3	0.655	1,175.2	1,137.9
	Carrier Payment	163.7	0.123	2,259.2	2,095.5
	ED Payment	2.6	0.696	86.5	84.0
Utilization	Any Inpatient (%)	-0.3	0.743	14.3	14.6
	Inpatient Days	0.2	0.235	1.2	1.1
	Any Outpatient (%)	1.2	0.265	73.3	72.1
	Outpatient Days	-0.1	0.590	3.9	3.9
	Carrier Days	0.4	0.196	16.9	16.5
	ED Visits	0.0	0.446	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	-0.2	0.843	23.0	23.2
	Breast Cancer (%)	1.6	0.330	61.7	60.1
	Cardiovascular (%)	1.0	0.486	66.1	65.2

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.26: Local Vendor B Arm 1 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 1	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	1.2	0.999	6,837.2	6,835.9
	Inpatient Payment	-90.7	0.878	2,473.3	2,564.1
	Outpatient Payment	57.6	0.833	1,312.4	1,254.8
	Carrier Payment	97.1	0.682	2,241.4	2,144.3
	ED Payment	13.9	0.357	97.0	83.1
Utilization	Any Inpatient (%)	-2.3	0.279	12.6	15.0
	Inpatient Days	-0.2	0.505	0.9	1.1
	Any Outpatient (%)	0.9	0.700	76.5	75.7
	Outpatient Days	-0.5	0.146	3.9	4.3
	Carrier Days	-0.8	0.225	16.2	17.0
	ED Visits	-0.0	0.955	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	0.5	0.859	21.9	21.4
	Breast Cancer (%)	-0.7	0.844	59.4	60.1
	Cardiovascular (%)	-3.3	0.315	61.5	64.8

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.27: Local Vendor B Arm 2 versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arm 2	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	976.1	0.301	7,486.8	6,510.7
	Inpatient Payment	122.8	0.835	2,615.7	2,492.8
	Outpatient Payment	440.5	0.106	1,567.5	1,127.1
	Carrier Payment	19.8	0.933	2,189.8	2,170.0
	ED Payment	-6.6	0.664	83.4	89.9
Utilization	Any Inpatient (%)	-1.7	0.440	13.1	14.7
	Inpatient Days	0.2	0.523	1.2	1.0
	Any Outpatient (%)	0.5	0.830	76.3	75.8
	Outpatient Days	-0.3	0.375	4.0	4.3
	Carrier Days	-0.4	0.593	16.5	16.9
	ED Visits	-0.0	0.476	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	-1.1	0.669	20.8	21.9
	Breast Cancer (%)	-0.9	0.803	59.3	60.2
	Cardiovascular (%)	-0.7	0.824	63.2	63.9

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

Exhibit C.28: Local Vendor B Arms 1 and 2 Pooled versus Arm 3 Regression Results

Outcome Type	Outcome	Coefficient[a]	P-value	Regression-Adjusted Mean Arms 1 and 2 Pooled	Regression-Adjusted Mean Arm 3
Medicare Payment	Total Medicare Payments	489.5	0.549	7,000.1	6,510.5
	Inpatient Payment	15.7	0.975	2,540.0	2,524.3
	Outpatient Payment	249.8	0.290	1,357.0	1,107.1
	Carrier Payment	58.3	0.776	2,196.1	2,137.8
	ED Payment	3.7	0.779	89.0	85.3
Utilization	Any Inpatient (%)	-2.0	0.284	13.5	15.5
	Inpatient Days	-0.0	0.987	1.0	1.0
	Any Outpatient (%)	0.7	0.731	76.2	75.5
	Outpatient Days	-0.4	0.177	4.1	4.4
	Carrier Days	-0.6	0.313	16.6	17.1
	ED Visits	-0.0	0.657	0.3	0.3
Preventive Screening	Colorectal Cancer (%)	-0.3	0.886	21.5	21.8
	Breast Cancer (%)	-0.8	0.797	59.6	60.4
	Cardiovascular (%)	-2.1	0.474	63.0	65.1

[a] *, **, *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.