Wellness Prospective Evaluation
Final Report

January 2019 (Revised)

Acumen, LLC

Kristy Piccinini
Michael Wernecke
Lucy Yao
Noy Birger
Anzer Habibulla
Yue Jin
Krishan Kumar
Stephen McKean
Katherine Pierce
Dimitra Politi
Jason Wyman

Westat, Inc.

Lois Olinger
Lauren Mercincavage
Jennifer Nooney
Zhiqun Tang
KC Lee
Hyunshik Lee

Submitted to:
Erin Colligan, PhD
Center for Medicare and Medicaid Innovation
Centers for Medicare & Medicaid Services (CMS)
7500 Security Blvd.
Mail Stop: WB-06-05
Baltimore, MD 21244

Contract Number: HHSM-500-2011-00012I; HHSM-500-T0011

Acumen, LLC
500 Airport Blvd., Suite 365
Burlingame, CA 94010

EXECUTIVE SUMMARY

In the Affordable Care Act (ACA), Section 4202, subsection (b), Congress mandated that the Centers for Medicare & Medicaid Services (CMS) conduct an independent evaluation of wellness programs focusing on the following priority areas:

- Chronic disease management (CDM);
- Physical activity, nutrition, and obesity (PANO);
- Falls prevention (FP); and
- Mental health

CMS contracted with Acumen, LLC, and its partner, Westat, Inc., ("the Acumen team") to conduct a prospective evaluation of evidence-based wellness programs. The Acumen team identified six national evidence-based programs with a primary focus on CDM, PANO, and FP, listed in Executive Summary Table 1.[1]

Executive Summary Table 1: Evaluated Wellness Programs

Chronic Disease Management	Physical Activity, Nutrition, and Obesity	Falls Prevention
Chronic Disease Self-Management Program (CDSMP) & Tomando Control de su Salud (Spanish-speaking CDSMP)	EnhanceFitness	A Matter of Balance
Diabetes Self-Management Program (DSMP) & Programa de Manejo Personal de la Diabetes (Spanish-speaking DSMP)	Fit & Strong!	Stepping On

The Wellness Prospective Evaluation assesses the impact of these wellness programs on the health, disease self-management behavior, functional status, health-related quality of life, health service utilization and Medicare costs to of Medicare fee-for-service (FFS) beneficiaries. The goal of the study is to determine whether broader Medicare beneficiary participation or Medicare coverage of wellness programs would be beneficial.[2]

Overall, key evaluation findings can be summarized as follows:

- The pattern of effects was generally consistent with the focus areas and design of the programs. For example, PANO programs improved participants' self-reported strength

[1] The Acumen team did not identify any suitable programs primarily focused on mental health, although some programs treated mental health as a secondary focus.

[2] Medicare costs analyzed in this report refer to Parts A and B and pharmaceutical (Part D) expenditures, and do not take into account the cost of administering wellness programs. For a qualitative study of program operations and costs, please see: "Report to Congress: The Centers for Medicare & Medicaid Services' Evaluation of Community-based Wellness and Prevention Programs under Section 4202(b) of the Affordable Care Act," found here: https://innovation.cms.gov/Files/reports/CommunityWellnessRTC.pdf

and physical activity levels, and FP programs improved multiple measures of physical health and body strength.

- PANO and FP programs had consistently positive impacts on self-reported mental health.
- Confidence in balance improved among program participants in all three priority areas.
- Outpatient emergency room (ER) expenditures decreased among PANO participants and home health expenditures decreased among FP participants. Emergency room utilization actually increased among CDM program participants. There is no evidence of program effects on healthcare utilization among PANO or FP participants, or on expenditures among CDM participants.

Executive Summary Table 2 through Executive Summary Table 4 show select quantitative findings of the evaluation, which used a differences-in-differences (DiD) estimation methodology. Findings are presented for CDM, PANO, and FP programs, respectively. Adjusted means are reported for both the pre-intervention and post-intervention periods. In addition, estimates of the effect of wellness programs are shown for both the first and second six-month period following program participation ("interim estimates"), as well as for the entire post-intervention 12-month period ("cumulative estimates").[3]

Executive Summary Table 2: Key Evaluation Findings for CDM Wellness Programs

Outcome	Adjusted Means				Interim DiDs		Cumulative DiD		
	Pre-Intervention		Post-Intervention		0-6 months	7-12 months	DiD (90% CI)	p-value	% Relative Diff.
	Part.	Comp.	Part.	Comp.	DiD (90% CI)	DiD (90% CI)			
Mental Health									
Role Emotional Subscale	44.1	45.5	44.3	45.2	1.41** (0.4,2.5)	-0.92 (-2.1,0.2)	0.52 (-0.6,1.6)	0.43	1.2%
Mental Health Subscale	50.6	52.0	50.3	51.6	1.06** (0.3,1.9)	-0.81 (-1.7,0.1)	0.26 (-0.6,1.1)	0.63	0.5%
Balance									
Confidence in Balance	51.4	59.3	52.0	55.4	3.33** (0.7,5.9)	1.03 (-1.3,3.3)	4.53*** (1.8,7.3)	0.01	8.8%
Emergency Room (ER) Visits per 1,000 Beneficiaries									
Count of ER Visits	289.8	275.1	391.1	226.3	41.84 (-37.2,120.9)	108.70** (33, 184.4)	149.97* (21.5,278.4)	0.06	51.8%

[3] Means are adjusted for covariates (gender, age, race, income, education, urban/rural indicators, and dual Medicare eligibility indicators) included in the models. The analytic sample consists of new program enrollees surveyed over a 15-month period in 2014 and 2015, and respondents to a national survey fielded in 2015, who were matched to program enrollees based on demographic, clinical, and self-reported information ("comparison group"). Both program participants and the comparison group were surveyed six and twelve months after initial survey waves. For more information regarding the identification of comparison groups for this evaluation, see: "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

Notes: Part.: Wellness program participants. Comp.: Comparison group. DiD: Differences-in-Differences; CI: Confidence Interval; The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate; p-value: probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data; *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01; % Relative Diff: Relative difference, calculated as the cumulative DiD estimate divided by the baseline participant adjusted mean, and expressed as a percentage.

Executive Summary Table 3: Key Evaluation Findings for PANO Wellness Programs

Outcome	Adjusted Means				Interim DiDs		Cumulative DiD		
	Pre-Intervention		Post-Intervention		0-6 months	7-12 months	DiD (90% CI)	p-value	% Relative Diff.
	Part.	Comp.	Part.	Comp.	DiD (90% CI)	DiD (90% CI)			
Physical Health									
Role Physical Subscale	45.3	45.2	45.6	45.2	0.93** (0.2,1.6)	-0.53 (-1.3,0.2)	0.25 (-0.5,1.0)	0.59	0.6%
Mental Health									
Mental Components Summary Score	53.0	52.9	53.5	52.4	0.81* (0.0,1.6)	0.31 (-0.5,1.1)	1.03** (0.2,1.9)	0.04	1.9%
Vitality Subscale	51.9	52.2	51.6	51.2	0.41 (-0.3,1.1)	0.33 (-0.3,1.0)	0.73* (0.0,1.4)	0.08	1.4%
Social Functioning Subscale	50.0	49.9	50.0	49.2	0.85* (0.1,1.7)	-0.13 (-0.9,0.7)	0.74 (-0.1,1.6)	0.14	1.5%
Role Emotional Subscale	47.3	47.4	48.3	47.4	1.09* (0.1,2.1)	-0.01 (-0.9,0.8)	1.12* (0.1,2.1)	0.07	2.4%
Mental Health Subscale	53.0	52.9	53.5	52.4	0.45 (-0.3,1.1)	0.51 (-0.2,1.2)	0.96** (0.2,1.8)	0.05	1.8%
Physical Strength and Balance									
Aerobic Activity	5.1	5.0	5.1	4.8	0.19* (0.0,0.3)	0.03 (-0.1,0.2)	0.23** (0.1,0.4)	0.03	4.5%
Strength and Flexibility	0.7	0.7	0.8	0.6	0.15*** (0.1,0.2)	-0.03 (-0.1,0.0)	0.14*** (0.1,0.2)	0.00	20.6%
Confidence in Balance	63.5	67.3	63.1	64.3	0.90 (-1.4,3.3)	1.15 (-0.7,3.0)	2.56* (0.3,4.8)	0.06	4.0%
Outpatient ER Expenditures per Beneficiary									
Total Outpatient ER	$363.15	$313.07	$307.85	$383.75	-55.94 (-143.1, 31.2)	-70.12** (-125.4, -14.9)	-125.98* (-248.1, -3.9)	0.09	-34.7%

Notes: Part.: Wellness program participants. Comp.: Comparison group. DiD: Differences-in-Differences; CI: Confidence Interval; The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate; p-value: probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data; *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01; % Relative Diff: Relative difference, calculated as the cumulative DiD estimate divided by the baseline participant adjusted mean, and expressed as a percentage.

Executive Summary Table 4: Key Evaluation Findings for FP Wellness Programs

Outcome	Adjusted Means				Interim DiDs		Cumulative DiD		
	Pre-Intervention		Post-Intervention		0-6 months	7-12 months	DiD (90% CI)	p-value	% Relative Diff.
	Part.	Comp.	Part.	Comp.	DiD (90% CI)	DiD (90% CI)			
Physical Health									
Role Physical Subscale	42.1	43.0	42.0	42.3	0.75** (0.2,1.3)	-0.11 (-0.6,0.4)	0.64* (0.1,1.2)	0.05	1.5%
Bodily Pain Subscale	45.1	46.1	45.2	45.6	0.54* (0.0,1.1)	0.15 (-0.4,0.7)	0.59* (0.1,1.1)	0.05	1.3%
Mental Health									
Mental Components Summary Score	51.7	51.8	52.1	51.4	0.94*** (0.3,1.5)	-0.28 (-0.8,0.3)	0.81** (0.3,1.4)	0.02	1.6%

Outcome	Adjusted Means				Interim DiDs		Cumulative DiD		
	Pre-Intervention		Post-Intervention		0-6 months	7-12 months	DiD (90% CI)	p-value	% Relative Diff.
	Part.	Comp.	Part.	Comp.	DiD (90% CI)	DiD (90% CI)			
Role Emotional Subscale	45.1	45.9	45.6	45.1	1.51*** (0.8,2.2)	-0.28 (-1.0,0.4)	1.22*** (0.5,1.9)	0.01	2.7%
Mental Health Subscale	51.7	51.8	51.9	51.5	0.74** (0.2,1.3)	-0.21 (-0.7,0.3)	0.56* (0.0,1.1)	0.09	1.1%
Physical Strength and Balance									
Aerobic Activity	4.7	4.6	4.4	4.5	-0.19** (-0.3,-0.1)	0.07 (-0.0,0.2)	-0.12 (-0.2,0.0)	0.11	-2.5%
Strength and Flexibility	0.6	0.5	0.6	0.5	0.04* (0.0,0.1)	0.01 (-0.0,0.0)	0.05** (0.0,0.1)	0.02	8.8%
Confidence in Balance	50.7	56.0	51.5	52.7	2.66*** (1.0,4.3)	1.00 (-0.4,2.4)	4.12*** (2.5,5.7)	0.00	8.1%

Notes: Part.: Wellness program participants. Comp.: Comparison group. DiD: Differences-in-Differences; CI: Confidence Interval; The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate; p-value: probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data; *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01; % Relative Diff: Relative difference, calculated as the cumulative DiD estimate divided by the baseline participant adjusted mean, and expressed as a percentage.

The key driver of improvements in self-reported measures of physical and mental health appears to be the relative stability of many outcome measures over time among participants, compared to the decline in outcome measures within the comparison group.[4] These findings indicate that wellness programs may have protective effects against deterioration in health and activity that naturally occur with aging, as opposed to generating notable improvements in self-reported health and activity for participants.

The small, statistically significant positive effects of the PANO and FP programs on mental health suggest that enhanced mental well-being may be an important secondary benefit of participation, resulting from both lifestyle changes (e.g. increased physical activity) and knowledge gained from programs, or from the social act of program participation.

The unexpected, significant increase in emergency room (ER) visits among CDM participants may be related to the demographics of this population. Specifically, CDM program participants (and their matched comparison group) have comparatively lower income and education levels than program participants in FP and PANO programs. Low socio-economic status, even after controlling for access to health insurance, is associated with a preference for utilizing ER services for primary care needs.[5] It is possible, therefore, than an unintended consequence of CDM programs, which encourage regular interactions with physicians for the

[4] Notable exceptions include statistically significant participant gains in strength and flexibility activities for FP and PANO programs.

[5] See, for example: Kangovi, S., F. K. Barg, T. Carter, J. A. Long, R. Shannon, and D. Grande. "Understanding Why Patients Of Low Socioeconomic Status Prefer Hospitals Over Ambulatory Care." *Health Affairs* 32, no. 7 (July 2013): 1196-203, doi:10.1377/hlthaff.2012.0825.

management of chronic conditions, is increased ER utilization to cover primary care needs. These findings are consistent with other studies showing that when increases in healthcare utilization occur among beneficiaries with socio-economic characteristics similar to those of CDM participants, they affect multiple settings, including the ER.[6] Furthermore, it is possible that increases in primary care utilization lead to increased ER utilization, if primary care providers advise patients to go to the ER for their more urgent medical needs.

Decreases in outpatient ER expenditures among PANO program participants and in home health expenditures among FP participants are consistent with a lower intensity of healthcare utilization. However, all findings should be interpreted with caution due to small sample sizes.

The analysis presented in this report is subject to a number of limitations. First, the observational nature of this study implies that estimated effects may be biased due to unobserved differences between the treatment and the comparison groups. While the analysis improves upon many other observational studies of wellness programs by explicitly taking self-selected program participation into account during the identification of comparison groups, it remains possible that the approach did not fully account for self-selection effects.[7] Second, we were unable to assess impacts of individual programs due to low enrollment numbers for program participants, and subsequently small sample sizes. Instead, we pooled participants of the two wellness programs within each priority area into a single sample to ensure that we had adequately powered analyses. As a result, the analysis cannot make conclusions about the effectiveness of any individual program. Third, the analyses of self-reported outcomes focused on the subsample of beneficiaries who responded to the 12-month follow-up survey and may thus be subject to response bias. Weighting methods were used to control for survey non-response, and the results (presented in this report) were very similar to the unweighted results, indicating that response bias is minimal. Fourth, claims-based analyses (expenditures, utilization, incidences of falls/fractures, adherence) focus on Medicare FFS beneficiaries,[8] and suffer from a number of limitations related to the small sample sizes included in the analysis. Statistical power is low and, for many claims-based outcomes, there is only a small number of participant and comparison group beneficiaries with nonzero observations driving the statistical estimates. The analysis of Part D claims (for

[6] See, for example, Finkelstein, A. N., S.L. Taubman, H.L. Allen, B.J. Wright, and K. Baicker. "Effect of Medicaid Coverage on ED Use - Further Evidence from Oregon's Experiment." *The New England Journal of Medicine* 375, no. 16 (October 2016): 1505-1507.

[7] For more information regarding the identification of comparison groups for this evaluation, see: "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[8] Beneficiaries enrolled in FFS cannot be combined with beneficiaries enrolled in Medicare Advantage (MA) for claims-based analyses, because the data sources and the way information is reported differ across the two cohorts. Beneficiaries in enrolled in MA could not be analyzed in this evaluation, due to the small sample sizes and number of beneficiaries with nonzero observations.

medication adherence measures) suffers from even lower sample sizes, given that Part D is optional and there are additional restrictions required for adherence calculations. Low statistical power makes it hard to detect an effect, especially if its size is small, and might explain why there is little evidence of impact of wellness programs on utilization, expenditures, and medication adherence outcomes in this evaluation.

Given these limitations, the lack of consistently positive findings of the claims-based analyses is not at odds with the positive findings of the survey-based analyses (self-reported physical and mental health, physical activity and strength, confidence in balance). Survey-based outcomes, particularly those related to self-reported mental health and wellbeing, measure concepts that often differ from those embodied in claims-based metrics. Therefore, a change reflected in self-reported outcomes will not necessarily be detected in the analysis of Medicare claims. In addition, the lack of consistent findings of the claims-based analyses may be due to the short post-intervention observation period. The improvement in self-reported health may have more sustainable downstream effects on medical costs and healthcare utilization over a longer post-intervention observation period.

This report differs from prior studies along four dimensions: (i) research setting and research design; (ii) study population; (iii) source of the data analyzed; and (iv) duration of the follow-up period. Differences in these four dimensions may explain differences in findings between this evaluation and prior studies. Unlike previous studies,[9] this evaluation is based on an observational, "real-word" setting that takes selection into wellness programs into account. In addition, the claims-based analysis on utilization and expenditure outcomes focuses on beneficiaries enrolled in Medicare FFS whereas many prior studies focus on managed care populations that may have different demographic and health characteristics.[10] Finally, this analysis relied on administrative and self-reported data and observed outcomes over a one-year follow-up period, while other studies utilized different data sources and post-intervention periods.

In sum, although there is limited evidence of cost savings in this evaluation, the observed protective effects of wellness programs, particularly those focused on PANO and FP, on physical and mental health, physical activity, body strength, and confidence in balance may pay dividends

[9] See, for example: Brady, Teresa J., et al. "A Meta-Analysis of Health Status, Health Behaviors, and Health Care Utilization Outcomes of the Chronic Disease Self-Management Program." *Preventing Chronic Disease* 10 (January 2013); and Alva, Maria L., et al. "Impact of The YMCA of the USA Diabetes Prevention Program on Medicare Spending and Utilization." *Health Affairs* 36, no. 3 (March 2017): 417-424.

[10] See, for example: Lorig, Kate R., et al. "Effect of a Self-Management Program on Patients with Chronic Disease." *Effective Clinical Practice* 4, no. 6 (November-December 2001): 256-262.; and Ackermann, Ronald T., et al. "Healthcare Cost Differences with Participation in a Community-Based Group Physical Activity Benefit for Medicare Managed Care Health Plan Members." *Journal of the American Geriatrics Society* 56, no. 8 (August 2008): 1459-1465.

in the future. Further studies are needed to explore whether a longer follow-up period or a larger sample size yield more promising effects on expenditure and utilization outcomes.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

1 INTRODUCTION

Community-based wellness and chronic disease prevention programs ("wellness programs") aim to promote healthier lifestyles, lower beneficiary health risks, and ultimately improve health outcomes. Wellness programs have the potential both to improve the health of Medicare beneficiaries and to reduce spending in the Medicare program.

In the Affordable Care Act (ACA), Section 4202, subsection (b), Congress mandated that the Centers for Medicare & Medicaid Services (CMS) conduct an independent evaluation of wellness programs focusing on the following four priority areas:

(i) Chronic disease management (CDM);

(ii) Physical activity, nutrition, and obesity (PANO);

(iii) Falls prevention (FP); and

(iv) Mental health.

CMS contracted with Acumen, LLC, and its partner, Westat, Inc., ("the Acumen team") to conduct a prospective evaluation of evidence-based wellness programs. The Acumen team identified six national evidence-based programs with a primary focus on CDM, PANO, and FP, described in Table 1.1.[11] The Acumen team did not identify any evidence-based programs primarily focused on mental health that met the inclusion criteria, although some programs treated mental health as a secondary focus.

Table 1.1: Overview of Wellness Programs Included in the Prospective Evaluation

Wellness Program	Description	Duration
Chronic Disease Management		
Chronic Disease Self-Management Program (CDSMP)	Group class for individuals with one or more chronic conditions, and their caregivers or significant others, focusing on: • Chronic disease risk and symptoms • Diet and exercise • Medication use • Communication with health care providers	6 weeks One 2.5-hour class per week
Diabetes Self-Management Program (DSMP)	Group class for individuals with diabetes, and their caregivers or significant others. This program is similar to CDSMP, but focuses only on diabetes.	6 weeks One 2.5-hour class per week
Physical Activity, Nutrition, and Obesity		
Enhance Fitness	Group exercise class for older adults focusing on: • Stretching • Cardiovascular endurance • Strength training • Balance and flexibility	Ongoing classes Three 1-hour classes per week

[11] Detailed descriptions of each national evidence-based program are available in "Wellness Prospective Evaluation Report on Baseline Survey Efforts and Qualitative Study of Program Operations and Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. March 2016. Available at: https://innovation.cms.gov/Files/reports/communitywellnessprgms-frstevalrpt.pdf

Wellness Program	Description	Duration
Fit & Strong!	Group exercise class targeted at sedentary and de-conditioned adults with lower extremity mobility challenges, focusing on: • Health education • Goal-setting • Problem solving • Stretching and balance • Low-impact aerobics • Strength training	8 weeks Three 1.5-hour classes per week
Falls Prevention		
A Matter of Balance	Group class for older adults to: • Reduce fear of falling • Set realistic goals for increasing activity • Change the environment to reduce falls risk factors	8 weeks One 2-hour class per week
Stepping On	Group class for older adults to understand their risk of falls, coping behaviors, and safety strategies in everyday life, including: • Falls history and future risk • Home hazards • Safe footwear and clothing • Vision as it relates to falls • Community mobility • Medication risks • Strength and balance exercises	7 weeks One 2-hour class per week; plus one booster session 3 months post-program

The Wellness Prospective Evaluation aims to assess the impact of the wellness programs presented in Table 1.1 on Medicare beneficiary health, utilization, and costs to determine whether broader Medicare beneficiary participation or Medicare coverage of wellness programs would be beneficial. Specifically, this Final Report addresses the following research questions:

- *Research Question 1:* What was the effect of participation in a wellness program on key self-reported physical and mental health outcomes, disease self-management behavior, functional status, and health-related quality of life?
- *Research Question 2:* Did participation in wellness programs lead to reductions in key health service utilization and expenditure outcomes?

This Final Report presents findings from baseline, six-month, and twelve-month follow-up survey and claims-based analyses. Section 2 describes analytic methods used in this report. Section 3 summarizes results on the effects of participation in wellness programs on self-reported physical and mental health, disease self-management behavior (e.g., self-reported adherence to medications), functional status (e.g., levels of physical activity and body strength), and health-related quality of life (e.g., confidence in balance). Section 4 presents findings on the effects of participation in wellness programs on health service utilization and costs, as well as medication adherence (using information from Part D claims) and incidence of falls and fractures (using information from Parts A and B claims). Section 5 draws global conclusions, synthesizing findings from the current and previous reports. Additional methodological details and results are available in the appendices.

2 METHODOLOGY

This section provides a brief summary of the methodology employed for the analytic sample construction and the empirical analyses presented in this evaluation. Detailed findings on the effect of wellness programs on self-reported health and health behaviors are presented in Section 3. Findings on the effect of wellness programs on healthcare utilization, expenditures, and medication adherence are presented in Section 4. This section is organized as follows: Section 2.1 describes the process of selecting the program participants and comparison groups that comprise the study samples. Sections 2.2 and 2.3 outline the methodology used to analyze the quantitative data. Sections 2.4 and 2.5 discuss methodological considerations specific to the survey-and claims-based analyses.

2.1 Selection of Program Participants and Comparison Group

To collect a sample of wellness program participants, the evaluation team partnered with 75 organizations offering wellness programs, and conducted baseline surveys of new program enrollees from October 2014 to December 2015. New program enrollees were eligible to participate in the baseline survey if they were enrolled in Medicare and 66 years of age and older. Table 2.1 shows the survey data collection design for the program participant sample. As shown in Table 2.1, the baseline survey was provided on-site to eligible new program participants.[12] The six-month survey was fielded by mail to the baseline survey respondents, and the twelve-month survey was fielded by mail to the six-month survey respondents. Details of the sampling, fielding, and weighting of the baseline and six-month surveys can be found in the "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs."[13] That report also contains the baseline and six-month survey instruments, while the twelve-month instruments are shown in Appendix E of this report.

The comparison group originated from a national sample of non-institutionalized Medicare beneficiaries.[14] These respondents completed surveys similar in content to those completed by wellness program participants, with additional questions on beneficiaries' readiness to participate in wellness programs and make lifestyle changes. Baseline surveys for the national sample were fielded from January to December 2015, roughly coinciding with

[12] Completed surveys of individuals who did not meet the eligibility criteria were excluded from the study, and those individuals did not receive follow-up surveys.

[13] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[14] Women with diabetes were oversampled, because they are disproportionately represented among program participants, and oversampling improved Acumen's ability to identify comparison groups for evaluation purposes.

survey fielding dates for wellness program participants. National survey respondents were also surveyed at six and twelve months after their first survey fielding date.[15] Table 2.1 also shows the survey data collection design for the national sample. Similarly to program participants, only national survey respondents who completed the baseline survey received a six-month follow-up. The twelve-month survey was mailed to those who both completed the six-month survey and also met criteria for matching to the participant sample (for details on matching criteria and the timing of twelve-month surveys, see the "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs" and Appendix A.5).[16]

Table 2.1: Medicare Beneficiary-Level Primary Data Collection Design

Survey	Wellness Program Participant Sample	National Sample
Baseline	Administered onsite at enrollment to new program participants (deemed eligible to participate in the survey) over a 15-month enrollment period Wellness program attendance records are also collected for those participants for whom baseline surveys have been received.	12 waves at 1-month intervals, by mail
6-Month Follow-Up Survey	Administered at corresponding 6-month points, by mail, to all program participants who were eligible for survey participation and who completed the baseline survey	12 waves at corresponding 1-month intervals, by mail, to all beneficiaries who completed the baseline survey
12-Month Follow-Up Survey	Administered at corresponding 12-month point, by mail, to all program participants who completed the 6-month follow-up survey	12 waves at corresponding 1-month intervals, by mail, to beneficiaries who met criteria for matching, and who completed the 6-month follow-up survey

Medicare enrollment and claims information was extracted for both program participants and the national sample members who completed the baseline survey. Using both self-reported information and claims data, program participants were matched to national survey respondents, and 1:1 propensity score matching was performed separately by ACA priority area and by Medicare enrollment category (fee for service or Medicare Advantage). To mitigate selection bias, only those national survey respondents with high self-reported readiness to participate in a wellness program and/or make lifestyle changes were considered for matching.[17] Propensity score matching ensured covariate balance on a variety of important predictive characteristics

[15] Some twelve-month surveys for matched national survey respondents were fielded later, in early 2017, after matching criteria were finalized.

[16] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[17] This approach may not fully account for selection into participation, if self-reported readiness to participate in wellness programs and/or make lifestyle changes is measured with error or is an unreliable predictor of program participation.

(e.g., medical conditions, baseline healthcare utilization and expenditures) while also ensuring exact matches on particularly important beneficiary characteristics (e.g., race, age, sex). The matching algorithm identified a well-balanced baseline sample of participant and comparison beneficiaries. Depending on the program, 23 to 36 percent of participants could not be appropriately matched to a national survey respondent, and thus were excluded from the analysis.[18]

2.2 Differences-in-Differences Methodology

A differences-in-differences (DiD) design was employed for the quantitative analyses of self-reported and claims-based outcomes. DiD estimation compares the change in the average for an outcome of interest among program participants to the change in the average for the same outcome among the comparison group, each measured relative to a pre-intervention baseline time period. The DiD estimator automatically controls for differences in characteristics that remain constant over time, on average, between program participants and the comparison group. The effect of wellness programs is identified by looking at differences in the trend of an outcome over the observation period. For double robustness, DiD models also control for urban/rural status, dual eligibility status, gender, race, age, education, and income.[19] More details about the DiD model can be found in Appendix A.3.

The effect of participation in wellness programs was estimated separately by ACA priority area. A program-specific analysis (within each priority area) was not feasible due to small enrollment numbers at the program level, which affected statistical power. As a result, the analysis cannot make conclusions about the relative effectiveness of any individual wellness programs. The claims-based analysis focuses on beneficiaries enrolled in fee for service (FFS). Beneficiaries enrolled in FFS cannot be combined with beneficiaries enrolled in Medicare Advantage (MA) for claims-based analyses, because the data sources and the way information is reported differ across the two cohorts. We attempted to study the impacts of the programs on beneficiaries in enrolled in MA, but were unable to do so due to the small sample sizes and number of beneficiaries with nonzero observations. Heteroscedasticity-robust standard errors are reported for all analyses.

[18] More information on the matching methodology, and pre- and post-matching covariate summaries, see "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[19] All survey-based and most claims-based models have been estimated with and without covariates, and DiD estimates are very similar across the two model specifications. The incidence of falls and fractures could not be reliably estimated with covariates due to small sample sizes and low numbers of beneficiaries with nonzero observations.

Three sets of DiD models were estimated for each self-reported outcome, corresponding to comparisons across the following two points in time: (1) baseline and six-month follow-up surveys, (2) six-month and twelve-month follow-up surveys, and (3) baseline and twelve-month follow-up surveys. The outcomes of interest explored in the survey-based analysis are discussed in Section 2.4.1.

The claims-based analyses also include three sets of DiD models, but they produce comparisons across time periods, rather than across single points in time. Specifically, one set of DiD estimates compares the year prior to participation in wellness programs or response to the national survey ("baseline period") to the first six months post intervention ("Interim Estimates: 0-6 months"). Another set of estimates compares the baseline period to the second six months post intervention ("Interim Estimates: 7-12 months"). Finally, a third set of estimates compares the baseline period to the entire post-intervention year ("Cumulative Estimates"). The claims-based outcomes of interest are discussed in Section 2.5.2.

2.3 Intention-to-Treat and Treatment-on-the-Treated

Two types of analyses were performed: intention-to-treat (ITT) and average-treatment-on-the-treated (ATT). The ITT analysis answers the question "What is the effect of participating in wellness programs?" and is based on a sample of matched program participants, irrespective of whether they completed a wellness program. The ATT analysis, which is based only on those matched beneficiaries who completed the program, answers the question "What is the effect of completing a wellness program?" The ITT analysis is more appropriate for the main policy question of the whether wellness programs are a worthwhile service for the Medicare population. Also, the ITT analysis is free from potential bias in the case where healthier beneficiaries are more likely to complete wellness programs. In addition, the matching algorithm is based on beneficiaries' readiness to participate in, rather than completion of, wellness programs. For these reasons, the main body of this report focuses on ITT findings. ATT results are included in Appendix D. The results of the ATT analysis on both self-reported and claims-based outcomes are very similar to the results of the ITT analysis.

2.4 Self-Reported Health and Health Behaviors Analyses

This section discusses the measures and methodological issues particular to the analysis of self-reported health outcomes and behaviors. Section 2.4.1 presents the measures for the analysis, and Section 2.4.2 discusses survey response rates and weighting to account for survey non-response.

2.4.1 Self-reported Health, Wellbeing, and Behavior Measures Collected

Our survey measured a total of 15 self-reported outcomes, representing areas targeted by wellness programs for improvement. They included overall physical and mental health, physical

activity, falls and balance, and medication adherence. All measure specifications for survey-based outcomes are presented in Appendix A.2.

- **Overall physical and mental health:** The Short Form Health Survey 36v2 (SF-36[20]) was used to measure overall physical and mental health as a function of key subdomains related to roles and functioning. The subdomains include:
  - *Physical functioning* - a 10-item scale that assesses performance of physical activities such as self-care, walking, moderate physical activities, and vigorous physical activities.
  - *Bodily pain* – a 2-item scale that assesses intensity, duration, and frequency of bodily pain and limitations in usual activities due to pain.
  - *Role physical* - a 4-item scale that assesses the degree to which a person performs their typical role activities (e.g., work or other activities).
  - *General health* - a 5-item scale that assesses beliefs and evaluations of a person's overall health.
  - *Vitality* - a 4-item scale that assesses a person's feelings of energy and the absence of fatigue.
  - *Social functioning* - a 2-item scale that assesses the degree to which a person's health problems interfered with normal social activities.
  - *Role emotional* - a 3-item scale that assesses role limitations related to mental health.
  - *Mental health* - a 5-item scale that assesses a person's emotional, cognitive and intellectual status, such as the degree to which a person feels nervous, depressed, calm, peaceful, and happy.
- **Physical activity:** The Rapid Assessment of Physical Activity (RAPA[21]) aerobics and strength/flexibility scales measure the amount and intensity of the respondent's usual physical activities (RAPA 1); and the level of activities undertaken to increase muscle strength and flexibility (RAPA 2).
- **Falls and balance:** Respondents were asked to provide the number of times they had fallen in the past six months. They also completed a series of six items measuring beneficiary confidence in balance, known as the Activities-specific Balance Confidence (ABC-6) scale.[22,23] These items ask respondents to rate their confidence in remaining

[20] See https://campaign.optum.com/content/optum/en/optum-outcomes/what-we-do/health-surveys/sf-36v2-health-survey.html and QualityMetric Health Outcomes™ Scoring Software 4.5 User's Guide (2004, 2007, 2009, 2010, 2011) for more technical details.
[21] Topolski TD, LoGerfo J, Patrick DL, Williams B, Walwick J, Patrick MB. "The Rapid Assessment of Physical Activity (RAPA) Among Older Adults." *Preventing Chronic Disease* 3, no. 4 (October 2006): A118.
[22] Peretz C, Herman T, Hausdorff J, Giladi, N. (2006). "Assessing Fear of Falling: Can a Short Version of the Activities-Specific Balance Confidence Scale Be Useful?" *Movement Disorders* 21, no. 12 (September 2006): 2101–2105.
[23] Schepens S, Goldberg A, Wallace M. "The short version of the Activities-specific Balance Confidence (ABC) scale: its validity, reliability, and relationship to balance impairment and falls in older adults." *Archives of Gerontology and Geriatrics* 51, no. 1 (July-August 2010): 9-12.

steady for specific activities such as standing on their tiptoes and reaching for something above their heads or stepping onto and off of an escalator.

- **Medication adherence:** The Morisky-4 medication adherence scale[24] measures problems remembering to take medications and stopping medications when feeling better or worse. Percentages are based only on respondents who take medications.

2.4.2 Survey Response and Weighting

Not everyone in the participant and comparison samples responded to the six- and twelve-month surveys. Nonresponse may be due to survey refusal, death, and institutionalization. Table 2.2 shows the survey completion rate at twelve months for the matched samples in each ACA priority area. Between 57.5 percent and 66.5 percent of the matched samples completed surveys for all three time points and are included in the final analyses.

Table 2.2: Follow-Up Survey Respondents for the Matched Samples

Group	Starting Sample	Six-Month Survey Completes	Twelve-Month Survey Completes	Sample Completion Rate at Twelve Months**
CDM				
National Respondents	920	734	585	63.6%
Participant Respondents	920	641	529	57.5%
PANO				
National Respondents	1,046	850	693	66.3%
Participant Respondents	1,046	764	656	62.7%
FP				
National Respondents	2,013	1,628	1,339	66.5%
Participant Respondents	2,013	1,471	1,252	62.2%

** Completes/Starting sample

To reduce bias due to differential nonresponse at twelve months between participants and the comparison group, nonresponse adjustment weights were used. For weighting purposes at six and twelve months, each matched sample was treated as a census at baseline. Both the weighting and analytic strategies treat the matched samples as having independent national and participant components as opposed to sets of two matched individuals. This allows us to preserve sample size when only one individual in a matched pair responds.[25] More details about the weighting strategy for the twelve-month survey can be found in Appendix A.

[24] Morisky DE, Green LW, Levine DM. "Concurrent and Predictive Validity of a Self-Reported Measure of Medication Adherence." *Medical Care* 24, no. 1 (January 1986): 67-74.
[25] Schafer, J.L., and Kang, J. "Average causal effect from nonrandomized studies: A practical guide and simulated example." *Psychological Methods* 13, no. 4 (December 2008): 279-313.

2.5 Healthcare Utilization, Expenditure, and Medication Adherence Analyses

The analyses of claims-based utilization, expenditure, and adherence outcomes include matched program participants and national survey respondents, and use information from claims data covering a two-year period: the 12 months prior to the start of wellness program participation or response to the national survey, and the 12 months following. Because participation and survey response dates differ across beneficiaries, the calendar periods of observation also vary. Participation dates range from October 2014 to December 2015,[26] while survey response dates range from January 2015 to March 2016, so the full observation period ranges from October 2014 to March 2017, with each beneficiary observed over a two-year period.

This section discusses methodological topics specific to the claims-based analyses. Section 2.5.1 outlines the data sources used in the quantitative analysis. Section 2.5.2 presents the outcomes of interest. Section 2.5.3 describes the enrollment restrictions and study inclusion criteria.

2.5.1 Sources of Data and Price Standardization

The claims-based analyses relied on beneficiary enrollment information from Medicare's Enrollment Data Base (EDB) and Parts A and B claims data from the Common Working File (CWF). Part D Prescription Drug Event (PDE) data were also used for the analyses of pharmaceutical expenditures and medication adherence.

FFS expenditure data included in these analyses were standardized to control for regional differences in the cost of care (due to labor costs and practice expenses).[27] In addition, all expenditures are reported in 2011 US dollars.

2.5.2 Utilization, Expenditures, and Medication Adherence Outcomes

The claims-based analyses focused on outcomes measuring healthcare utilization, medical and pharmaceutical expenditures, and medication adherence. All measure specifications for claims-based outcomes are presented in Appendix A.1. Utilization outcomes include the number of inpatient (IP) admissions (all-cause and unplanned), length of stay, and number of emergency room (ER) visits. The incidence of falls and fractures, defined as the number of beneficiaries with at least one fall- or fracture-related claim, is also analyzed. Expenditure

[26] There are a few cases of program participants with program start dates in January or February 2016. These cases correspond to beneficiaries with missing or invalid initial program attendance dates. For these beneficiaries, the baseline survey response dates were used instead.

[27] More information about expenditure standardization methodology is available in CMS Standardization Methodology For Allowed Amount (CMS), available at https://www.qualitynet.org.

outcomes include total Parts A and B expenditures, total Part D expenditures (for the subsample of beneficiaries with Part D coverage), IP expenditures, outpatient ER and non-ER expenditures, physician and ancillary services expenditures, durable medical equipment expenditures, and home health expenditures.[28] Medication adherence was estimated for the following five drug classes, based on their importance for the management of chronic conditions and their high prevalence in the population:

- Beta blockers (for the management of hypertension and heart arrhythmias)
- Calcium channel blockers (for the management of hypertension and heart arrhythmias)
- Diabetes medications
- Renin angiotensin system (RAS) Antagonists (for the management of hypertension)
- Statin medications (anti-cholesterol medications for the management of cardiovascular disease)

We used the Pharmacy Quality Alliance (PQA) proportion of days covered (PDC) metric assessing the proportion of days with prescription coverage for the drug classes listed above; this metric has been endorsed by the National Quality Forum (NQF).[29] Medication adherence was measured in two ways: (1) change in average PDC, and (2) change in the proportion of beneficiaries who are highly adherent. High adherence is defined as having a PDC of at least 80 percent, following the PQA's definition. The PDC threshold is established at 80 percent based on clinical study results demonstrating that this is the level above which the medication has a reasonable likelihood of achieving the most health benefit.

2.5.3 Enrollment Restrictions and Study Inclusion Criteria

Program participants and matched comparison beneficiaries were included in the claims-based analyses only if they had complete claims information during the entire baseline period, and, depending on model specification, the first or second six months of the post-intervention period. To have complete claims information, program participants and matched comparison beneficiaries must have been continuously enrolled in Medicare during the baseline period, and the first or second six-month period following program participation (or national survey receipt).[30] Beneficiaries who were continuously enrolled in Medicare but switched between FFS and MA were excluded from quantitative analyses because the data-generating processes of these two types of claims are not comparable.

---

[28] The Acumen team also considered the rate of readmissions and skilled nursing facility expenditures as potential analytic outcomes, but these could not be estimated due to the very low sample sizes of beneficiaries with nonzero observations for these outcomes.

[29] See http://pqaalliance.org/

[30] Observations in claims-based analyses correspond to beneficiary-six-months. A beneficiary-six-month observation is included in the analysis if the beneficiary had complete claims information (continuous enrollment) over that six-month time period. For details on exact model specification, please see Appendix A.4.

Outcomes related to pharmaceutical utilization and expenditures required further sample restrictions. Apart from Part D enrollment, which was necessary for the pharmaceutical expenditures outcome, the analysis of adherence required that a beneficiary have at least two prescriptions for the relevant drug class, on two unique dates, covering at least 91 days within the observation period.

In addition, beneficiaries receiving hospice care or end stage renal disease (ESRD) treatment during the observation period (baseline, first, or second six-month post-intervention period) were excluded from the analysis. These beneficiaries are potential outliers, characterized by short life expectancy and atypical health resource utilization.

3 WELLNESS PROGRAM EFFECTS ON SELF-REPORTED HEALTH AND HEALTH BEHAVIORS

This section describes results from analyses of wellness program impacts on self-reported measures of health, wellbeing, and health behavior collected through national and participant surveys at baseline, six months, and twelve months. As described in Section 2, a DiD approach was employed, where changes over time in wellness program participants are compared with changes over time in a matched comparison group. Section 3.1 describes the sample of respondents who completed surveys at all three time points and briefly discusses survey panel attrition at twelve months. Section 3.2 presents results at six and twelve months, broken out by ACA priority area. Section 3.3 discusses these findings.

3.1 Characteristics of Twelve-Month Survey Respondents

This section presents descriptive statistics on the matched samples of beneficiaries who responded to both the six and twelve month surveys, broken out by priority area. As shown in Table 3.1, the matched samples of twelve-month survey respondents differ in demographic composition across the three ACA priority areas. These differences are consistent with those found among baseline and six-month survey respondents.[31] FP program participants and their comparison sample tended to be older and less racially diverse, while CDM program participants and their comparison sample tended to be more racially and ethnically diverse, have lower levels of income and education, and be more likely dual eligible for Medicare and Medicaid. In contrast, PANO program participants and their comparison sample had the highest levels of income and education.

The participant and comparison samples were relatively well matched at twelve months despite panel attrition, with the exception of statistically significant differences in the proportion of urban residents across all ACA priority areas.[32,33] PANO program comparators also had

[31] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[32] Note that statistical tests reported in Table 3.1 are between participants and comparators within an ACA priority area, not across ACA priority areas.

[33] Differences in urban residency may imply differences in access to health care services (see, for example: Goins RT, Williams KA, Carter MW, Spencer SM, Solovieva T. "Perceived barriers to health care access among rural older adults: a qualitative study." *The Journal of Rural Health* 21, no. 3 (June 2005): 206-13.; and Rosenthal TC, Fox C. "Access to health care for the rural elderly." Journal of the American Medical Association 284, no. 16 (October 2000): 2034-6). The quantitative analysis takes this into account in two ways: (i) the DiD estimator controls automatically for any permanent differences between the treatment and comparison groups (e.g. in access to care due to differences in urban residency status) as discussed in Section 2.2; (ii) urban residency status indicators have been added as regressors to the statistical models, to control for the effect of urban residency on outcomes of interest.

significantly lower income and educational attainment, and FP program comparators had a significantly higher proportion of Hispanic beneficiaries. To address these differences, the samples were weighted for survey nonresponse, such that respondents to each follow-up survey were weighted to reflect the characteristics of the full matched samples at baseline. The similarity between unweighted and weighted sample characteristics (shown in Appendix Table C.1) reflects the fact that attrition did not meaningfully change the composition of the matched samples. In addition, as discussed in Section 2, the DiD model specifications also included demographic characteristics as covariates, which further limits any remaining differences (after weighting) from biasing results.

Table 3.1: Unweighted Characteristics Survey-Based Analytic Samples

Characteristic (measured at baseline)	ACA Priority Area					
	CDM		PANO		FP	
	Part. N=529	Comp. N=585	Part. N=656	Comp. N=693	Part. N=1,252	Comp. N=1,339
Average Age[a]	74.6	75.0	74.2	74.5	77.0	77.1
% Female[a]	78.1	78.6	82.8	83.4	77.3	76.3
Race/ethnicity[a]						
% White	77.5	75.6*	83.1	81.7	92.0	92.5
% Black/African American	19.1	23.3	13.4	15.4	5.0	5.2
% Hispanic	1.9	0.5	0.8	0.7	1.6	0.8
% Asian	0.4	0.2	0.5	1.0	1.6	0.3
% Native American	0.0	0.0	0.2	0.1	0.2	0.1
% Other	1.1	0.5	2.1	1.0	0.7	1.1
% Urban[a]	69.6	79.0***	85.4	77.5***	70.8	76.5***
% Dual[a]	13.0	15.4	5.0	6.6	8.9	8.3
Income[b]						
% less than $20,000	51.8	53.9	39.0	44.6	43.9	44.9
% $20,000-$49,999	27.0	26.2	30.0	29.3	31.9	30.6
% $50,000-$99,999	17.0	15.6	24.5	21.7	19.3	19.0
% $100,000 or more	4.2	4.4	6.4	4.5	5.0	5.6
Educational attainment[b]						
% less than high school	14.6	17.1*	9.2	14.1**	8.6	9.0
% high school graduate	25.3	29.2	23.0	24.1	29.6	30.8
% some college/2 year degree	47.1	39.3	44.8	41.4	41.5	40.9
% 4 year college graduate or higher	13.0	14.4	23.0	20.4	20.4	19.3

[a] Characteristics are identified through Medicare enrollment data.
[b] Characteristics are identified through baseline national and participant survey data.
Notes: Part.: Wellness program participants. Comp.: Comparison group. *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01. The p-value is the probability that, if there are no differences in characteristics between participants and the comparison group in each priority area, the observed differences could have occurred by chance in the data. Missing data are included in the lowest income and education categories, and among those of "other" race.

3.2 Survey-Based Program Impact Analysis

This section presents results of the impact of wellness program participation on self-reported health, wellbeing and other health behaviors. Section 3.2.1 provides a brief overview of the survey-based evaluation, and describes how to interpret the charts presented in the remainder

of Section 3. Sections 3.2.2 through 3.2.4 present results for each of the three ACA priority areas.

3.2.1 Overview of Survey-Based Evaluation

The DiD analysis results are presented in both charts and tables. The tables provide information about change over three time points: baseline to six months, six months to twelve months, and baseline to twelve months. For the cumulative estimates (baseline to twelve months), the tables present sample sizes (based on sample members non-missing on the measure), the DiD estimate along with its 90 percent confidence interval, and regression-adjusted means for participants and comparators at the beginning and end of the time frame. Additionally, the relative difference – defined as the DiD estimate divided by the participant mean at the beginning of the estimation period – is shown in each table. All results are weighted and adjusted for covariates as discussed in Section 2. All 15 outcomes are shown in the tables.

How to Interpret the DiD Charts

Each bar chart shows twelve month changes in each outcome:

- in brown stripe for the comparison group
- in blue stripe for the participants
- in teal for the difference of comparison group change and the participant change

Values above zero represent improvements in each outcome, while values below zero reflect deterioration.

Our charts focus on the baseline to twelve-month time horizon and only include variables with statistically significant DiD estimates. They also include single differences for participants and comparators, which are an important part of the story for the self-reported outcomes. For example, a positive DiD estimate can result from very little change in the participant group combined with a sharp decline in the comparison group. For many of our findings, that pattern is illustrated by the charts. It suggests that wellness program participation protects against deterioration in perceived health, wellbeing, and activity levels that may naturally occur as part of the aging process for Medicare beneficiaries.

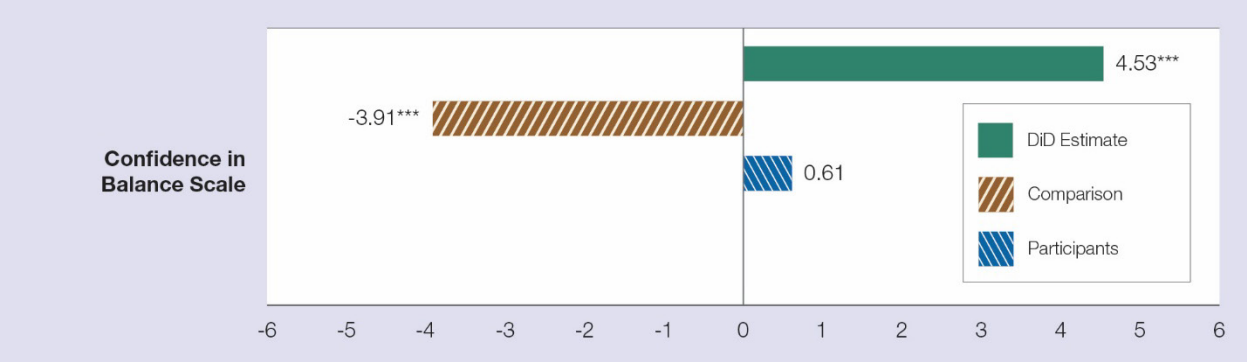
3.2.2 Twelve-month Survey Findings for CDM Programs

There were few statistically significant effects of CDM program participation on self-reported outcomes. Only confidence in balance showed a statistically significant positive effect at twelve months.

Figure 3.1 illustrates how the estimates of change for participants and the comparison group work together to create the statistically significant DiD estimate. For confidence in balance, improvements among participants, paired with declines among members of the comparison group, result in an overall improvement in confidence in balance among program

participants relative to the comparison group. The DiD estimate of 4.5 is calculated as the difference between the participant group change (0.6) and the comparison group change (-3.9).

Figure 3.1: Effects of Chronic Disease Management Programs at Twelve Months



Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. This figure shows the DiD estimate, along with single difference estimates for program participants and the comparison group.

As shown in Table 3.2, CDM programs did not improve self-reported physical health at any time point.

Table 3.2: Physical Health Results for Chronic Disease Management Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	506/571	529/577	504/577	506/577	529/585
Difference-in-Difference	0.21	0.60	0.31	0.25	-0.03
P-value	0.66	0.23	0.53	0.67	0.96
90% Confidence Interval	(-0.6,1.0)	(-0.2,1.4)	(-0.5,1.1)	(-0.7,1.2)	(-0.9,0.9)
Baseline Participant Mean	41.5	39.6	41.8	44.2	47.9
Twelve-Month Participant Mean	41.2	39.4	41.7	44.6	47.2
Baseline Comparison Mean	41.4	40.4	41.8	44.7	48.7
Twelve-Month Comparison Mean	40.9	39.5	41.3	44.9	48.0
Relative Difference	0.5%	1.5%	0.7%	0.6%	-0.1%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	510/577	529/578	509/577	509/583	529/585
Difference-in-Difference	-0.15	0.70	0.29	-0.47	0.17
P-value	0.71	0.14	0.57	0.37	0.73
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	488/565	529/570	487/569	489/578	529/585
Difference-in-Difference	0.62	-0.12	0.31	0.74	-0.20
P-value	0.12	0.78	0.49	0.13	0.69

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10%

level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

Table 3.3 illustrates that there were short-term mental health benefits (baseline to six months) of CDM program participation for role limitations of mental health and overall mental health. In contrast to the confidence in balance findings reported above, these short-term benefits were driven more by improvements in the participant group than deterioration for the comparison group. However, the benefits were small and did not persist at twelve months.

Table 3.3: Mental Health Results for Chronic Disease Management Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	506/571	528/579	527/579	503/573	529/579
Difference-in-Difference	0.04	0.11	-0.17	0.52	0.26
P-value	0.94	0.82	0.79	0.43	0.63
90% Confidence Interval	(-0.9,1.0)	(-0.7,0.9)	(-1.2,0.9)	(-0.6,1.6)	(-0.6,1.1)
Baseline Participant Mean	50.7	48.9	47.0	44.1	50.6
Twelve-Month Participant Mean	50.5	48.2	47.0	44.3	50.3
Baseline Comparison Mean	51.9	49.7	47.1	45.5	52.0
Twelve-Month Comparison Mean	51.6	48.9	47.3	45.2	51.6
Relative Difference	0.1%	0.2%	-0.4%	1.2%	0.5%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	510/577	528/584	528/584	507/574	528/584
Difference-in-Difference	0.72	-0.32	-0.35	1.41**	1.06**
P-value	0.18	0.55	0.58	0.03	0.03
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	488/565	527/579	527/579	486/567	528/579
Difference-in-Difference	-0.65	0.40	0.26	-0.92	-0.81
P-value	0.25	0.38	0.68	0.18	0.13

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

Table 3.4 highlights the persistent statistically significant finding of benefits for confidence in balance at twelve months, shown in Figure 3.1. Another notable finding from Table 3.4 is the absence of an effect on self-reported medication adherence, a commonly targeted outcome for improvement in CDM programs.

Table 3.4: Activity, Balance, and Medication Adherence Results for Chronic Disease Management Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	513/563	488/546	477/539	352/392	451/528
Difference-in-Difference	-0.14	0.02	0.05	4.53***	-0.02
P-value	0.30	0.64	0.15	0.01	0.83
90% Confidence Interval	(-0.4,0.1)	(-0.0,0.1)	(-0.0,0.1)	(1.8,7.3)	(-0.2,0.1)
Baseline Participant Mean	4.8	0.5	0.2	51.4	3.1
Twelve-Month Participant Mean	4.6	0.5	0.2	52.0	3.2
Baseline Comparison Mean	4.5	0.5	0.2	59.3	3.0
Twelve-Month Comparison Mean	4.4	0.5	0.2	55.4	3.2
Relative Difference	-2.9%	3.7%	24.3%	8.8%	-0.7%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	517/561	490/549	477/537	348/385	465/521
Difference-in-Difference	-0.13	0.01	-0.00	3.33**	-0.02
P-value	0.29	0.79	0.99	0.04	0.77
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	517/572	498/561	487/545	370/392	464/527
Difference-in-Difference	-0.07	0.00	0.03	1.03	0.00
P-value	0.58	0.93	0.27	0.46	0.97

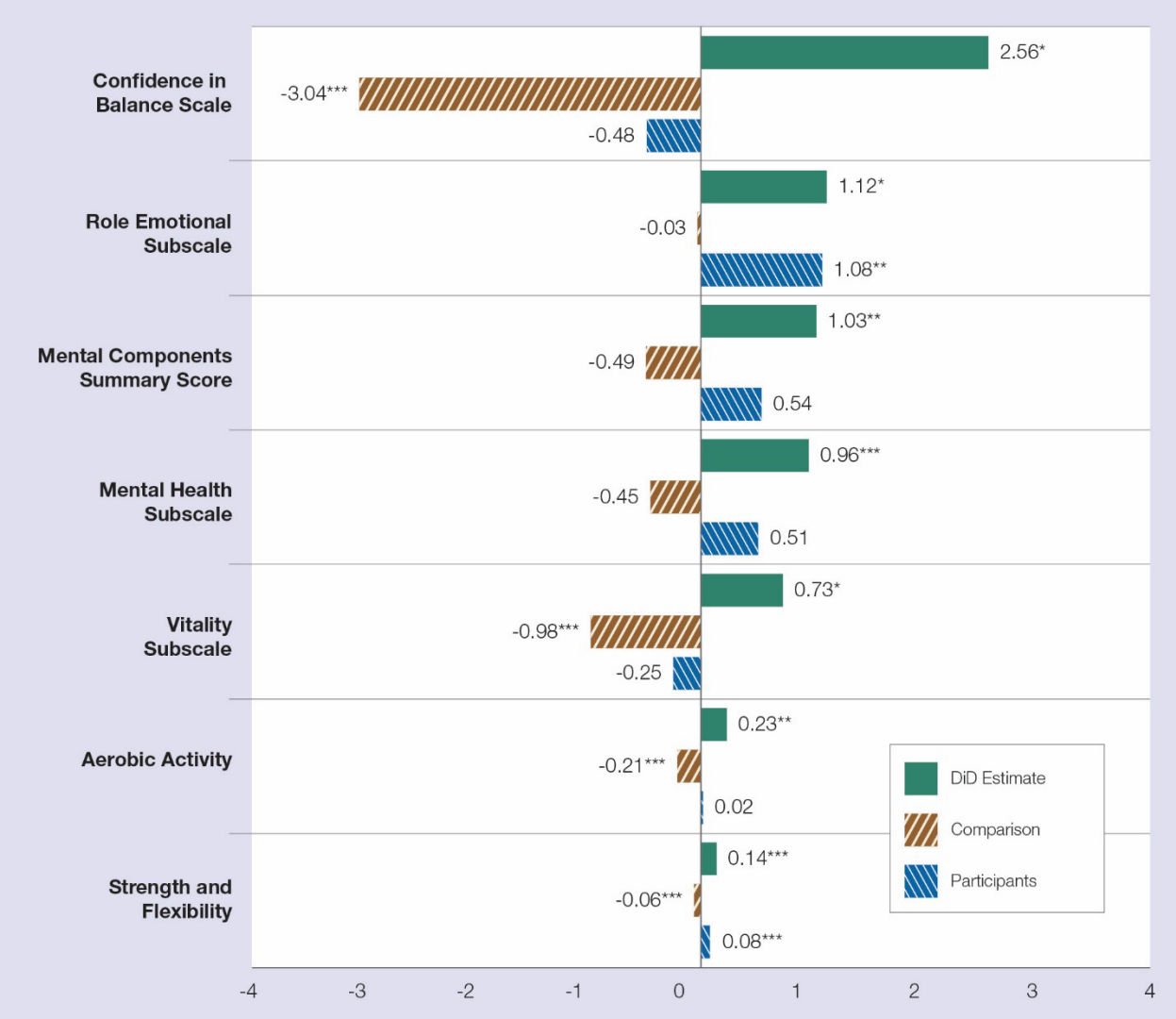
Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

3.2.3 Twelve-month Survey Findings for PANO Programs

PANO programs generated benefits for seven of the 15 total measures at twelve months (Figure 3.2). Three measures - aerobic activity, strength and flexibility, and confidence in balance - are related to physical health and health behaviors. The remaining four measures - role emotional subscale, mental health subscale, vitality subscale, and the mental components summary score - are related to mental wellbeing. For most measures, negligible improvements among participants are compared with declines in the comparison group. Notable exceptions

include statistically significant improvements among program participants in strength and flexibility training and role limitations related to mental health. The overall pattern suggests that PANO programs are protective against deterioration in physical activity levels and mental wellbeing that may occur over time due to aging.

Figure 3.2: Effects of Physical Activity, Nutrition, and Obesity Programs at Twelve Months



Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. This figure shows the DiD estimate, along with single difference estimates for program participants and the comparison group.

As shown in Table 3.5, PANO programs created a small benefit in terms of limitations related to physical health (role physical subscale measure) at six months, but that benefit did not persist at twelve months. In general, PANO programs had little impact on self-reported physical health.

Table 3.5: Physical Health Results for Physical Activity, Nutrition, and Obesity Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	631/676	656/680	631/679	629/687	656/693
Difference-in-Difference	-0.00	0.55	0.25	0.53	-0.23
P-value	1.00	0.15	0.59	0.31	0.57
90% Confidence Interval	(-0.6,0.6)	(-0.1,1.2)	(-0.5,1.0)	(-0.3,1.4)	(-0.9,0.4)
Baseline Participant Mean	45.4	44.1	45.3	47.1	52.4
Twelve-Month Participant Mean	45.0	43.9	45.6	47.6	51.3
Baseline Comparison Mean	45.4	44.5	45.2	47.2	51.6
Twelve-Month Comparison Mean	45.0	43.8	45.2	47.2	50.7
Relative Difference	0.0%	1.2%	0.6%	1.1%	-0.4%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	635/677	655/683	635/683	635/684	655/693
Difference-in-Difference	0.35	0.46	0.93**	0.69	-0.23
P-value	0.331	0.235	0.032	0.166	0.543
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	611/661	655/670	611/669	609/680	655/693
Difference-in-Difference	-0.29	0.17	-0.53	-0.23	0.00
P-value	0.44	0.63	0.24	0.68	0.99

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

In contrast, PANO programs generated benefits for mental health and wellbeing across all of the mental health measures. Many DiD estimates at six months increased at twelve months, some becoming significant only at the twelve-month mark (energy levels and overall mental health). Only social functioning showed reduced benefit at twelve months. As noted above, PANO programs appear to provide protection against deterioration rather than generate large improvements for participants. The relative differences reported in Table 3.6 indicate that these protective benefits are quite small at twelve months; the largest relative difference was only 2.4% for role limitations related to mental health.

Table 3.6: Mental Health Results for Physical Activity, Nutrition, and Obesity Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	631/676	656/689	656/689	628/679	656/689
Difference-in-Difference	1.03**	0.73*	0.74	1.12*	0.96**
P-value	0.04	0.08	0.14	0.07	0.05
90% Confidence Interval	(0.2,1.9)	(0.0,1.4)	(-0.1,1.6)	(0.1,2.1)	(0.2,1.8)
Baseline Participant Mean	53.0	51.9	50.0	47.3	53.0
Twelve-Month Participant Mean	53.5	51.6	50.0	48.3	53.5
Baseline Comparison Mean	52.9	52.2	49.9	47.4	52.9
Twelve-Month Comparison Mean	52.4	51.2	49.2	47.4	52.4
Relative Difference	1.9%	1.4%	1.5%	2.4%	1.8%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	635/677	655/688	654/687	631/682	655/688
Difference-in-Difference	0.81*	0.41	0.85*	1.09*	0.45
P-value	0.093	0.314	0.080	0.063	0.295
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	611/661	655/684	654/683	608/669	655/684
Difference-in-Difference	0.31	0.33	-0.13	-0.01	0.51
P-value	0.50	0.41	0.80	0.99	0.25

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

Not surprisingly, PANO programs increased the level of physical activity reported by respondents at twelve months (Table 3.7). The improvement was most pronounced for strength and flexibility, with a relative difference of 20.6 percent. For confidence in balance, participants were relatively stable over time, while confidence in balance declined significantly for the comparison group. This decline occurred between baseline and six-month follow-up surveys, and again between six-month and twelve-month follow-up surveys.

Table 3.7: Activity, Balance, and Medication Adherence Results for Physical Activity, Nutrition, and Obesity Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Number of Beneficiaries (Participants/Comparators)	648/675	604/648	604/647	447/475	545/587
Difference-in-Difference	0.23**	0.14***	0.01	2.56*	0.03
P-value	0.03	0.00	0.75	0.06	0.63
90% Confidence Interval	(0.1,0.4)	(0.1,0.2)	(-0.0,0.0)	(0.3,4.8)	(-0.1,0.1)
Baseline Participant Mean	5.1	0.7	0.2	63.5	3.2
Twelve-Month Participant Mean	5.1	0.8	0.2	63.1	3.3
Baseline Comparison Mean	5.0	0.7	0.2	67.3	3.2
Twelve-Month Comparison Mean	4.8	0.6	0.2	64.3	3.3
Relative Difference	4.5%	20.6%	6.0%	4.0%	0.9%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	647/674	600/648	602/651	455/457	546/581
Difference-in-Difference	0.19*	0.15***	0.03	0.90	-0.04
P-value	0.054	0.000	0.240	0.527	0.494
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	651/674	602/656	603/652	463/470	545/598
Difference-in-Difference	0.03	-0.03	-0.02	1.15	0.06
P-value	0.77	0.36	0.37	0.30	0.30

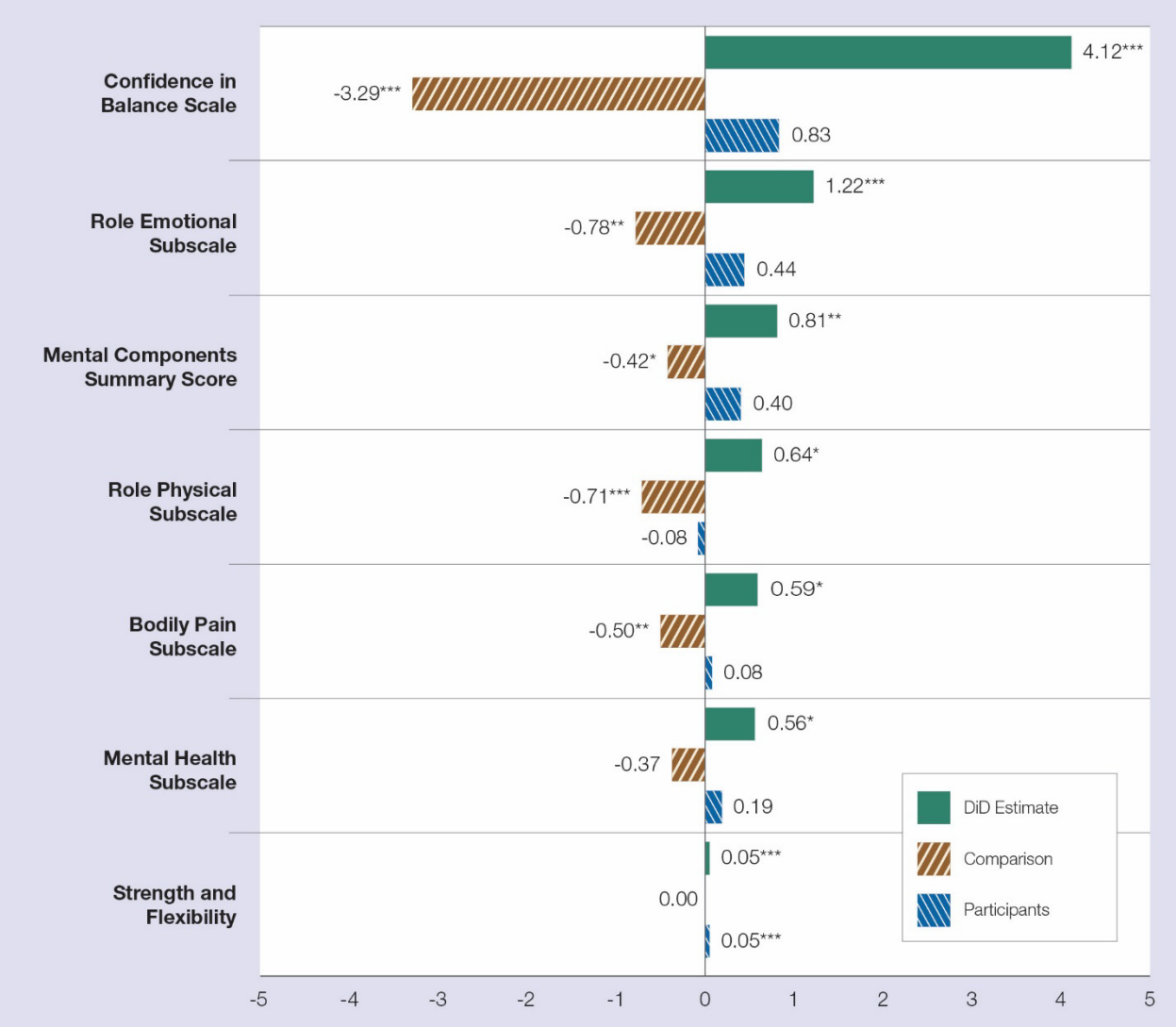
Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

3.2.4 Twelve-month Survey Findings for FP Programs

FP programs generated benefits for seven of the 15 total measures at twelve months (Figure 3.3). As expected, FP program effects were found for confidence in balance and strength and flexibility activities (an important component of balance training). FP programs may also have benefits for bodily pain and beneficiary ability to perform typical activities (role physical subscale).

Consistent with findings for PANO programs, FP program effects were also found for a number of the mental health measures, including the role emotional subscale, the mental health subscale, and the mental components summary score. Once again the estimated program effects were typically driven by deterioration in the comparison group more than improvement among program participants. A notable exception is the improvement for FP program participants in strength and flexibility activity.

Figure 3.3: Effects of Falls Prevention Programs at Twelve Months



Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. This figure shows the DiD estimate, along with single difference estimates for program participants and the comparison group.

Table 3.8 shows that the FP programs' effect on limitations to performing work and other activities (role physical subscale), as well as bodily pain, was present at six months and changed little in magnitude by twelve months. However, the effect is extremely small, with relative differences less than 2 percent.

Table 3.8: Physical Health Results for Falls Prevention Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Number of Beneficiaries (Participants/Comparators)	1,187/1,307	1,249/1,317	1,187/1,316	1,187/1,326	1,246/1,339
Difference-in-Difference	0.24	0.39	0.64*	0.59*	0.13
P-value	0.42	0.24	0.05	0.05	0.58
90% Confidence Interval	(-0.3,0.7)	(-0.2,0.9)	(0.1,1.2)	(0.1,1.1)	(-0.3,0.5)
Baseline Participant Mean	41.9	40.3	42.1	45.1	49.4
Twelve-Month Participant Mean	41.3	39.7	42.0	45.2	48.5
Baseline Comparison Mean	43.1	41.4	43.0	46.1	49.9
Twelve-Month Comparison Mean	42.2	40.4	42.3	45.6	48.9
Relative Difference	0.6%	1.0%	1.5%	1.3%	0.3%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,200/1,304	1,251/1,313	1,199/1,312	1,200/1,325	1,249/1,338
Difference-in-Difference	0.01	0.18	0.75**	0.54*	-0.13
P-value	0.96	0.53	0.02	0.08	0.61
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,141/1,274	1,249/1,291	1,141/1,289	1,144/1,319	1,244/1,338
Difference-in-Difference	0.25	0.22	-0.11	0.15	0.25
P-value	0.34	0.38	0.73	0.65	0.32

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

Similarly, Table 3.9 shows that the program benefit for mental health measures was present at six months, relatively stable at twelve months, and very small in magnitude (relative differences less than 3 percent).

Table 3.9: Mental Health Results for Falls Prevention Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	1,187/1,307	1,249/1,329	1,250/1,329	1,186/1,314	1,249/1,329
Difference-in-Difference	0.81**	0.30	0.52	1.22***	0.56*
P-value	0.02	0.28	0.12	0.01	0.09
90% Confidence Interval	(0.3,1.4)	(-0.2,0.8)	(-0.0,1.1)	(0.5,1.9)	(0.0,1.1)
Baseline Participant Mean	51.7	49.5	48.0	45.1	51.7
Twelve-Month Participant Mean	52.1	49.1	48.0	45.6	51.9

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Baseline Comparison Mean	51.8	50.1	48.0	45.9	51.8
Twelve-Month Comparison Mean	51.4	49.4	47.5	45.1	51.5
Relative Difference	1.6%	0.6%	1.1%	2.7%	1.1%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,200/1,304	1,249/1,329	1,246/1,328	1,192/1,309	1,250/1,329
Difference-in-Difference	0.94***	0.05	0.40	1.51***	0.74**
P-value	0.01	0.86	0.30	0.00	0.02
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,141/1,274	1,247/1,320	1,245/1,319	1,137/1,288	1,248/1,320
Difference-in-Difference	-0.28	0.20	0.10	-0.28	-0.21
P-value	0.39	0.48	0.79	0.50	0.47

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

FP program participants declined in levels of aerobic activity at six months relative to their comparators, though the effects were not statistically significant at twelve months (Table 3.10). However, FP programs provided benefits for levels of strength and flexibility and confidence in balance at six and twelve months. The magnitude of effect is moderate for both outcomes (relative difference of 8.8 and 8.1 percent, respectively). No program effect on self-reported number of falls was found at six or twelve months.

Table 3.10: Activity, Balance, and Medication Adherence Results for Falls Prevention Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	1,220/1,291	1,134/1,217	1,128/1,229	895/908	1,061/1,169
Difference-in-Difference	-0.12	0.05**	0.02	4.12***	-0.00
P-value	0.11	0.02	0.51	0.00	0.93
90% Confidence Interval	(-0.2,0.0)	(0.0,0.1)	(-0.0,0.1)	(2.5,5.7)	(-0.1,0.1)
Baseline Participant Mean	4.7	0.6	0.3	50.7	3.2
Twelve-Month Participant Mean	4.4	0.6	0.3	51.5	3.2
Baseline Comparison Mean	4.6	0.5	0.3	56.0	3.2
Twelve-Month Comparison Mean	4.5	0.5	0.2	52.7	3.3
Relative Difference	-2.5%	8.8%	5.9%	8.1%	-0.0%

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,223/1,279	1,152/1,213	1,149/1,219	928/930	1,076/1,174
Difference-in-Difference	-0.19**	0.04*	0.02	2.66***	0.01
P-value	0.01	0.05	0.36	0.01	0.84
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,223/1,305	1,151/1,270	1,150/1,258	915/957	1,085/1,200
Difference-in-Difference	0.07	0.01	-0.01	1.00	0.00
P-value	0.28	0.69	0.79	0.25	0.98

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

3.3 Discussion of Survey-Based Evaluation Findings

Across ACA priority areas, wellness programs showed consistently beneficial effects on many of the self-reported health and behavioral measures, but in most cases, these effects were small. Notable exceptions included the impact of PANO programs on strength and flexibility activities (relative difference of 20.6%), the impact of FP programs on strength and flexibility activities (8.8%) and confidence in balance (8.1%), and the impact of CDM programs on confidence in balance (8.8%). CDM programs had far fewer significant effects when compared with the other ACA priority areas.

For PANO and FP programs, the pattern of effects was generally consistent with the focus areas and design of the programs. For example, PANO programs improved participants' self-reported physical activity levels, and FP programs improved confidence in balance. Surprisingly, the most consistently positive program impacts were on mental health: both PANO and FP programs were associated with small, statistically significant positive effects on different aspects of mental health. This is an interesting finding, since the intended impacts of these programs are primarily medical and physical in nature. An important secondary benefit of participation is enhanced mental well-being, which may result from lifestyle changes (e.g., increased physical activity), knowledge gained from programs, or the social act of program participation.

An encouraging finding of the analysis was that many program effects persisted at twelve months. This is notable because most of the programs were time limited (6-8 weeks on average), and it might be expected that some program effects will fade over time as participants become

more distant from the intervention. Program effects on confidence in balance were actually larger at twelve months for all three ACA priority areas, and FP program effects on strength and flexibility activities were also larger at twelve months. For PANO, in addition to improved confidence in balance over time, program effects also increased between 6 and 12 months for several of the mental health subscales.

The results suggest that, overall, wellness programs may protect against deterioration in health and activity that naturally occurs with aging, as opposed to generating notable improvements in self-reported health and activity for participants. The key driver of these favorable findings appears to be minimal change over time across many outcome measures among participants, combined with a decline in outcome measures within the comparison group. Notable exceptions include statistically significant participant gains in strength and flexibility activities for FP and PANO programs. Overall, participants reported only slight, typically non-significant improvements in overall health and physical capabilities.

4 WELLNESS PROGRAM EFFECTS ON HEALTHCARE UTILIZATION, EXPENDITURE, AND MEDICATION ADHERENCE

Section 4 presents the evaluation of wellness program impacts on claims-based healthcare utilization, expenditure, and medication adherence outcomes. Section 4.1 describes the characteristics of the claims-based analytic sample. Section 4.2 presents the findings from the analysis and provides an overview and discussion of the limitations of the claims-based evaluation. Finally, Section 4.3 provides a discussion of our findings.

4.1 Characteristics of the Claims-Based Analytic Population

As described in Section 2.5, the analysis of wellness program impacts on healthcare utilization, expenditure, and medication adherence uses Medicare claims data and incorporates a different set of sample restrictions than the survey-based analysis. Table 4.1, which presents the analytic sample size for each priority area after the application of each restriction for the claims-based analysis, shows that across priority areas, about 40 percent of the matched sample was excluded from the analysis due to discontinuous enrollment in Medicare FFS (or enrollment in MA).

Table 4.1: Claims-Based Analytic Sample and Exclusions

Exclusions	CDM		PANO		FP	
	Part.	Comp.	Part.	Comp.	Part.	Comp.
Starting Sample Matched at Baseline	920	920	1,046	1,046	2,013	2,013
Sample Size After Successive Exclusions at Baseline						
Beneficiaries not continuously enrolled	533	533	595	595	1,166	1,166
Beneficiaries with ESRD	529	529	593	594	1,162	1,161
Beneficiaries receiving hospice care	529	528	593	593	1,160	1,161
Sample Size After Successive Exclusions at 0-6 Months Post-Intervention						
Beneficiaries not continuously enrolled	518	519	584	586	1,138	1,145
Beneficiaries with ESRD	517	519	584	586	1,138	1,145
Beneficiaries receiving hospice care	514	517	583	584	1,133	1,136
Sample Size After Successive Exclusions at 7-12 Months Post-Intervention						
Beneficiaries not continuously enrolled	510	508	576	578	1,107	1,125
Beneficiaries with ESRD	509	508	576	578	1,107	1,125
Beneficiaries receiving hospice care	506	502	574	578	1,097	1,114

Notes: Part.: Wellness program participants. Comp.: Comparison group. Counts presented in this table are the number of beneficiaries remaining in the analytic sample after each exclusion was made. Beneficiary observations were excluded from the analysis if beneficiaries did not have continuous enrollment in FFS, or if they received hospice care or ESRD treatment during the 12-month baseline period, or during the first or second six months after program participation (or survey receipt).

As shown in Table 4.2, the baseline demographic characteristics of the analytic sample for the analysis of wellness program impacts on healthcare utilization, expenditure, and medication adherence are generally similar to the survey-based analysis of self-reported health

and health behaviors (Section 3). FP program participants and their matched comparisons were older and less racially diverse, CDM program participants and matched comparison beneficiaries were more likely to be dually eligible for Medicare and Medicaid, and PANO participants and matched comparators were more likely to be female. However, there are some slight differences. For example, as shown in Table 4.2, there are fewer black beneficiaries in the claims-based analytic population for CDM programs, compared to those in the survey-based analytic population (Table 3.1). This difference is due to differences in Medicare enrollment; proportionally fewer black beneficiaries are continuously enrolled in Medicare FFS.[34] Beneficiaries across priority areas also differ in their baseline health status. PANO program participants are healthier, with lower counts of IP stays, ER visits, and expenditures. As expected, CDM program participants have higher pharmaceutical expenditure than program participants in other ACA priority areas.

The participant and comparison groups for the claims-based analysis on utilization, expenditure, and medication adherence are generally well matched. As shown in Table 4.2, similar to the analytic population for the survey-based analysis, there are differences between participant and comparison group beneficiaries in some demographic characteristics such as urban residence status, race, and dual eligibility status.[35]

Table 4.2: Baseline Characteristics of Claims-Based Analytic Samples

Characteristic (baseline year)	ACA Priority Area					
	CDM		PANO		FP	
	Part. N=529	Comp. N=528	Part. N=593	Comp. N=593	Part. N=1,160	Comp. N=1,161
Average Age	75.3	75.6	74.7	74.8	77.8	77.5
% Female	77.1	77.5	83.1	83.0	76.6	76.6
Race						
% White	81.9*	83.9	83.1	83.8	90.9**	92.6
% Black	15.1*	15.2	14.5	14.5	5.1**	5.3
% Other	3.0*	0.9	2.4	1.7	4.0**	2.2
% Dual Eligible	10.4	12.5	4.0*	6.6	9.0	8.7
% Urban	63.7***	72.5	86.5***	73.5	70.4	73.2
Evaluation and Management (E&M) Visits						
% E&M Visits: 0	2.5	2.3	5.1	4.9	2.8	3.1
% E&M Visits: 1-10	55.2	58.1	65.9	69.1	59.2	58.8
% E&M Visits: 11+	42.3	39.6	29.0	26.0	37.9	38.1
IP Stays						
% 0 IP Stays (Prior Year)	83.7	84.1	88.9	89.0	85.0	84.2
% 1 IP Stay (Prior Year)	12.5	12.1	7.4	6.7	11.0	11.9
% 2+ IP Stays (Prior Year)	3.8	3.8	3.7	4.2	4.0	3.9

[34] Among beneficiaries matched at baseline, 15% of CDM beneficiaries enrolled in Medicare FFS and 33% of CDM beneficiaries not enrolled in Medicare FFS identified as black.

[35] Urban/rural status indicators, dual eligibility status indicators, gender, race, age, education, and income variables have been added as covariates to all estimation models, with the exception of falls and factures and home health outcomes, where low sample size did not allow for the addition of model covariates.

Characteristic (baseline year)	ACA Priority Area					
	CDM		PANO		FP	
	Part. N=529	Comp. N=528	Part. N=593	Comp. N=593	Part. N=1,160	Comp. N=1,161
ER Visits						
% ER Visits: 0	72.0	72.2	79.6	78.4	71.9	74.7
% ER Visits: 1	18.3	17.0	14.5	16.2	19.4	17.7
% ER Visits: 2+	9.6	10.8	5.9	5.4	8.7	7.6
Total Part A and B Cost per Beneficiary	$7,612	$7,417	$5,327	$5,306	$6,617	$6,632
IP Cost per Beneficiary	$1,767	$1,814	$1,376	$1,295	$1,565	$1,499
Part D Cost per Beneficiary	$4,042	$4,570	$2,495	$2,667	$2,951	$2,809

Notes: Part.: Program participants. Comp.: Comparison group. IP: Inpatient; ER: Emergency Room. *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01. The p-value is the probability that, if there are no differences in characteristics between participants and the comparison group in each priority area, the observed differences could have occurred by chance in the data. Part D cost per beneficiary only accounts for beneficiaries who have Part D coverage. E&M visits do not include annual wellness visits or visits to FQHCs.

4.2 Claims-Based Program Impact Analysis

This section presents findings from the evaluation of the effect of wellness program participation on healthcare utilization, expenditures, and medication adherence using information from Medicare FFS claims data. Section 4.2.1 provides an overview and discussion of the limitations of the claims-based evaluation. Sections 4.2.2 through 4.2.4 present analytic findings by ACA priority area.

4.2.1 Limitations and Overview of Claims-Based Evaluation

The claims-based evaluation estimates the effect of wellness program participation on healthcare utilization, expenditures, and medication adherence by priority area for beneficiaries enrolled in Medicare FFS.[36] Appendix Section C.1 presents summary statistics for each analytic cohort. As discussed in Section 4.1, demographic characteristics of the claims-based sample are generally similar to the survey-based sample.

The conclusions drawn from the claims-based evaluation are subject to a number of limitations related to the small sample sizes available for analysis, and the small number of nonzero observations for many utilization and expenditures outcomes. The analytic sample size differs across ACA priority areas,[37] but, in general, statistical power is low for the claims-based analyses. Low statistical power makes it harder for the analysis to detect an effect of

[36] The Acumen team explored the incidence of mental health claims as an additional outcome, given the survey-based findings on mental health improvements for program participants. Acumen ran exploratory DiD analyses on the incidence of mental health claims, defined as the proportion of beneficiaries with at least one mental health related claim. Acumen used an "umbrella" definition of mental health utilization, which included high-intensity events (e.g., suicides), but also low-intensity outcomes (e.g., visits to a mental health practitioner). There was no impact of program participation in any priority area on this outcome, and, since this measure was only intended for an exploratory analysis and is not a validated measure of mental health utilization, these findings are not reported here.

[37] The number of participant and comparison beneficiaries is almost double in FP programs compared to CDM or PANO programs. See Appendix Table C.1 for details.

participation in wellness programs, especially if the size of the effect is small. In addition, for many claims-based outcomes, due to the nature of the metrics (for example, counts of IP admissions, or ER visits), there is only a small proportion of participant and comparison group beneficiaries with nonzero observations driving the statistical estimates. Because sample sizes are small, this implies that, for many outcomes, the number of nonzero observations is very low (Appendix Section C.1). For example, the calculation of the incidence of falls/fractures requires beneficiaries to have at least one fall- or fracture-related claim, and there are very few beneficiaries in each priority area who contribute nonzero observations for this outcome. One exception is total Parts A and B and physician/ancillary expenditures, which take positive values for most beneficiaries (Appendix Table C.3). The analysis of Part D claims suffers from even lower sample sizes given further sample restrictions required for adherence calculations. As a result, adherence findings should be interpreted with caution.

Low sample sizes cannot be remediated by pooling together multiple cohorts of analysis. For example, Medicare FFS beneficiaries cannot be combined with MA beneficiaries into a single analytic cohort, because the data generating processes differ across the two types of claims data. This evaluation focuses on beneficiaries enrolled in FFS, because sample sizes of beneficiaries enrolled in MA cohort were too small for the purposes of statistical analysis. In addition, programs in each priority area have different scopes and attract different types of beneficiaries. As shown on Table 4.2, there are differences in baseline demographic and health characteristics across the three priority areas. As a result, pooling beneficiaries across priority areas into a single analytic cohort would produce results that are of limited value to policy makers, and would be hard to interpret.

The following sections describe the claims-based analytic results by priority area. These sections present findings for utilization outcomes (ER visits, IP admissions, and incidence of falls/fractures[38]); expenditures on total medical, IP, outpatient ER, outpatient non-ER, and physician and ancillary services; and average adherence and rates of high adherence (PDC ≥ 80%) to beta blockers, calcium channel blockers, diabetes medication, RAS antagonists, and statins. Full results of the cumulative analyses are presented in tables, with information on the number of nonzero and total observations,[39] DiD point estimates, p-values, 90% confidence intervals, and adjusted baseline and post-intervention means for program participants and the comparison group. Relative differences are also presented, defined as the ratio of the DiD point estimate divided by the baseline participant mean. Point estimates, p-values, and sample sizes are also shown for the interim analyses (0-6 months and 7-12 months post-intervention).

[38] The numerator for the incidence of fall/fractures is defined as the number of beneficiaries who had at least one falls- or fracture-related claim.
[39] As discussed in Appendix Section A.2, observations correspond to beneficiary-half years.

The Acumen team also analyzed unplanned IP admissions and length of stay, total expenditures on Part D, durable medical equipment, and home health. There were no statistically significant interim or cumulative effects, other than a cumulative decrease in home health spending for FP programs, and thus these findings are presented in the Appendix Section C.3 rather than the main report.

4.2.2 Findings for CDM Programs

Among CDM wellness program participants, there were more ER visits and increased adherence to calcium channel blockers relative to the comparison group across the entire twelve-month post-intervention period. There were no statistically significant findings on expenditures.

As shown in Table 4.3, ER visits increased substantially, by 150 per 1,000 beneficiaries; however, this finding was not paired with statistically significant increases in outpatient ER expenditures (Table 4.4).

As shown in Table 4.5, average adherence to calcium channel blockers among CDM participants increased by 5 percentage points, driven by increased adherence in the first six months post-intervention. Adherence improved across both program participants and the comparison group, but the improvement was bigger among participants than among the comparison group. Similarly, the rate of highly adherent participants (PDC ≥ 80%) increased by 14 percentage points (Figure 4.1 and Appendix Table C.7).

Table 4.3: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, CDM Programs

Measures	CDM		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	197/1,020	113/1,020	130/1,020
Difference-in-Difference	149.97*	36.47	-13.90
P-value	0.06	0.52	0.58
90% Confidence Interval	(21.5, 278.4)	(-56.0, 128.9)	(-55.2, 27.4)
Baseline Participant Mean	289.78	239.41	117.20
Intervention Period Participant Mean	391.07	309.15	127.45
Baseline Comparison Mean	275.06	204.63	123.11
Intervention Period Comparison Mean	226.38	238.01	147.26
Relative Difference	51.8%	15.2%	-11.9%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	94/514	59/514	48/514
Difference-in-Difference	41.84	25.13	-39.98
P-value	0.38	0.50	0.16
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	103/506	54/506	82/506
Difference-in-Difference	108.70**	11.43	12.58

Measures	CDM		
	ER Visits	Inpatient Admissions	Falls/ Fractures
P-value	0.02	0.73	0.68

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income, except for the falls/fracture outcome due to small sample size. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage

Table 4.4: Medicare Expenditures per Beneficiary, CDM Programs

Measures (2011 USD)	CDM				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	995/1,020	111/1,020	197/1,020	690/1,020	991/1,020
Difference-in-Difference	$1,282.11	$325.69	$99.15	- $8.89	$363.05
P-value	0.25	0.56	0.20	0.98	0.28
90% Confidence Interval	(-534.6, 3,098.8)	(-603.9, 1,255.3)	(-27.0, 225.3)	(-484.4, 466.6)	(-186.7, 912.8)
Baseline Participant Mean	$9,494.50	$2,160.98	$212.45	$1,888.21	$3,093.40
Intervention Period Participant Mean	$10,685.00	$2,677.83	$262.34	$1,920.81	$3,227.00
Baseline Comparison Mean	$9,186.05	$2,127.67	$190.87	$1,891.84	$2,976.96
Intervention Period Comparison Mean	$9,094.08	$2,318.83	$141.61	$1,933.33	$2,747.52
Relative Difference	13.5%	15.1%	46.7%	-0.5%	11.7%
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	503/514	58/514	94/514	349/514	501/514
Difference-in-Difference	$890.93	$364.99	$45.13	- $111.33	$187.30
P-value	0.19	0.27	0.36	0.56	0.35
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	492/506	53/506	103/506	341/506	490/506
Difference-in-Difference	$395.58	- $37.63	$54.41	$103.06	$173.75
P-value	0.55	0.92	0.17	0.53	0.35

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means

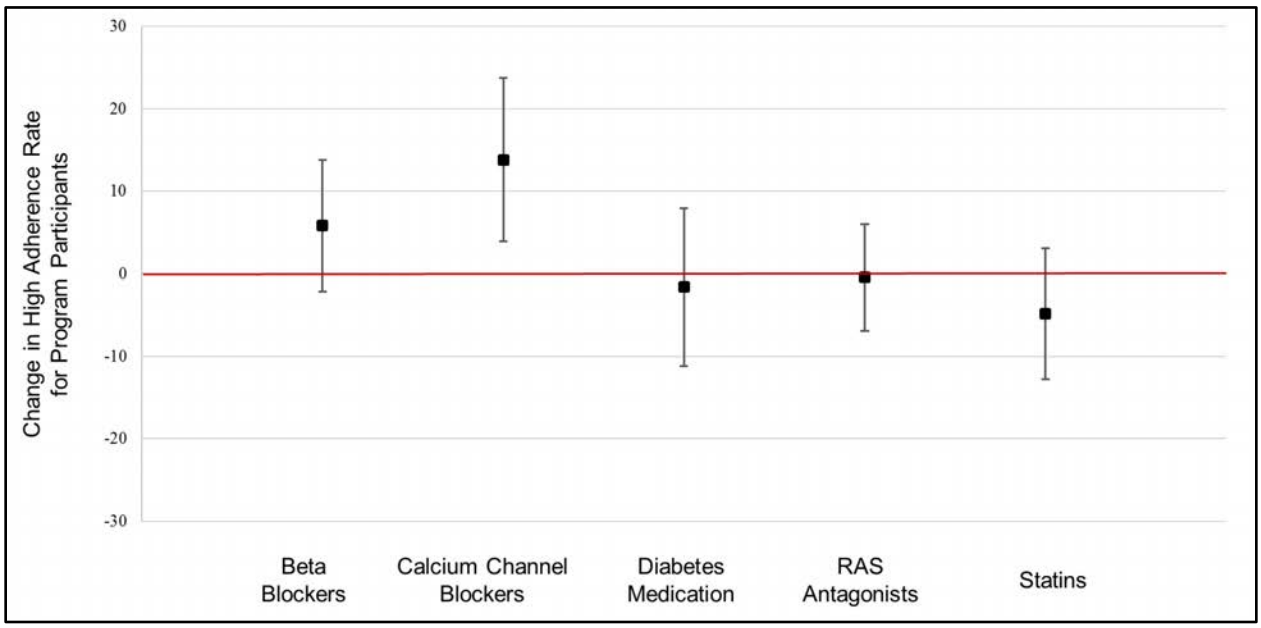
are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Table 4.5: Medication Adherence (Average Proportion of Days Covered), CDM Programs

Measures (Average PDC)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	248	190	168	341	326
Difference-in-Difference	2.01	5.30*	-0.05	0.46	-2.27
P-value	0.33	0.05	0.98	0.78	0.23
90% Confidence Interval	(-1.4, 5.4)	(0.8, 9.7)	(-3.9, 3.8)	(-2.3, 3.2)	(-5.4, 0.8)
Baseline Participant Mean	$88.36	90.16	90.76	91.01	85.60
Intervention Period Participant Mean	$93.64	99.41	95.07	95.30	91.60
Baseline Comparison Mean	$89.34	92.05	91.93	91.14	82.95
Intervention Period Comparison Mean	$92.60	96.02	96.24	94.97	91.20
Relative Difference	2.3%	5.9%	-0.1%	0.5%	-2.7%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	126	96	88	164	167
Difference-in-Difference	1.78	7.10**	-1.01	0.26	-3.14
P-value	0.42	0.02	0.69	0.89	0.14
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	122	94	80	177	159
Difference-in-Difference	2.32	3.17	1.15	0.69	-1.26
P-value	0.31	0.26	0.66	0.70	0.54

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Figure 4.1: High Adherence Rate (PDC ≥ 80%), Cumulative Estimates, CDM Programs



Notes: The y-axis represents the percentage point change in the rate of highly adherent beneficiaries in the year following program participation. The solid circle represents the estimated change in high adherence rate for each drug class, and the vertical lines show the 90% confidence interval for each estimate. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. Appendix Table C.7 presents full estimation results.

4.2.3 Findings for PANO Programs

There were lower outpatient ER expenditures and increased adherence to diabetes medications among PANO program participants. Healthcare utilization findings for PANO program participants were not statistically significant, though many utilization outcomes had negative point estimates.

Outpatient ER expenditures decreased by about $126 per beneficiary among PANO participants, driven by decreases in the second six months post-intervention. Expenditures among participants dropped, whereas they increased for comparators (Table 4.7). While not statistically significant, other expenditure outcomes also had negative point estimates.

Among PANO program participants, there was an increase of 12 percentage points in the rate of highly adherent participants who take diabetes medications (Figure 4.2 and Appendix Table C.10). This estimate was driven by big improvements among program participants, and relative stability among the comparison group. Cumulative average adherence was not statistically significant (Table 4.8).

Table 4.6: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, PANO Programs

Measures	PANO		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	144/1,157	70/1,157	132/1,157
Difference-in-Difference	-58.08	-50.83	0.03
P-value	0.44	0.39	1.00
90% Confidence Interval	(-180.6, 64.4)	(-147.6, 45.9)	(-34.7, 34.8)
Baseline Participant Mean	440.72	359.47	70.83
Intervention Period Participant Mean	430.47	331.40	114.09
Baseline Comparison Mean	440.57	337.15	96.12
Intervention Period Comparison Mean	488.41	359.81	139.35
Relative Difference	-13.2%	-14.1%	0.0%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	84/583	31/583	51/583
Difference-in-Difference	12.28	-32.09	-3.66
P-value	0.78	0.30	0.88
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	61/574	39/574	81/574
Difference-in-Difference	-71.38*	-18.81	3.78
P-value	0.08	0.61	0.89

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income, except for the falls/fracture outcome due to small sample size. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Table 4.7: Medicare Expenditures per Beneficiary, PANO Programs

Measures (2011 USD)	PANO				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	1,121/1,157	70/1,157	144/1,157	749/1,157	1,144/1,157
Difference-in-Difference	- $564.84	- $205.09	- $125.98*	- $298.16	- $114.77
P-value	0.50	0.63	0.09	0.17	0.67
90% Confidence Interval	(-1,929.6, 799.9)	(-909.4, 499.2)	(-248.1, -3.9)	(-656.6, 60.3)	(-562.8, 333.3)
Baseline Participant Mean	$4,910.60	$1,944.78	$363.15	$689.05	$1,422.91
Intervention Period Participant Mean	$5,241.07	$1,866.02	$307.85	$717.20	$1,526.08
Baseline Comparison Mean	$4,998.61	$1,897.52	$313.07	$680.02	$1,342.80

Measures (2011 USD)	PANO				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Intervention Period Comparison Mean	$5,893.92	$2,023.85	$383.75	$1,006.34	$1,560.73
Relative Difference	-11.5%	-10.5%	-34.7%	-43.3%	-8.1%
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	559/583	31/583	84/583	379/583	555/583
Difference-in-Difference	- $235.04	- $135.45	- $55.94	- $111.62	- $31.58
P-value	0.63	0.59	0.29	0.39	0.82
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	562/574	39/574	60/574	370/574	559/574
Difference-in-Difference	- $332.39	- $70.93	- $70.12**	- $187.76	- $84.27
P-value	0.52	0.79	0.04	0.16	0.66

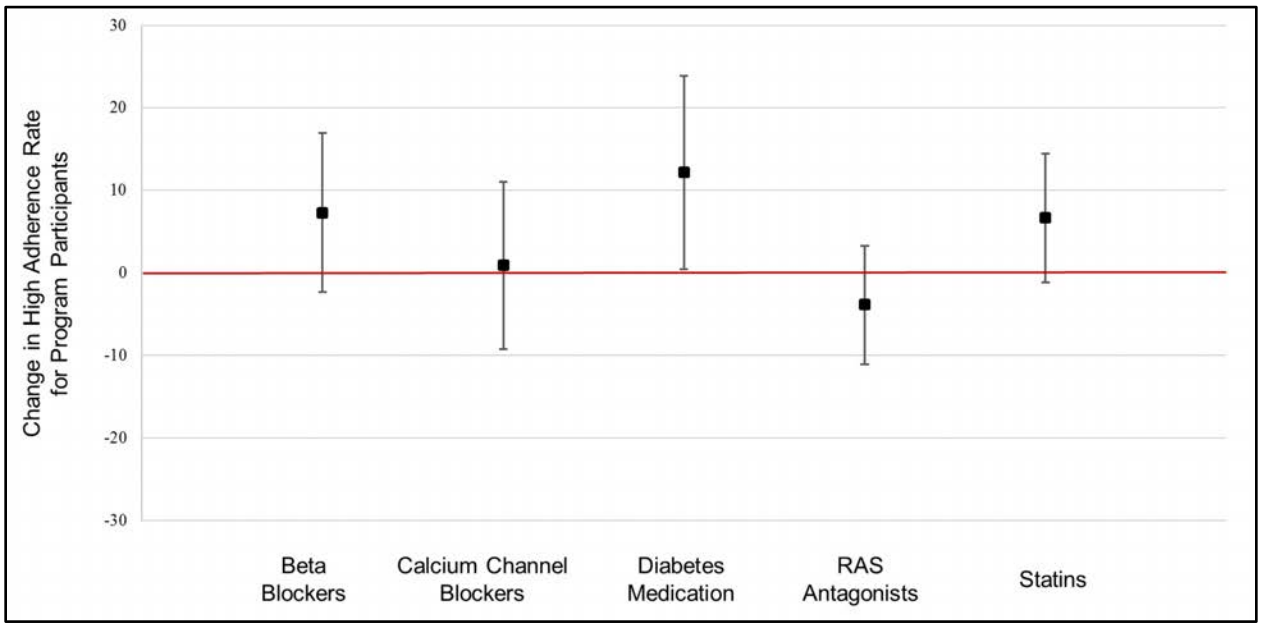
Notes*p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Table 4.8: Medication Adherence (Average Proportion of Days Covered), PANO Programs

Measures (Average PDC)	PANO				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	182	139	98	274	302
Difference-in-Difference	0.39	-1.29	3.72	-1.69	1.12
P-value	0.85	0.58	0.21	0.32	0.56
90% Confidence Interval	(-3.1, 3.9)	(-5.1, 2.5)	(-1.1, 8.6)	(-4.4, 1.1)	(-2.0, 4.2)
Baseline Participant Mean	86.11	88.99	84.11	86.98	87.09
Intervention Period Participant Mean	89.94	91.61	88.79	90.99	94.70
Baseline Comparison Mean	85.51	85.60	86.34	85.56	90.12
Intervention Period Comparison Mean	88.97	89.69	87.18	91.24	96.62
Relative Difference	0.5%	-1.4%	4.4%	-1.9%	1.3%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	88	65	51	134	143
Difference-in-Difference	-0.25	-4.13	1.01	-1.33	0.67
P-value	0.92	0.12	0.76	0.47	0.75
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	94	74	47	140	159
Difference-in-Difference	1.04	1.24	6.65**	-2.13	1.55
P-value	0.66	0.63	0.04	0.24	0.44

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Figure 4.2: High Adherence Rate (PDC ≥ 80%), Cumulative Estimates, PANO Programs



Notes: The y-axis represents the percentage point change in the rate of highly adherent beneficiaries in the year following program participation. The solid circle represents the estimated change in high adherence rate for each drug class, and the vertical lines show the 90% confidence interval for each estimate. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. Appendix Table C.10 presents full estimation results.

4.2.4 Findings for FP Programs

Among FP wellness program participants, there was increased adherence to diabetes medications. Point estimates for major utilization outcomes were negative, but there were no statistically significant findings. Home health expenditures decreased by about $173 per beneficiary among FP participants (see Appendix Table C.12).

There was an increase of 11 percentage points in the rate of highly adherent participants who take diabetes medications (Figure 4.3 and Appendix Table C.13), which, similarly to findings for PANO programs, is due to big improvements among participants, and relative

stability among the comparison group. There were no statistically significant cumulative average adherence findings for beneficiaries participating in FP programs (Table 4.11).

Table 4.9: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, FP Programs

Measures	FP		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	356/2,230	220/2,230	376/2,230
Difference-in-Difference	-28.94	-31.47	-19.30
P-value	0.58	0.40	0.30
90% Confidence Interval	(-113.9, 56.0)	(-92.7, 29.8)	(-50.1, 11.5)
Baseline Participant Mean	534.21	229.73	153.45
Intervention Period Participant Mean	569.27	273.51	168.61
Baseline Comparison Mean	513.76	220.85	150.73
Intervention Period Comparison Mean	577.75	296.02	185.19
Relative Difference	-5.4%	-13.7%	-12.6%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	169/1,133	106/1,133	160/1,133
Difference-in-Difference	-24.76	-18.00	-25.23
P-value	0.41	0.45	0.23
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	187/1,097	114/1,097	216/1,097
Difference-in-Difference	-4.05	-13.51	-13.18
P-value	0.91	0.56	0.56

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income, except for the falls/fracture outcome due to small sample size. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Table 4.10: Medicare Expenditures per Beneficiary, FP Programs

Measures (2011 USD)	FP				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	2,181/2,230	216/2,230	355/2,230	1,453/2,230	2,171/2,230
Difference-in-Difference	- $190.19	- $311.47	- $37.79	$12.03	$114.75
P-value	0.78	0.37	0.39	0.94	0.56
90% Confidence Interval	(-1,326.4, 946.1)	(-883.5, 260.6)	(-110.5, 34.9)	(-235.9, 260)	(-211.6, 441.1)

Measures (2011 USD)	FP				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Baseline Participant Mean	$6,975.88	$1,315.72	$365.11	$1,310.64	$2,647.24
Intervention Period Participant Mean	$7,953.60	$1,718.09	$377.11	$1,440.25	$2,757.41
Baseline Comparison Mean	$7,021.79	$1,265.70	$340.06	$1,303.44	$2,695.15
Intervention Period Comparison Mean	$8,189.70	$1,979.54	$389.84	$1,421.02	$2,690.57
Relative Difference	-2.7%	-23.7%	-10.4%	0.9%	4.3%
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	1,106/1,133	105/1,133	169/1,133	787/1,133	1,101/1,133
Difference-in-Difference	$205.74	- $99.83	- $37.60	$94.22	$31.86
P-value	0.64	0.63	0.21	0.30	0.79
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	1,075/1,097	101/1,097	176/1,097	713/1,097	1,070/1,097
Difference-in-Difference	- $406.77	- $214.66	$0.08	- $84.55	$83.73
P-value	0.34	0.37	1.00	0.39	0.46

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

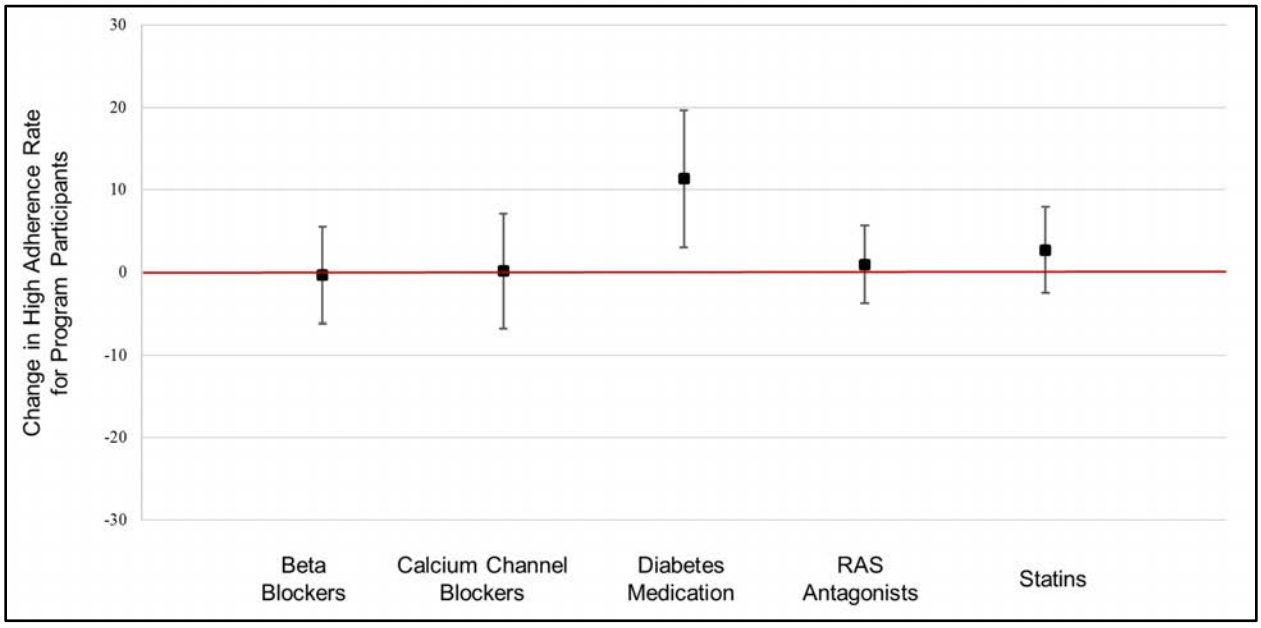
Table 4.11: Medication Adherence (Average Proportion of Days Covered), FP Programs

Measures (Average PDC)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	505	350	184	642	669
Difference-in-Difference	-1.06	-0.23	3.38	1.26	0.41
P-value	0.45	0.89	0.15	0.28	0.75
90% Confidence Interval	(-3.4, 1.3)	(-3.0, 2.5)	(-0.4, 7.2)	(-0.7, 3.2)	(-1.7, 2.5)
Baseline Participant Mean	89.32	88.74	85.56	87.35	85.09
Intervention Period Participant Mean	92.37	93.45	91.63	92.12	90.86
Baseline Comparison Mean	88.54	87.74	86.49	87.27	86.04
Intervention Period Comparison Mean	92.65	92.67	89.22	90.77	91.40
Relative Difference	-1.2%	-0.3%	4.0%	1.4%	0.5%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	254	180	95	329	347
Difference-in-Difference	-1.96	-0.56	4.38*	1.02	0.06
P-value	0.22	0.76	0.08	0.43	0.97
Interim Estimates: 7-12 Months					

Measures (Average PDC)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Total Participant Observations in the Post-Intervention Period	251	170	89	313	322
Difference-in-Difference	-0.21	0.08	2.14	1.48	0.73
P-value	0.90	0.97	0.41	0.25	0.61

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Figure 4.3: High Adherence Rate (PDC ≥ 80%), Cumulative Estimates, FP Programs



Notes: The y-axis represents the percentage point change in the rate of highly adherent beneficiaries in the year following program participation. The solid circle represents the estimated change in high adherence rate for each drug class, and the vertical lines show the 90% confidence interval for each estimate. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. Appendix Table C.13 presents full estimation results.

4.3 Discussion of Claims-Based Evaluation Findings

Increases in utilization among CDM participants suggest these beneficiaries may have increased interactions with their providers, consistent with CDM program goal. Decreases in expenditures among PANO and FP participants suggest lower intensity of healthcare utilization.

However, these findings, along with observed increases in medication adherence among CDM, PANO, and FP program participants to certain drug classes, should be interpreted with caution given the low analytic sample sizes.

The increase in ER visits among CDM program participants is unexpected. This finding may be a consequence of the small sample size and the fact that only a modest proportion of beneficiaries utilize ER services. Another possible explanation is related to the demographics of the CDM FFS population. Specifically, the CDM FFS matched participant population tends to have lower levels of income and education (Table 3.1) than the participant population in other priority areas. Low socio-economic status is associated with a preference for utilizing ER services for primary care needs, even after controlling for access to health insurance.[40] This is consistent with CDM programs' aim to improve health-related self-efficacy and communication with providers among participants. These findings are also consistent with other studies showing that when increases in healthcare utilization occur among beneficiaries with similar socio-economic characteristics to those of CDM participants, they affect multiple settings, including the ER.[41]

A decrease in outpatient ER expenditures among PANO program participants is consistent with lower intensity of healthcare utilization. Other expenditure outcomes and most utilization outcomes, including ER visits, also had negative point estimates, but were not statistically significant. The decrease in home health expenditures among FP program participants is driven by a drop in home health expenditures among participants and an increase among comparators. While not statistically significant, point estimates for outpatient ER and inpatient expenditures, ER visits, unplanned inpatient admissions, and falls/fractures outcomes were also negative, suggesting lower healthcare utilization among FP participants.

Participation in a wellness program is not generally associated with consistently significant effects on healthcare utilization, expenditures, and medication adherence. The lack of consistent claims-based findings may signify a lack of effect of the programs on Medicare utilization and spending, but it could also be due to the sample limitations discussed in Section 4.2.1, or to the fact that the post-intervention period was limited to only one year. In addition, the observational nature of this study implies that estimated effects may be biased due to unobserved differences between the treatment and the comparison groups. The analysis improves upon most

[40] Kangovi, S., F. K. Barg, T. Carter, J. A. Long, R. Shannon, and D. Grande. "Understanding Why Patients Of Low Socioeconomic Status Prefer Hospitals Over Ambulatory Care." *Health Affairs* 32, no. 7 (July 2013): 1196-203, doi:10.1377/hlthaff.2012.0825

[41] See, for example, Finkelstein, A. N., S.L. Taubman, H.L. Allen, B.J. Wright, and K. Baicker. "Effect of Medicaid Coverage on ED Use - Further Evidence from Oregon's Experiment." *The New England Journal of Medicine* 375, no. 16 (October 2016): 1505-1507).

other observational studies of wellness programs by explicitly taking into account selection into participation for the identification of comparison groups, but it is possible that the approach did not fully account for selection.[42]

[42] For more information regarding the identification of comparison groups for this evaluation, see: "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

5 CONCLUSION

The Wellness Prospective Evaluation sought to: (1) describe the overall distribution of readiness to engage with wellness programs in the Medicare population, (2) evaluate program impacts on health behaviors, self-reported health outcomes, and claims-based measures of utilization and costs, and (3) describe program operations and costs. This section presents a synthesis of our findings related to all three aims.

Readiness to Participate in Wellness Programs

Nearly a quarter (24%) of Medicare beneficiaries are "ready" to participate in a wellness program, based on a composite readiness index developed from a nationally representative survey of Medicare beneficiaries.[43] Beneficiaries were more likely to be ready to participate in wellness programs if they:

- Were younger (66-74 years), female, or non-white;
- Were aware of wellness programs in the community or online;
- Participated in a wellness program in the past two years;
- Had high self-efficacy or patient activation;
- Received a physician recommendation to participate;
- Reported having a higher body mass index (BMI); a chronic condition, such as arthritis, diabetes, or pre-diabetes; or more physical or mental limitations.

Those without a high school degree and those with transportation difficulty had lower levels of readiness. Interestingly, social support was slightly lower among those who were ready to participate in a wellness program. This finding may signal additional social needs among those ready for behavior change and program participation, or it may suggest that those with more social support had their wellness needs met outside the context of a wellness program.

The strong effects on readiness of program awareness, prior participation, and physician recommendations suggest that demand for wellness programs could increase if promotion efforts for such programs in the community and among healthcare providers also increased.

Program Impacts

The following themes emerged from the analysis of the effects of wellness program participation on self-reported and claims-based outcomes:

[43] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

- PANO and FP programs showed consistently beneficial effects on many of the self-reported health and behavioral measures, but in most cases the effects were small. CDM programs showed a benefit for only confidence in balance.
- The pattern of effects was generally consistent with the focus areas and design of the programs. For example, PANO programs improved participants' self-reported strength and physical activity levels, and FP programs improved multiple measures of physical health and body strength.
- The most consistent benefits for PANO and FP programs were on wellbeing, as measured by the mental health subscales. Although the programs did not specifically target mental health, the results suggest that program participation, lifestyle changes, and increased physical activity have benefits in this area.
- Many program effects persisted at 12 months beyond baseline, which is encouraging, since most wellness program were only 6-8 weeks long.[44] Duration of program participation may have contributed to benefits detected in PANO programs at 12 months, although benefits also persisted in the shorter-term FP and CDM programs.
- The results suggest that wellness programs have protective effects against deterioration in health, mental health, and activity levels that may occur over time due to aging. For most measures and ACA priority areas, favorable results occurred due to a decline within the comparison group, rather than statistically significant gains among participants.
- Outpatient ER expenditures decreased among PANO participants and home health expenditures decreased among FP participants. Utilization (ER visits) actually increased among CDM program participants.[45]There is no evidence of program effects on healthcare utilization among PANO or FP participants, or on expenditures among CDM participants.
- Average adherence among CDM program participants only improved for calcium channel blockers, whereas the proportion of beneficiaries who are highly adherent to diabetes medication increased among PANO and FP participants. However, adherence estimates are based on very small sample sizes, and should be interpreted with caution.

In general, there were few strong benefits of program participation for self-reported health, wellbeing, and health behaviors. This general conclusion is consistent with several explanations, the simplest of which is that wellness programs do not strongly benefit Medicare beneficiaries in a way that would be reflected in the self-reported outcomes included in the surveys. Another explanation is that the observation period is too short to capture stronger effects on health and health behaviors. Notably, some program effects (e.g., confidence in balance) observed at six months post-intervention increased at twelve months. If it takes time for behavioral changes to translate into improved health and wellbeing, then the one-year follow-up period may be too short to observe large benefits of program participation.

[44] About half of PANO participants were engaged in an ongoing wellness program.
[45] Participants are limited to those enrolled in FFS.

Findings for cost and utilization outcomes do not offer a consistent conclusion. The increase in ER visits observed among CDM program participants may be related to a higher inclination to seek medical help in the post-intervention period. However, findings for CDM program participants should be interpreted with caution, given the low sample sizes and the possibility that they represent false positives. The lack of consistent empirical findings for PANO and FP programs may indicate no significant impact of wellness programs on Medicare utilization and spending, but it may also be due to the low sample sizes available for analysis. Many outcomes have a very low number of non-zero observations driving the estimates. The lack of consistent findings may also be due to the short post-evaluation observation period. It is possible that a longer time horizon would have permitted identification of an impact of wellness programs on healthcare utilization and expenditures.

Qualitative Study of Program Operations and Costs

In 2015, the Acumen team conducted site visits to ten wellness programs, to get information on wellness program operations and costs.[46] The following best practices, challenges, and lessons learned were identified:

- Large and multi-site coordinators have centralized portions of workforce management, marketing, fidelity monitoring, and data reporting to create operational efficiencies.
- Organizations have leveraged partnerships with local health systems and universities to recruit leaders and guest experts.
- In-person and word-of-mouth marketing strategies are most effective for recruiting Medicare beneficiaries.
- Transportation services or translators are used to engage rural or refugee populations, who are harder to reach.
- The majority of organizations are able to conduct simple analyses of program data, but only a few organizations maintain robust data collection and cost reporting systems.

Across wellness programs, operational costs for program delivery ranged from $100 to $500 per participant.[47] These estimates may be lower than the amount needed to sustain or scale up wellness program delivery, because they do not include facility costs for class locations, and because many programs rely heavily on volunteer labor.

Most organizations rely on grant- or contract-based funding to support wellness program operations, and reported that financial sustainability was an ongoing challenge, emphasizing the

[46] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[47] The analysis of operational costs included large organizations with mature wellness program operations, and thus the generalizability of these findings is limited.

need for reliable and sustainable funding streams. Several respondents noted that the unreliability of funding streams hinders success, creates "fear that programs will go away," and presents a challenge to the scale up and spread of wellness programs, because organizations are reluctant to create infrastructure and expand the workforce.

Comparison with Prior Studies

The prospective evaluation presented in this report differs from prior studies along the following dimensions: (i) research setting and research design; (ii) study population; (iii) source of the data analyzed; and (iv) duration of the follow-up period. Similarities and differences along these dimensions may explain differences in findings between the prospective evaluation and prior studies. The prospective evaluation is based on an observational, "real-world" study, which takes into account selection into wellness programs, whereas previous studies either take place in a randomized-controlled setting, or do not account for selective program participation.[48] In addition, the claims-based analysis focuses on Medicare FFS beneficiaries, while many prior studies focus on populations in managed care, who may have different demographic and health characteristics.[49] Finally, the prospective evaluation relied on both CMS administrative data sources and self-reported data, observed over a one-year follow-up, while prior studies used different data sources (e.g., self-reported data only, data from managed care organizations) and post-intervention period durations (e.g., 3 years).

Most peer-reviewed studies of wellness programs have found positive effects of program participation on physical and mental health and health behaviors.[50,51] Similar to the results of this prospective evaluation, a handful of studies have found only modest or no effects of CDM programs on health behaviors, physical and mental health, and health status.[52,53] The prospective evaluation findings for PANO and FP programs are also consistent with prior literature pointing

[48] See, for example: Brady, Teresa J., et al. "A Meta-Analysis of Health Status, Health Behaviors, and Health Care Utilization Outcomes of the Chronic Disease Self-Management Program." *Preventing Chronic Disease* 10 (January 2013); and Alva, Maria L., et al. "Impact of The YMCA of the USA Diabetes Prevention Program on Medicare Spending and Utilization." *Health Affairs* 36, no. 3 (March 2017): 417-424.
[49] See, for example: Lorig, Kate R., et al. "Effect of a Self-Management Program on Patients with Chronic Disease." *Effective Clinical Practice* 4, no. 6 (November-December 2001): 256-262.; and Ackermann, Ronald T., et al. "Healthcare Cost Differences with Participation in a Community-Based Group Physical Activity Benefit for Medicare Managed Care Health Plan Members." *Journal of the American Geriatrics Society* 56, no. 8 (August 2008): 1459-1465.
[50] Ory et al. "Successes of a National Study of the Chronic Disease Self-Management Program: Meeting the Triple Aim of Health Care Reform." *Medical Care* 51, no. 11 (November 2013): 992-998.
[51] Lorig et al. "A Diabetes Self-Management Program: 12-Month Outcome Sustainability from a Nonreinforced Pragmatic Trial." *Journal of Medical Internet Research* 18, no. 12 (December 2016): e322.
[52] Ersek M, Turner JA, Cain KC, Kemp CA. "Results of a Randomized Controlled Trial to Examine the Efficacy of a Chronic Pain Self-Management Group for Older Adults." *Pain* 138 no. 1 (August 2008): 29-40.
[53] Haas M, Groupp E, Kraemer D, Brummel-Smith K, Sharma R, Granger B, Attwood M, Fairweather A. "Chronic Disease Self-Management Program for Low-Back Pain in the Elderly." *Journal of Manipulative and Physiological Therapeutics* 28, no. 4 (May 2005): 228-237.

to positive program effects on anxiety and depression, activity levels, and confidence in balance.[54,55,56]

A prior, retrospective study of community-based wellness and prevention programs ("retrospective evaluation") most closely resembles the analysis of utilization and expenditure outcomes included in the prospective evaluation presented in this report.[57] The retrospective evaluation found no evidence of increased total Parts A and B expenditures for participants of CDSMP, though it found increases in both outpatient ER expenditures and outpatient ER visits. In addition, and contrary to the findings outlined in this report, the retrospective evaluation found that EnhanceFitness and Matter of Balance programs decreased total Parts A and B expenditures and unplanned admissions.

The differences in findings between the retrospective and the prospective evaluation may be due to differences in the research design of the two studies. Specifically, the matching algorithm for this prospective evaluation took into account selection into the wellness programs, as well as many other socio-economic variables included in the beneficiary survey. Controlling for participation selection and including more socio-economic variables in the matching makes the conclusions of the prospective evaluation more robust to bias from unobserved differences between program participants and the comparison group, which could lead to different trends in healthcare utilization and expenditures. For example, if beneficiaries willing to participate in PANO programs are wealthier and better educated than the average Medicare beneficiary, then ignoring these aspects of program participation may result in a comparison group with faster-increasing healthcare utilization and expenditures than PANO program participants, biasing the DiD estimates.

The differences in findings between the prospective and the retrospective evaluation may also be due to the smaller sample sizes available for the prospective evaluation. For example, point estimates of the effect of PANO and FP programs on unplanned admissions are negative, but not statistically significant (see Appendix C.3). This finding may be due to the low statistical power of the claims-based analysis in the prospective evaluation, which limits the ability to

[54] Hughes SL, Seymour RB, Campbell RT, Desai P, Huber G, Chang HJ. "Fit and Strong!: Bolstering Maintenance of Physical Activity Among Older Adults With Lower-extremity Osteoarthritis." *American Journal of Health Behavior* 34, no. 6 (November-December 2010): 750-763.
[55] Hughes SL, Seymour RB, Campbell RT, et al. "Long-Term Impact of Fit and Strong! on Older Adults With Osteoarthritis." *The Gerontologist* 46, no. 6 (December 2006): 801-814.
[56] Alexander et al. "Effect of the Matter of Balance Program on Balance Confidence in Older Adults." *The Journal of Gerontopsychology and Geriatric Psychiatry* 28, no. 4 (2015): 183-9.
[57] "Report to Congress: The Centers for Medicare & Medicaid Services' Evaluation of Community-based Wellness and Prevention Programs under Section 4202(b) of the Affordable Care Act." Centers for Medicare & Medicaid Services (CMS). Available at: https://innovation.cms.gov/Files/reports/CommunityWellnessRTC.pdf.

detect statistically significant effects of wellness programs on utilization and expenditure outcomes.

Increases in ER utilization for CDM participants, found in this prospective evaluation, also differ from other peer-reviewed literature, which found either no effect[58,59,60] or a reduction in utilization and expenditures.[61,62] Differences in the age of the study population may help explain these differences. Liddy et al. (2015) found that, in general, there were no significant changes in ER use, physician visits, or hospitalizations among adults aged 19 years or older who participated in a Canadian CDSMP program.[63] However, among older adults (66 years and older), widows, and those who were more severely ill (measured by number of chronic conditions), ER and physician visits increased following participation in the CDSMP program. In contrast to the population in the prospective evaluation, which included Medicare beneficiaries over the age of 65, most of the studies focused on CDM programs included a younger population. In fact, a review summarizing the literature on the impact of the Stanford CDSMP noted that, of the 25 quantitative articles reviewed, only 8 included a population with mean age greater than 64 years.[64] Studies of CDM programs that predominantly included the elderly have, similarly to this prospective evaluation, tended to find only a modest or no significant effect on pain improvement, general health, self-efficacy, and self-care.[65,66]

The prospective evaluation studied wellness programs administered in diverse delivery settings across the country, while some prior work has looked at effects of wellness programs implemented in more carefully controlled, academic environments. This difference may

[58] Gitlin LN, Chernett NL, Harris LF, Palmer D, Hopkins P, Dennis MP. "Harvest Health: Translation of the Chronic Disease Self-Management Program for Older African Americans in a Senior Setting." *Gerontologist* 48, no. 5 (July 2014): 698-705.
[59] Goeppinger J, Armstrong B, Schwartz T, Brady T. "Self-Management Education for Persons with Arthritis: Managing Comorbidity and Eliminating Health Disparities." *Arthritis and Rheumatism* 57, no. 6 (August 2007): 1081-1088.
[60] Jerant A, Moore-Hill M, Franks P. "Home-based Peer Led Chronic Illness Self-Management Training: Findings from a 1-year Randomized Controlled Trial." *Annals of Family Medicine* 7, no. 4 (July 2009): 319-327.
[61] Lorig et al. "Chronic Disease Self-Management Program: 2-year Health Status and Health Care Utilization Outcomes." *Medical Care* 39, no. 11 (November 2001): 1217-23.
[62] Ory et al. "Successes of a National Study of the Chronic Disease Self-Management Program: Meeting the Triple Aim of Health Care Reform." *Medical Care* 51, no. 11 (November 2013): 992-998.
[63] Liddy C, Johnston S, Guilcher S, Irving H, Hogel M, Jaglal S. "Impact of a chronic disease self-management program on healthcare utilization in eastern Ontario, Canada." *Preventive Medicine Reports* 2, (July 2015): 586-590.
[64] "Evaluation Design for the Chronic Disease Self-Management Program Implement in AoA funded Settings." Agency for Healthcare Research and Quality (AHRQ): IMPAQ International, LLC and Abt Associates, Inc. February 2011. Available at: https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/final-reports/aoa/aoachronic-apb.pdf.
[65] Ersek M, Turner JA, Cain KC, Kemp CA. "Results of a Randomized Controlled Trial to Examine the Efficacy of a Chronic Pain Self-Management Group for Older Adults." *Pain* 138 no. 1 (August 2008): 29-40.
[66] Haas M, Groupp E, Kraemer D, Brummel-Smith K, Sharma R, Granger B, Attwood M, Fairweather A. "Chronic Disease Self-Management Program for Low-Back Pain in the Elderly." *Journal of Manipulative and Physiological Therapeutics* 28, no. 4 (May 2005): 228-237.

contribute to more variation in program implementation across delivery settings. However, qualitative findings from this prospective evaluation's site visits suggest a high degree of fidelity to the original national program, and materials prepared by national program sponsors were frequently employed in the local setting.

Scalability Assessment

The evaluation team's findings suggest that nearly one out of four beneficiaries is ready to participate in a community wellness program, indicating that there is demand for such programs among the Medicare population. The findings of this evaluation also show that wellness programs, particularly those promoting physical activity, offer physical and mental health benefits to their participants by preventing age-related deterioration. Most organizations that currently deliver wellness programs, however, have a limited number of staff and resources. Large organizations play an important role in the scalability of such programs, because they may offer efficiencies and support high-quality program delivery. The scalability of wellness programs is also dependent on the stability of public and other funding streams, which would enable increased outreach and education to trusted providers and community organizations, particularly in rural areas and among lower-resourced organizations.

Conclusion

While evaluation findings do not provide conclusive evidence that evidence-based wellness programs have a significant impact on utilization and expenditures, self-reported outcomes related to beneficiary physical and mental health modestly improved between baseline and 12 months. These two sets of results are not inconsistent: Self-reported benefits related to mental health and wellbeing may not necessarily result in impacts on health care utilization or costs. In addition, the improvement in self-reported health may have more sustainable impacts on costs and utilization over a longer post-intervention observation period. Currently, there is no consistent evidence of cost savings. However, the observed protective effects of wellness programs, particularly those focused on PANO and FP, on physical and mental health, physical activity, body strength, and confidence in balance may pay dividends in the future. Further studies are needed to explore whether a longer follow-up period or a larger sample size yield more promising effects on expenditure and utilization outcomes.

APPENDIX A – ANALYTIC METHODOLOGY DETAILS

This section presents analytic methodology details. Section A.1 and Section A.2 present claims-based and survey-based outcome measures specifications, respectively. Section A.3 summarizes the differences-in-differences methodology. Section A.4 describes the structure of the claims data used in the claims-based analysis on healthcare utilization, expenditure, and medication adherence. Section A.5 discusses matching criteria and the timing of twelve-month survey fielding in the survey-based analysis on self-reported health and health behaviors. Finally, Section A.6 describes the twelve-month survey weighting methodology.

A.1 Claims-Based Outcome Measure Specifications

The tables below define the claims-based outcome measures presented for the Wellness Prospective Evaluation Final Report. Appendix Table A.1 provides definitions of key terms used in the outcome measure definitions, and Appendix Table A.2 provides definitions of the outcome measures themselves.

Appendix Table A.1: Definitions of Terms Used in Outcome Measure Definitions

Term	Definition
Expenditure	All expenditure measures represent Medicare payments. Cost data for all Parts A and B expenditure measures are standardized using the CMS payment standardization methodology to remove differences due to geographic variation in Medicare payment rates and variation among classes of providers.[67] Parts A and B costs are also adjusted monthly for inflation (2011 base year) using the Bureau of Labor Statistics Consumer Price Index for medical care services. Cost data are not risk-adjusted.
Beneficiary	Beneficiaries must be continuously enrolled in Medicare Parts A and B Fee For Service for one year prior to the program's intervention date through the intervention period of interest. Beneficiaries who switch between FFS and MA are not included in the analysis. If a beneficiary dies, the beneficiary will be included in the six-month period in which he or she died and not in any subsequent six-month periods.
Proportion of Days Covered (PDC)	PDC was calculated by examining Part D claims for each medication in question to determine the proportion of days during the 12 month period when an individual possessed any of the specified medications. For inclusion in either the numerator or denominator, patients required at least two prescriptions and 91 total days of prescriptions.[68]

Appendix Table A.2: Definitions of Claims-based Outcome Measures

Measure	Definition
ER Visits	ER Visit Rate per 1,000 beneficiaries Numerator: Number of beneficiaries with at least one outpatient ER claim or observational stay with no inpatient admission on the same day * 1,000. Denominator: Total number of beneficiaries.
Inpatient Admissions	Inpatient Admission Rate per 1,000 beneficiaries. Numerator: Number of beneficiaries with at least one inpatient stay * 1,000. Denominator: Total number of beneficiaries.

[67] More information about expenditure standardization methodology is available in CMS Standardization Methodology For Allowed Amount (CMS), available at http://www.qualitynet.org/

[68] More information about adherence is available in PQA Performance Measures. Available at http://pqaalliance.org/

Measure	Definition
Unplanned Inpatient Admissions	Unplanned admission rate per 1,000 beneficiaries. Unplanned stays do not include stays that are planned or potentially planned stays without acute care. Numerator: Number of beneficiaries with at least one unplanned inpatient stay * 1,000. Denominator: Total number of beneficiaries.
Length of Stay	Average number of hospital days per 1,000 beneficiaries. Numerator: Total number of inpatient days * 1,000. Denominator: Total number of beneficiaries.
Falls/Fractures	Proportion of beneficiaries (per 1,000) with at least one fall- or fracture-related claim within the intervention period. Numerator: Number of beneficiaries who had at least one falls-related claim * 1,000 Denominator: Total number of beneficiaries
Total Part D Expenditures (2011 USD)	Total Part D Expenditures per beneficiary. Numerator: Total Part D claims costs Denominator: Total number of Beneficiaries
Total Parts A and B Expenditures (2011 USD)	Total Parts A and B Expenditures per beneficiary. Numerator: Total Parts A and B claims costs. Denominator: Total number of Beneficiaries
Inpatient Expenditures (2011 USD)	Inpatient Expenditures per beneficiary. Numerator: Total inpatient stay costs. Denominator: Total number of beneficiaries.
Outpatient ER Expenditures (2011 USD)	Outpatient ER Expenditures per beneficiary. Numerator: Total emergency room (ER)-only outpatient claim or observational stay claim (without an inpatient admission claim) costs. Denominator: Total number of beneficiaries.
Outpatient Non-ER Expenditures (2011 USD)	Outpatient Non-ER Expenditures per beneficiary. Numerator: Total non-ER outpatient claim costs. Denominator: Total number of beneficiaries.
Physician and Ancillary Expenditures (2011 USD)	Physician and Ancillary Expenditures per beneficiary. Numerator: Total Carrier/PB claim costs. Denominator: Total number of beneficiaries.
Durable Medical Equipment Expenditures (2011 USD)	Durable Medical Equipment Expenditures per beneficiary. Numerator: Total durable medical equipment claims costs. Denominator: Total number of beneficiaries.
Home Health Expenditures (2011 USD)	Home Health Expenditures per beneficiary. Numerator: Total home health claim costs. Denominator: Total number of beneficiaries.
Adherence to Beta Blockers (Average PDC)	Average PDC during the intervention period for beneficiaries taking at least one Beta Blocker.
Adherence to Calcium Channel Blockers (Average PDC)	Average PDC during the intervention period for beneficiaries taking at least one Calcium Blocker.
Adherence to Diabetes Medication (Average PDC)	Average PDC during the intervention period for beneficiaries taking at least one Diabetes Medication. Insulin users and ESRD beneficiaries are excluded from this calculation.
Adherence to RAS Antagonists (Average PDC)	Average PDC during the intervention period for beneficiaries taking at least one RAS Antagonist. ESRD beneficiaries are excluded from this calculation.
Adherence to Statins (Average PDC)	Average PDC during the intervention period for beneficiaries taking at least one Statin.
Adherence to Beta Blockers (PDC ≥ 80%)	Proportion of beneficiaries with PDC of at least 80% for a Beta Blocker during the intervention period.

Measure	Definition
Adherence to Calcium Channel Blockers (PDC ≥ 80%)	Proportion of beneficiaries with PDC of at least 80% for a Calcium Channel Blocker during the intervention period.
Adherence to Diabetes Medication (PDC ≥ 80%)	Proportion of beneficiaries with PDC of at least 80% for a Diabetes Medication during the intervention period. ESRD beneficiaries and insulin users were excluded from this calculation.
Adherence to RAS Antagonists (PDC ≥ 80%)	Proportion of beneficiaries with PDC of at least 80% for a RAS Antagonist during the intervention period. ESRD beneficiaries were excluded from this calculation.
Adherence to Statins (PDC ≥ 80%)	Proportion of beneficiaries with PDC of at least 80% for a Statin during the intervention period.

A.2 Survey-Based Outcome Measure Specifications

Appendix Table A.3 below defines the survey-based outcome measures presented for the Wellness Prospective Evaluation Final Report. All data were sourced from the National and Participant Surveys.

Appendix Table A.3: Definitions of Survey-based Outcome Measures

Measure	Participant Survey Item Numbering	Specification	Missing Data Rules
SF-36v2 Health Survey			
Physical Components Summary Score	q3a -q3j, q7, q8, q4a – q4d, q1, q11a – q11d	Produced through QualityMetric proprietary algorithm as a latent variable.	The Physical Component Summary (PCS) score can be calculated when seven physical health items are available and the Physical Functioning (PF) scale is not missing.
Physical Functioning Subscale	q3a -q3j	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Bodily Pain Subscale	q7, q8	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Role Physical Subscale	q4a – q4d	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
General Health Subscale	q1, q11a – q11d	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Mental Components Summary Score	q9a, q9e, q9g, q9i, q6, q10, q5a – q5c, q9b, q9c, q9d, q9f, q9h	Produced through QualityMetric proprietary algorithm as a latent variable.	The Mental Component Summary (MCS) score can be calculated when at least seven mental health items are available and the Mental Health (MH) scale is not missing.
Vitality Subscale	q9a, q9e, q9g, q9i	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Social Functioning Subscale	q6, q10	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.

Measure	Participant Survey Item Numbering	Specification	Missing Data Rules
Role Emotional Subscale	q5a – q5c	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Mental Health Subscale	q9b, q9c, q9d, q9f, q9h	Items are averaged and transformed to have a mean of 50	QualityMetric proprietary algorithm can score if at least one item is answered.
Rapid Assessment of Physical Activity			
Aerobic	q12a – q12g	Score as sedentary If "yes" to 12a Score as under-active if "yes" to 12b Score as under-active regular – light activities if "yes" to 12c Score as under-active regular if "yes" to 12d or 12e Score as active if "yes" to 12f or 12f	No treatment of missing data. Highest "yes" value is selected as the scale score.
Strength/Flexibility	q12h, q12i	RAPA_STRFLEX=1 if and only if "yes" to 12h. RAPA_STRFLEX=2 if and only if "yes" to 12i. RAPA_STRFLEX=3 if "yes" to BOTH RAPA_STRFLEX=0 if "no" to both.	Respondent must answer both items to score this measure.
Falls and Balance			
Falls in Past Six Months	q13	Yes/No	NA: single item.
Confidence in Balance (ABC) Scale	q17a-q17f	Average of valid answers where each is scored from 0% confidence to 100% confidence.	75% of items must be answered to score this scale.
Medication Adherence			
MAQ-4	q18-q21	Total of "no" responses is the scale score.	75% of items must be answered to score this scale.

A.3 Differences-in-Differences Methodology

The general DiD model can be illustrated as follows:

$$outcome_{it} = \beta_0 + \beta_1 \cdot program_i + \beta_2 \cdot post_t + \beta_3 \cdot (programXpost)_{it} + \beta_4 \cdot X_{it} + u_{it}$$

In the equation above, $outcome_{it}$ is the survey- or claims-based measure of interest for beneficiary i at time period t. $Program_i$ is an indicator variable equal to 1 if the observation refers to a program participant, and 0 otherwise. $Post_t$ is an indicator variable equal to 1 if the observation refers to the post-intervention period.[69] The interaction term $(programXpost)_{it}$ is an indicator variable equal to 1 if the observation refers to a program participant during the post-intervention period. X_{it} represents a vector (or set) of control variables representing the following survey- and claims-based demographic variables: urban/rural status, dual eligibility status,

[69] The post-intervention period is defined as the period following initial attendance date (for program participants) or baseline survey response date (for national survey respondents).

gender, race, age, education, and income.[70] The variable u_{it} is the error term. The coefficient of interest, which estimates the effect of program participation on the outcome of interest, is β_3.

A.4 Claims-Based Analysis: Structure of Claims Data

Each observation in the claims-based analysis corresponds to a beneficiary-six-month period. For example, if the outcome of interest is total medical expenditures, each observation corresponds to total medical costs incurred by a beneficiary in the sample over six months. Baseline observations for utilization and expenditure outcomes are generated by adding up the total number of events or expenditures incurred over the 12-month baseline period by each beneficiary, and dividing the sum by two. The incidence of falls and fractures, and adherence outcomes are not summed or divided in this way, since they correspond to averages over a given time period.

Interim analyses compare outcomes at baseline (expressed on a half-year basis) to outcomes during the first and second six-month period following program participation (or survey receipt). The cumulative analysis model specification also uses beneficiary-six-month observations, and introduces an extra time indicator variable in the main DiD model, along with its interaction with the program participation variable.

The cumulative analysis model is thus:

$$outcome_{it} = \beta_0 + \beta_1 \cdot program_i + \beta_2 \cdot (time = 1) + \beta_3 \cdot (time = 2) + \beta_4 \cdot [programX(time = 1)]_{it} + \beta_5 \cdot [programX(time = 2)]_{it} + \beta_6 \cdot X_{it} + u_{it}$$

In the above model, the variable (*time=1*) is an indicator variable for the first six months of the post-intervention period, whereas the variable (*time=2*) is an indicator variable for the second six months of the post-intervention period. For outcomes that correspond to counts (e.g. ER visits, IP admissions, total expenditures), the cumulative DiD estimate is the weighted sum of coefficients β_4 and β_5, weighted by the number of observations in the first and second half of the post-intervention period respectively. For outcomes that correspond to proportions (e.g. adherence outcomes, incidence of falls and fractures), the cumulative DiD is the weighted average of coefficients β_4 and β_5, weighted by the number of observations in the first and second half of the post-intervention period respectively.

Utilization outcomes, as well as the incidence of falls and fractures, are reported on a per 1,000 beneficiary basis, whereas expenditure and adherence outcomes are reported on a per

[70] All survey-based and most claims-based models have been estimated with and without covariates, and DiD estimates are very similar across the two model specifications. Some claims-based outcomes (incidence of falls and fractures, and home health expenditure) could not be reliably estimated with covariates, due to low sample size and number of beneficiaries with nonzero observations.

beneficiary basis. Adjusted means for the interim analyses correspond to half-year averages, whereas cumulative analysis means are yearly averages.

A.5 Survey-Based Analysis: Matching Criteria and Twelve-Month Survey Fielding

Twelve-month surveys were sent late to approximately 700 matched comparators who were successfully matched to program participants based on their willingness to make lifestyle changes (rather than their willingness to enroll in wellness programs). Prior to matching, it was assumed that all matched national respondents would be selected based on their willingness to participate in wellness programs, but this criterion did not produce enough potential matches for the study. As a result, beneficiaries who were willing to make lifestyle changes were also added to the pool of potential matches. Since twelve-month surveys were fielded prior to finalizing the matching process, the initial fielding focused on beneficiaries who were willing to participate in wellness programs, but did take into account beneficiaries who were willing to make lifestyle changes. As a result, some of the matched comparison sample did not receive their surveys on schedule, and responded to the twelve-month surveys substantially later than other respondents. Analysis of self-reported outcomes indicated that delayed response was not significantly related to most outcomes among all twelve-month respondents. To protect against possible impacts of delayed response, the number of months each respondent's survey fielding was delayed by was controlled for in the regression analyses, as an additional covariate in all DiD models using the twelve-month survey sample.

A.6 Survey-Based Analysis: Twelve-Month Survey Weighting Methodology

This section describes the weighting process for the matched sample respondents at twelve months. The overall goal of the twelve month survey weighting was to re-balance the matched samples after attrition at six and twelve months so that they better reflect the size and characteristics of the full matched samples. A similar process was undertaken to re-balance the six-month survey data.[71] The twelve-month weighting uses six-month weights as a starting point, and further adjusts them to reflect nonresponse in the twelve month survey among the six month survey respondents. Jackknife variance estimation was used based on a set of replicate weights, a common resampling procedure for complex survey designs.[72]

[71] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

[72] Wolter, K.M. *Introduction to Variance Estimation*. Springer: New York, 2007.

For weighting purposes at six and twelve months, each matched sample was treated as a census at baseline. Both the weighting and analytic strategies treat the matched samples as having independent national and participant components as opposed to sets of two matched individuals. This allows us to preserve sample size when only one individual in a matched pair responds. Nonresponse weighting adjustment was carried out within each of the six resulting samples, three participant sample and three comparator samples.

Imputation of Unknown Eligibility

Nonrespondents who were discovered to be deceased, institutionalized, or have speech/language issues during the course of survey fielding were coded as ineligible for the survey. For the majority of nonrespondents, however, it is not possible to directly determine eligibility status. Accounting for ineligibility is an important part of the weighting process, since the nonresponse weighting adjustment is done for only the eligible sample. Therefore, the first step in refining the six-month weights for nonresponse at twelve months was imputing eligibility status for nonrespondents for whom survey eligibility is unknown.[73] The missing eligibility status for eligibility unknown cases was imputed using the tree-building software GUIDE (Generalized, Unbiased, Interaction Detection, and Estimation).[74]

GUIDE is a tree algorithm that builds a classification or regression tree. As an option, it also produces a classification or regression forest. The GUIDE classification forest was used with the six-month survey data as auxiliary variables to impute the eligibility status for eligibility unknown cases in the twelve-month survey. The GUIDE classification forest produced an estimated probability that a sample unit is eligible – the forest works better than the tree for imputation.[75] Appendix Table A.4 presents the imputation result.

Appendix Table A.4: Original and Imputed Eligibility for the Twelve-Month Survey

Survey Type	Original Eligibility	Imputed Eligibility	Frequency	Percent
Participant	Unknown Eligibility	Ineligible	13	0.5
	Unknown Eligibility	Eligible	231	8.0
	Ineligible	Ineligible	15	0.5
	Eligible	Eligible	2,617	91.0
	Total		2,876	100.0

[73] There are other ways to handle unknown eligibility such as estimating the ineligibility rate using the rate among the known cases, which is often used. However, the imputation approach, when there are rich auxiliary data available (as in our case), is better in dealing with the unknown eligibility issue for nonresponse adjustment.
[74] Loh, W. Y. (2002). Regression Trees with Unbiased Variable Selection and Interaction Detection. *Statistica Sinica*, 12, 361–386. Loh, W. Y. (2009). Improving the precision of classification trees. *Annals of Applied Statistics*, 3, 1710–1737.
[75] Lee, H., and Jeong, D. "Missing data imputation using regression and classification tree software GUIDE." *Proceedings of the Survey Research Methods Section, American Statistical Association*. (Forthcoming).

Survey Type	Original Eligibility	Imputed Eligibility	Frequency	Percent
National	Unknown Eligibility	Ineligible	42	1.7
	Unknown Eligibility	Eligible	351	14.2
	Ineligible	Ineligible	14	0.6
	Eligible	Eligible	2,070	83.6
	Total		2,477	100.0
Combined	Unknown Eligibility	Ineligible	55	1.0
	Unknown Eligibility	Eligible	582	10.9
	Ineligible	Ineligible	29	0.5
	Eligible	Eligible	4,687	87.6
	Total		5,353	100.0

There were 244 nonrespondents with unknown eligibility (8.5 percent) among participant survey invitees, of which 231 were imputed to be eligible. The original eligibility rate among the eligibility known cases is 99.4 percent, which becomes slightly reduced to 99.0 percent after imputation. The eligibility rate for the national survey was lower; 99.3 percent in the original and 97.7 percent after imputation.

Response Rates

Frequency distributions of the twelve-month survey samples by response status and the corresponding response rates (based on imputed eligibility status) are presented in Appendix Table A.5. While the overall completion rate at twelve months for the full matched samples, shown in Section 2.4, is closer to 60 percent, the twelve-month survey response rate is higher than 80 percent for all samples because the starting sample included only those matched sample members who responded at six months.

Appendix Table A.5: Twelve-Month Survey Samples by Response Status and Rates

Response Type	PANO Programs				CDM Programs				FP Programs			
	Participant		National		Participant		National		Participant		National	
	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%	Freq	%
Response	656	85.9	693	81.5	529	82.5	585	79.7	1,252	85.1	1,339	82.3
Nonresponse	102	13.4	143	16.8	108	16.9	129	17.6	201	13.7	251	15.4
Ineligible	6	0.8	14	1.7	4	0.6	19	2.7	18	1.2	33	2.3
Total	764		850		641		734		1,471		1,628	
Total Eligible	758		836		637		715		1,453		1,595	
Response Rate (%)	86.5		82.9		83.0		81.8		86.2		83.9	

Notes: All counts and percentages are based on matched sample members who responded to the six month survey. Response rates are calculated as the number of respondents divided by the number of eligible respondents. The total row shows the twelve-month survey sample size, which is equivalent to the number of six-month respondents.

Nonresponse Weighting Adjustment

The GUIDE classification forest was used with the same auxiliary variables used for imputation of unknown eligibility to estimate the response propensity for only eligible respondents and nonrespondents. The estimated response propensity was then used to form quintiles to use as weighting classes for nonresponse adjustment for the twelve-month survey.

Within each weighting class, the nonresponse adjustment factor was first calculated as the ratio of the sum of the six month survey (nonresponse-adjusted) weights for the initial sample to the sum of the six month weights for the twelve month survey respondents. This adjustment factor is the same for all respondents within the same weighting class. The nonresponse-adjusted twelve-month survey weight was then obtained by multiplying this factor with the six-month survey weight. The nonresponse adjustment procedure was carried out separately for each of the six samples.

Descriptive statistics of the nonresponse-adjusted weights are shown in Appendix Table A.6 for the six- and twelve- month surveys along with the estimated design effect (which indicates how much a stratified design biases results relative to a simple random sample) based on the variation of the weights.[76] The table shows a slight increase in the design effect from that of the six-month survey weights. This was expected because the starting weight for the twelve-month nonresponse adjustment was the six-month nonresponse-adjusted weights, and the twelve-month adjustment introduced more variation. Fortunately, the extra variation introduced by the twelve-month survey nonresponse adjustment was quite small.

Appendix Table A.6: Descriptive Statistics of the Nonresponse-adjusted Weights and Design Effect

Program Type	Survey Type	Survey Month	Sample Size	Mean	Standard Deviation	Design Effect
PANO	Participants	6m	765	1.359	0.616	1.206
		12m	656	1.568	0.736	1.221
	National	6m	850	1.212	0.236	1.038
		12m	693	1.459	0.296	1.041
CDM	Participants	6m	641	1.420	0.717	1.255
		12m	529	1.711	0.903	1.279
	National	6m	736	1.227	0.227	1.034
		12m	585	1.506	0.290	1.037
FP	Participants	6m	1471	1.360	0.408	1.090
		12m	1252	1.577	0.495	1.099
	National	6m	1631	1.206	0.156	1.017
		12m	1339	1.439	0.219	1.023

Using Weights in Analysis

When analyzing the data for each survey (baseline, six-month, and twelve-month) separately, the final weights along with corresponding replicate weights developed for that

[76] For a non-cluster sample design, the design effect can be estimated by a simple formula given by Kish (1992), $1 + C^2$, where C^2 is the squared coefficient of variation (i.e., relative variance) of the weights. This measure provides how much the sampling efficiency is lost because of variable weights against equal weights (of the simple random sample) when estimating the population mean. A design effect of 1.221 means that the respondent sample size of 656 for the twelve-month survey in the PANO program is equivalent to 537 (= 656/1.221) of a simple random sample – this design effect-adjusted sample size is called the effective sample size. The effective sample size is further reduced as the design effect is greater than and further away from one.

particular survey are used. For the longitudinal analysis presented in this report, the twelve-month survey weights were used for analysis of all time points among the set of matched sample members responding at twelve months. Specifically, the twelve-month survey weights were attached to records containing data from baseline, six months, and twelve months for all respondents to the twelve-month survey. Matched sample members who did not respond at twelve months were excluded from the longitudinal analysis.

For estimation of the variance, 200 jackknife replicate weights were developed for the twelve-month survey. All analyses were completed in SAS version 9.4 using procedures for complex survey designs that permit application of replicate weights. The regression-adjusted means and DiD estimates were generated using the SURVEYREG procedure.

APPENDIX B– INTENTION-TO-TREAT (ITT) ANALYSIS SINGLE DIFFERENCE TABLES

This appendix presents the survey-based ITT analysis single difference results in Section B.1 and the claims-based ITT analysis single difference results in Section B.2.

B.1 Survey-Based ITT Analysis Single Difference Tables

Appendix Table B.1 through Appendix Table B.9 present survey-based single difference analysis results across priority areas.

Appendix Table B.1: Difference in Six Month and Twelve Month from Baseline Means for Physical Health Status Outcomes, CDM Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Physical Components Summary Score				
Number of Beneficiaries	577	510	571	506
Score Difference	0.05	-0.10	-0.49*	-0.28
90% Confidence Interval	(-0.3,0.4)	(-0.7,0.4)	(-0.9,-0.1)	(-1.0,0.4)
P-Value	0.83	0.76	0.06	0.51
Physical Functioning Subscale				
Number of Beneficiaries	578	529	577	529
Score Difference	-0.42	0.29	-0.84***	-0.24
90% Confidence Interval	(-0.9,0.1)	(-0.3,0.9)	(-1.3,-0.4)	(-0.9,0.5)
P-Value	0.15	0.46	0.00	0.58
Role Physical Subscale				
Number of Beneficiaries	577	509	577	504
Score Difference	0.14	0.43	-0.45	-0.14
90% Confidence Interval	(-0.4,0.6)	(-0.2,1.1)	(-1.0,0.0)	(-0.8,0.5)
P-Value	0.65	0.30	0.14	0.73
Bodily Pain Subscale				
Number of Beneficiaries	583	509	577	506
Score Difference	0.49	0.02	0.14	0.39
90% Confidence Interval	(-0.0,1.0)	(-0.7,0.7)	(-0.4,0.7)	(-0.4,1.2)
P-Value	0.11	0.96	0.69	0.42
General Health Subscale				
Number of Beneficiaries	585	529	585	529
Score Difference	-0.16	0.01	-0.76**	-0.78*
90% Confidence Interval	(-0.7,0.3)	(-0.7,0.7)	(-1.3,-0.2)	(-1.5,-0.0)
P-Value	0.60	0.98	0.02	0.09

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.2: Difference in Six Month and Twelve Month from Baseline Means for Mental Health Status Outcomes, CDM Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Mental Components Summary Score				
Number of Beneficiaries	577	510	571	506
Score Difference	0.09	0.81*	-0.26	-0.22
90% Confidence Interval	(-0.4,0.6)	(0.1,1.5)	(-0.9,0.3)	(-1.0,0.5)
P-Value	0.77	0.06	0.47	0.62
Vitality Subscale				
Number of Beneficiaries	584	528	579	528
Score Difference	0.22	-0.10	-0.83***	-0.72*
90% Confidence Interval	(-0.3,0.7)	(-0.8,0.6)	(-1.3,-0.3)	(-1.4,-0.1)
P-Value	0.48	0.82	0.01	0.08
Social Functioning Subscale				
Number of Beneficiaries	584	528	579	527
Score Difference	0.39	0.04	0.19	0.02
90% Confidence Interval	(-0.2,0.9)	(-0.8,0.9)	(-0.4,0.8)	(-0.8,0.9)
P-Value	0.25	0.94	0.61	0.97
Role Emotional Subscale				
Number of Beneficiaries	574	507	573	503
Score Difference	-0.21	1.20**	-0.37	0.15
90% Confidence Interval	(-0.9,0.4)	(0.4,2.0)	(-1.1,0.3)	(-0.7,1.0)
P-Value	0.59	0.02	0.40	0.77
Mental Health Subscale				
Number of Beneficiaries	584	528	579	529
Score Difference	-0.07	0.99***	-0.47	-0.22
90% Confidence Interval	(-0.6,0.4)	(0.4,1.6)	(-1.0,0.1)	(-0.9,0.5)
P-Value	0.82	0.01	0.16	0.61

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.3: Difference in Six Month and Twelve Month from Baseline Means for Activity, Balance, and Medication Adherence Measures, CDM Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Aerobic Activity				
Number of Beneficiaries	561	517	563	513
Score Difference	-0.03	-0.16	-0.07	-0.22*
90% Confidence Interval	(-0.2,0.1)	(-0.3,0.0)	(-0.2,0.1)	(-0.4,-0.0)
P-Value	0.73	0.11	0.39	0.05
Strength and Flexibility				
Number of Beneficiaries	549	490	546	488
Score Difference	-0.01	0.00	-0.02	0.00
90% Confidence Interval	(-0.0,0.0)	(-0.0,0.0)	(-0.1,0.0)	(-0.1,0.0)
P-Value	0.68	0.99	0.40	0.95

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Presence of Falls in Last Six Months				
Number of Beneficiaries	537	477	539	477
Score Difference	0.01	0.01	-0.01	0.04
90% Confidence Interval	(-0.0,0.0)	(-0.0,0.0)	(-0.0,0.0)	(-0.0,0.1)
P-Value	0.66	0.74	0.60	0.16
Confidence in Balance				
Number of Beneficiaries	385	348	392	352
Score Difference	-1.63*	1.70	-3.91***	0.61
90% Confidence Interval	(-3.0,-0.2)	(-0.5,3.9)	(-5.6,-2.3)	(-1.7,2.9)
P-Value	0.05	0.20	0.00	0.66
Medication Adherence				
Number of Beneficiaries	521	465	528	451
Score Difference	0.10**	0.08	0.13***	0.11*
90% Confidence Interval	(0.0,0.2)	(-0.0,0.2)	(0.1,0.2)	(0.0,0.2)
P-Value	0.02	0.22	0.01	0.08

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.4: Difference in Six Month and Twelve Month from Baseline Means for Physical Health Status Outcomes, PANO Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Physical Components Summary Score				
Number of Beneficiaries	677	635	676	631
Score Difference	-0.49**	-0.14	-0.43*	-0.43
90% Confidence Interval	(-0.9,-0.1)	(-0.6,0.3)	(-0.8,-0.0)	(-0.9,0.1)
P-Value	0.04	0.62	0.07	0.16
Physical Functioning Subscale				
Number of Beneficiaries	683	655	680	656
Score Difference	-0.48*	-0.01	-0.71***	-0.16
90% Confidence Interval	(-0.9,-0.0)	(-0.5,0.5)	(-1.2,-0.3)	(-0.6,0.3)
P-Value	0.07	0.97	0.01	0.57
Role Physical Subscale				
Number of Beneficiaries	683	635	679	631
Score Difference	-0.70***	0.23	0.01	0.26
90% Confidence Interval	(-1.1,-0.3)	(-0.3,0.8)	(-0.4,0.5)	(-0.4,0.9)
P-Value	0.01	0.50	0.97	0.50
Bodily Pain Subscale				
Number of Beneficiaries	684	635	687	629
Score Difference	-0.09	0.60	0.01	0.54
90% Confidence Interval	(-0.6,0.4)	(-0.1,1.3)	(-0.5,0.5)	(-0.2,1.2)
P-Value	0.76	0.14	0.98	0.21
General Health Subscale				

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Number of Beneficiaries	693	655	693	656
Score Difference	-0.42**	-0.65**	-0.88***	-1.11***
90% Confidence Interval	(-0.8,-0.1)	(-1.2,-0.1)	(-1.3,-0.5)	(-1.6,-0.6)
P-Value	0.05	0.04	0.00	0.00

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.5: Difference in Six Month and Twelve Month from Baseline Means for Mental Health Status Outcomes, PANO Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Mental Components Summary Score				
Number of Beneficiaries	677	635	676	631
Score Difference	-0.49**	-0.14	-0.43*	-0.43
90% Confidence Interval	(-0.9,-0.1)	(-0.6,0.3)	(-0.8,-0.0)	(-0.9,0.1)
P-Value	0.04	0.62	0.07	0.16
Vitality Subscale				
Number of Beneficiaries	688	655	684	655
Score Difference	-0.69***	-0.27	-0.30	0.03
90% Confidence Interval	(-1.1,-0.3)	(-0.8,0.3)	(-0.7,0.1)	(-0.5,0.6)
P-Value	0.01	0.41	0.21	0.92
Social Functioning Subscale				
Number of Beneficiaries	687	654	683	654
Score Difference	-0.56*	0.29	-0.13	-0.25
90% Confidence Interval	(-1.1,-0.1)	(-0.3,0.9)	(-0.7,0.4)	(-0.9,0.4)
P-Value	0.06	0.45	0.70	0.51
Role Emotional Subscale				
Number of Beneficiaries	682	631	669	608
Score Difference	-0.14	0.95**	0.19	0.19
90% Confidence Interval	(-0.7,0.4)	(0.2,1.7)	(-0.4,0.8)	(-0.4,0.8)
P-Value	0.69	0.05	0.58	0.62
Mental Health Subscale				
Number of Beneficiaries	688	655	684	655
Score Difference	-0.14	0.31	-0.31	0.20
90% Confidence Interval	(-0.6,0.3)	(-0.3,0.9)	(-0.8,0.1)	(-0.4,0.8)
P-Value	0.57	0.38	0.27	0.55

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.6: Difference in Six Month and Twelve Month from Baseline Means for Activity, Balance, and Medication Adherence Measures, PANO Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Aerobic Activity				
Number of Beneficiaries	674	647	675	648
Score Difference	-0.15**	0.04	-0.21***	0.02
90% Confidence Interval	(-0.3,-0.0)	(-0.1,0.2)	(-0.3,-0.1)	(-0.1,0.2)
P-Value	0.03	0.60	0.00	0.83
Strength and Flexibility				
Number of Beneficiaries	648	600	648	604
Score Difference	-0.07***	0.08***	-0.06***	0.08***
90% Confidence Interval	(-0.1,-0.0)	(0.0,0.1)	(-0.1,-0.0)	(0.0,0.1)
P-Value	0.00	0.00	0.00	0.00
Presence of Falls in Last Six Months				
Number of Beneficiaries	651	602	647	604
Score Difference	-0.01	0.02	-0.01	0.00
90% Confidence Interval	(-0.0,0.0)	(-0.0,0.0)	(-0.0,0.0)	(-0.0,0.0)
P-Value	0.67	0.18	0.51	0.83
Confidence in Balance				
Number of Beneficiaries	457	455	475	447
Score Difference	-1.41*	-0.51	-3.04***	-0.48
90% Confidence Interval	(-2.7,-0.1)	(-2.5,1.4)	(-4.6,-1.5)	(-2.2,1.2)
P-Value	0.08	0.67	0.00	0.63
Medication Adherence				
Number of Beneficiaries	581	546	587	545
Score Difference	0.06	0.02	0.10**	0.13***
90% Confidence Interval	(-0.0,0.1)	(-0.1,0.1)	(0.0,0.2)	(0.1,0.2)
P-Value	0.13	0.67	0.01	0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.7: Difference in Six Month and Twelve Month from Baseline Means for Physical Health Status Outcomes, FP Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Physical Components Summary Score				
Number of Beneficiaries	1,304	1,200	1,307	1,187
Score Difference	-0.52***	-0.51**	-0.86***	-0.62***
90% Confidence Interval	(-0.8,-0.3)	(-0.9,-0.2)	(-1.2,-0.5)	(-1.0,-0.2)
P-Value	0.00	0.02	0.00	0.01
Physical Functioning Subscale				
Number of Beneficiaries	1,313	1,251	1,317	1,249
Score Difference	-0.61***	-0.43*	-0.99***	-0.60**

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
90% Confidence Interval	(-0.9,-0.3)	(-0.8,-0.1)	(-1.4,-0.6)	(-1.0,-0.2)
P-Value	0.00	0.05	0.00	0.02
Role Physical Subscale				
Number of Beneficiaries	1,312	1,199	1,316	1,187
Score Difference	-0.57***	0.18	-0.71***	-0.08
90% Confidence Interval	(-0.9,-0.2)	(-0.2,0.6)	(-1.1,-0.4)	(-0.5,0.3)
P-Value	0.01	0.44	0.00	0.76
Bodily Pain Subscale				
Number of Beneficiaries	1,325	1,200	1,326	1,187
Score Difference	-0.14	0.40*	-0.50**	0.08
90% Confidence Interval	(-0.5,0.2)	(0.0,0.8)	(-0.8,-0.2)	(-0.3,0.5)
P-Value	0.48	0.10	0.01	0.72
General Health Subscale				
Number of Beneficiaries	1,338	1,249	1,339	1,246
Score Difference	-0.35**	-0.48**	-1.04***	-0.91***
90% Confidence Interval	(-0.6,-0.1)	(-0.8,-0.2)	(-1.3,-0.8)	(-1.2,-0.6)
P-Value	0.04	0.01	0.00	0.00

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.8: Difference in Six Month and Twelve Month from Baseline Means for Mental Health Status Outcomes, FP Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Mental Components Summary Score				
Number of Beneficiaries	1,304	1,200	1,307	1,187
Score Difference	0.11	1.05***	-0.42*	0.40
90% Confidence Interval	(-0.3,0.5)	(0.6,1.5)	(-0.8,-0.0)	(-0.0,0.8)
P-Value	0.67	0.00	0.09	0.10
Vitality Subscale				
Number of Beneficiaries	1,329	1,249	1,329	1,249
Score Difference	-0.03	0.02	-0.69***	-0.40*
90% Confidence Interval	(-0.4,0.3)	(-0.4,0.4)	(-1.0,-0.4)	(-0.7,-0.0)
P-Value	0.86	0.93	0.00	0.06
Social Functioning Subscale				
Number of Beneficiaries	1,328	1,246	1,329	1,250
Score Difference	-0.01	0.39	-0.50**	0.01
90% Confidence Interval	(-0.5,0.5)	(-0.1,0.8)	(-0.9,-0.1)	(-0.4,0.4)
P-Value	0.96	0.15	0.02	0.96
Role Emotional Subscale				
Number of Beneficiaries	1,309	1,192	1,314	1,186
Score Difference	-0.37	1.14***	-0.78**	0.44
90% Confidence Interval	(-0.9,0.1)	(0.6,1.7)	(-1.3,-0.3)	(-0.1,1.0)
P-Value	0.22	0.00	0.01	0.17

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Mental Health Subscale				
Number of Beneficiaries	1,329	1,250	1,329	1,249
Score Difference	0.02	0.76***	-0.37	0.19
90% Confidence Interval	(-0.4,0.4)	(0.4,1.1)	(-0.8,0.0)	(-0.2,0.6)
P-Value	0.92	0.00	0.12	0.41

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

Appendix Table B.9: Difference in Six Month and Twelve Month from Baseline Means for Activity, Balance, and Medication Adherence Measures, FP Programs

Measures	Comparison (Six Month - Baseline)	Participants (Six Month - Baseline)	Comparison (Twelve Month - Baseline)	Participant (Twelve Month - Baseline)
Aerobic Activity				
Number of Beneficiaries	1,279	1,223	1,291	1,220
Score Difference	0.05	-0.14**	-0.15***	-0.27***
90% Confidence Interval	(-0.0,0.1)	(-0.2,-0.0)	(-0.2,-0.1)	(-0.4,-0.2)
P-Value	0.31	0.02	0.00	0.00
Strength and Flexibility				
Number of Beneficiaries	1,213	1,152	1,217	1,134
Score Difference	0.02	0.06***	0.00	0.05***
90% Confidence Interval	(-0.0,0.0)	(0.0,0.1)	(-0.0,0.0)	(0.0,0.1)
P-Value	0.18	0.00	0.98	0.00
Presence of Falls in Last Six Months				
Number of Beneficiaries	1,219	1,149	1,229	1,128
Score Difference	-0.07***	-0.05***	-0.08***	-0.06***
90% Confidence Interval	(-0.1,-0.0)	(-0.1,-0.0)	(-0.1,-0.1)	(-0.1,-0.0)
P-Value	0.00	0.00	0.00	0.00
Confidence in Balance				
Number of Beneficiaries	930	928	908	895
Score Difference	-0.58	2.08***	-3.29***	0.83
90% Confidence Interval	(-1.7,0.6)	(0.9,3.3)	(-4.4,-2.2)	(-0.4,2.0)
P-Value	0.41	0.00	0.00	0.26
Medication Adherence				
Number of Beneficiaries	1,174	1,076	1,169	1,061
Score Difference	0.03	0.04	0.04	0.04
90% Confidence Interval	(-0.0,0.1)	(-0.0,0.1)	(-0.0,0.1)	(-0.0,0.1)
P-Value	0.22	0.24	0.15	0.24

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. Includes all program participants, regardless of program completion, and their comparators.

B.2 Claims-Based ITT Analysis Single Difference Tables

Appendix Table B.10 through Appendix Table B.21 present the ITT single difference analysis results for healthcare service utilization, the incidence of falls and fractures, expenditure, and medication adherence by priority area.

Appendix Table B.10: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Rate of ER Visits, Rate of Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 beneficiaries, CDM Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
ER Visits						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	97	98	69	106	166	197
Difference	8.37	50.21	-57.64**	51.06	-48.68	101.28*
90% Confidence Interval	(-42.5, 59.2)	(-10.3, 110.8)	(-103.4, -11.9)	(-9.1, 111.2)	(-130.8, 33.4)	(2.5, 200.1)
P-value	0.79	0.17	0.04	0.16	0.33	0.09
Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	45	59	43	54	88	113
Difference	22.12	47.25*	11.60	23.03	33.47	69.94*
90% Confidence Interval	(-18.2, 62.5)	(1.8, 92.7)	(-25.1, 48.3)	(-17.4, 63.5)	(-27.7, 94.6)	(0.7, 139.2)
P-value	0.37	0.09	0.60	0.35	0.37	0.10
Unplanned Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	33	49	35	47	68	96
Difference	19.73	48.32**	16.43	41.16*	35.65	89.06**
90% Confidence Interval	(-16.9, 56.4)	(7.9, 88.8)	(-15.8, 48.6)	(4.2, 78.2)	(-18.7, 90.0)	(27.6, 150.5)
P-value	0.38	0.05	0.40	0.07	0.28	0.02
Length of Stay						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	45	58	43	52	88	110
Difference	9.07	201.33	95.93	156.35	104.54	355.90
90% Confidence Interval	(-244.4, 262.5)	(-65.1, 467.8)	(-217.9, 409.8)	(-128.7, 441.4)	(-364.9, 573.9)	(-83.4, 795.2)
P-value	0.95	0.21	0.62	0.37	0.71	0.18
Falls/Fractures						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	72	48	78	82	150	130
Difference	16.16	-23.82	32.27	44.85**	24.15	10.25
90% Confidence Interval	(-18.2, 50.5)	(-55.0, 7.4)	(-3.2, 67.8)	(9.4, 80.3)	(-5.6, 53.9)	(-18.4, 38.9)
P-value	0.44	0.21	0.14	0.04	0.18	0.56

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis:

comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates, except Falls/Fractures, are regression-adjusted for these covariates.

Appendix Table B.11: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Healthcare Expenditures per Beneficiary, CDM Programs

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Part D						
Total Participant Observations in the Post-Intervention Period	368	368	357	359	725	727
Nonzero Participant Observations in the Post-Intervention Period	362	361	352	354	714	715
Difference	$44.38	$118.72	- $274.09	$317.44	- $226.86	$439.62
90% Confidence Interval	(-846.2, 934.9)	(-465.4, 702.8)	(-1,084.9, 536.7)	(-303.9, 938.8)	(-1669, 1,215.3)	(-610.3, 1,489.5)
P-value	0.94	0.74	0.58	0.40	0.80	0.49
Total Medical						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	499	503	488	492	987	995
Difference	- $36.06	$854.86	- $57.80	$337.78	- $91.96	$1,190.15
90% Confidence Interval	(-750.3, 678.2)	(0.3, 1709.5)	(-849.9, 734.3)	(-417.6, 1,093.2)	(-1,335.4, 1,151.5)	(-134.9, 2,515.2)
P-value	0.93	0.10	0.90	0.46	0.90	0.14
Inpatient						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	44	58	42	53	86	111
Difference	- $27.37	$337.62	$219.26	$181.63	$191.16	$516.85
90% Confidence Interval	(-375.8, 321.1)	(-77.9, 753.1)	(-276.4, 714.9)	(-200.4, 563.6)	(-490.7, 873)	(-115.2, 1148.9)
P-value	0.90	0.18	0.47	0.43	0.65	0.18
Outpatient ER						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	97	94	69	103	166	197
Difference	- $9.73	$35.40	- $39.13	$15.28	- $49.25	$49.89
90% Confidence Interval	(-55.0, 35.6)	(-32.0, 102.8)	(-82.3, 4.1)	(-33.9, 64.5)	(-128.5, 30)	(-48.0, 147.8)
P-value	0.72	0.39	0.14	0.61	0.31	0.40
Outpatient Non-ER						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	347	349	331	341	678	690
Difference	$69.81	- $41.52	- $29.74	$73.32	$41.49	$32.60
90% Confidence Interval	(-179.8, 319.4)	(-229.9, 146.8)	(-236.1, 176.6)	(-98.5, 245.1)	(-327.5, 410.5)	(-267.9, 333.1)
P-value	0.65	0.72	0.81	0.48	0.85	0.86
Physician and Ancillary						

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	494	501	485	490	979	991
Difference	- $56.48	$130.82	- $173.80	- $0.04	- $229.44	$133.60
90% Confidence Interval	(-306.4, 193.4)	(-84.8, 346.4)	(-397.3, 49.8)	(-208.0, 207.9)	(-635.9, 177.0)	(-236.7, 503.9)
P-value	0.71	0.32	0.20	1.00	0.35	0.55
Durable Medical Equipment						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	183	209	170	213	353	422
Difference	$3.24	- $11.06	- $7.44	- $38.58	- $4.22	- $49.58
90% Confidence Interval	(-49.2, 55.7)	(-83.0, 60.9)	(-63.7, 48.9)	(-104.2, 27.1)	(-95.0, 86.5)	(-171.8, 72.6)
P-value	0.92	0.80	0.83	0.33	0.94	0.51
Home Health						
Total Participant Observations in the Post-Intervention Period	517	514	502	506	1,019	1,020
Nonzero Participant Observations in the Post-Intervention Period	30	40	28	35	58	75
Difference	$27.59	$84.96	- $19.60	$16.18	$7.94	$100.68
90% Confidence Interval	(-65.7, 120.9)	(-31.8, 201.8)	(-98.3, 59.1)	(-87.8, 120.2)	(-136.3, 152.2)	(-83.1, 284.4)
P-value	0.63	0.23	0.68	0.80	0.93	0.37

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.12: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (Average Proportion of Days Covered), CDM Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	112	126	111	122	223	248
Difference	3.76**	5.54***	2.64	4.96***	3.27**	5.28***
90% Confidence Interval	(1.3, 6.3)	(3, 8.1)	(0.0, 5.3)	(2.3, 7.6)	(0.9, 5.6)	(2.9, 7.7)
P-value	0.01	< 0.01	0.10	< 0.01	0.02	< 0.01
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	85	96	82	94	167	190
Difference	2.32	9.42***	5.89***	9.06***	3.95*	9.25***
90% Confidence Interval	(-1.6, 6.2)	(6.2, 12.6)	(2.5, 9.3)	(5.9, 12.2)	(0.6, 7.3)	(6.3, 12.2)
P-value	0.33	< 0.01	< 0.01	< 0.01	0.05	< 0.01
Diabetes Medication						

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	97	88	94	80	191	168
Difference	4.37***	3.36*	4.12**	5.27***	4.32***	4.27**
90% Confidence Interval	(1.8, 7)	(0.1, 6.7)	(1.4, 6.8)	(2.0, 8.5)	(1.8, 6.8)	(1.3, 7.3)
P-value	0.01	0.10	0.01	0.01	< 0.01	0.02
RAS Antagonists						
Total Participant Observations in the Post-Intervention Period	174	164	165	177	339	341
Difference	3.56***	3.81***	4.08***	4.77***	3.84***	4.3***
90% Confidence Interval	(1.4, 5.7)	(1.5, 6.1)	(2.0, 6.2)	(2.7, 6.8)	(1.9, 5.8)	(2.3, 6.3)
P-value	0.01	0.01	< 0.01	< 0.01	< 0.01	< 0.01
Statins						
Total Participant Observations in the Post-Intervention Period	158	167	153	159	311	326
Difference	8.36***	5.21***	8.01***	6.74***	8.25***	5.98***
90% Confidence Interval	(5.8, 10.9)	(2.8, 7.6)	(5.5, 10.5)	(4.5, 9.0)	(5.9, 10.6)	(3.9, 8.0)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.13: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (PDC ≥ 80%), CDM Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	112	126	111	122	223	248
Difference	2.45	9.72***	4.32	9.02**	3.60	9.43***
90% Confidence Interval	(-4.2, 9.1)	(3.6, 15.8)	(-2.2, 10.8)	(2.9, 15.2)	(-2.2, 9.4)	(3.9, 15)
P-value	0.55	0.01	0.27	0.02	0.30	0.01
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	85	96	82	94	167	190
Difference	1.46	20.09***	9.08*	17.85***	5.23	18.98***
90% Confidence Interval	(-7.1, 10.0)	(12.9, 27.3)	(1.4, 16.7)	(10.5, 25.2)	(-1.9, 12.4)	(12.1, 25.8)
P-value	0.78	< 0.01	0.05	0.00	0.23	< 0.01
Diabetes Medication						
Total Participant Observations in the Post-Intervention Period	97	88	94	80	191	168
Difference	7.63*	2.57	7.14	9.81**	7.57*	5.95
90% Confidence Interval	(0.6, 14.7)	(-5.8, 11.0)	(0.0, 14.3)	(2.5, 17.1)	(1.2, 14.0)	(-1.1, 13.0)
P-value	0.08	0.62	0.10	0.03	0.05	0.17
RAS Antagonists						

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	174	164	165	177	339	341
Difference	4.43	3.90	7.72**	7.39**	6.15**	5.7**
90% Confidence Interval	(-0.9, 9.8)	(-1.4, 9.3)	(2.7, 12.7)	(2.6, 12.2)	(1.5, 10.8)	(1.2, 10.2)
P-value	0.17	0.23	0.01	0.01	0.03	0.04
Statins						
Total Participant Observations in the Post-Intervention Period	158	167	153	159	311	326
Difference	17.88***	10.81***	15.95***	13.54***	17.02***	12.17***
90% Confidence Interval	(11.8, 24.0)	(4.5, 17.1)	(9.5, 22.4)	(7.4, 19.7)	(11.3, 22.7)	(6.6, 17.7)
P-value	< 0.01	0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.14: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Rate of ER Visits, Rate of Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 beneficiaries, PANO Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
ER Visits						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	74	84	89	60	163	144
Difference	12.96	25.24	35.12	-36.26	47.87	-10.24
90% Confidence Interval	(-32.2, 58.2)	(-30.0, 80.5)	(-9.4, 79.7)	(-85.3, 12.8)	(-26.4, 122.1)	(-105.8, 85.3)
P-value	0.64	0.45	0.20	0.22	0.29	0.86
Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	42	31	42	39	84	70
Difference	4.60	-27.49	18.34	-0.47	22.56	-28.28
90% Confidence Interval	(-24.2, 33.4)	(-68, 13)	(-17.6, 54.2)	(-49.6, 48.6)	(-29.9, 75)	(-108.6, 52)
P-value	0.79	0.26	0.40	0.99	0.48	0.56
Unplanned Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	28	25	37	26	65	51
Difference	7.70	-21.15	37.14*	-1.01	44.2*	-22.51
90% Confidence Interval	(-15.8, 31.1)	(-58.7, 16.4)	(5.7, 68.6)	(-47.6, 45.6)	(0.5, 87.9)	(-98.8, 53.8)
P-value	0.59	0.35	0.05	0.97	0.10	0.63

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Length of Stay						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	42	30	42	39	84	69
Difference	36.25	-178.38	113.37	-13.89	149.15	-193.74
90% Confidence Interval	(-149.4, 221.9)	(-390.8, 34.1)	(-103.0, 329.8)	(-296.7, 268.9)	(-180.8, 479.1)	(-627.1, 239.6)
P-value	0.75	0.17	0.39	0.94	0.46	0.46
Falls/Fractures						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	68	51	94	81	162	132
Difference	20.32	16.65	66.51***	70.29***	43.23***	43.26***
90% Confidence Interval	(-9.2, 49.9)	(-9.2, 42.6)	(34.4, 98.7)	(40.8, 99.8)	(17.3, 69.2)	(20.1, 66.4)
P-value	0.26	0.29	< 0.01	< 0.01	0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates, except Falls/Fractures, are regression-adjusted for these covariates.

Appendix Table B.15: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Healthcare Expenditures per Beneficiary, PANO Programs

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Part D						
Total Participant Observations in the Post-Intervention Period	406	381	401	374	807	755
Nonzero Participant Observations in the Post-Intervention Period	398	360	393	355	791	715
Difference	- $82.67	- $17.15	- $34.30	- $200.25	- $116.43	- $215.77
90% Confidence Interval	(-583.4, 418)	(-492.3, 458)	(-596.2, 527.5)	(-584, 183.5)	(-1,020.1, 787.3)	(-961.7, 530.1)
P-value	0.79	0.95	0.92	0.39	0.83	0.63
Total Medical						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	557	559	561	562	1,118	1,121
Difference	$275.14	$40.09	$621.86	$289.47	$895.32	$330.47
90% Confidence Interval	(-270.8, 821.1)	(-546.3, 626.5)	(-1.5, 1,245.2)	(-289.6, 868.5)	(-52.4, 1,843.0)	(-648.1, 1309.1)
P-value	0.41	0.91	0.10	0.41	0.12	0.58
Inpatient						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Nonzero Participant Observations in the Post-Intervention Period	43	31	41	39	84	70
Difference	$50.36	- $85.09	$76.89	$5.96	$126.34	- $78.76
90% Confidence Interval	(-213.7, 314.4)	(-400.6, 230.4)	(-215.1, 368.8)	(-313.0, 324.9)	(-333.4, 586.1)	(-610.2, 452.7)
P-value	0.75	0.66	0.67	0.98	0.65	0.81
Outpatient ER						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	74	84	88	60	162	144
Difference	$52.46	- $3.49	$17.29	-52.83**	$70.68	- $55.30
90% Confidence Interval	(-19.9, 124.8)	(-51.2, 44.3)	(-16.8, 51.3)	(-96.1, -9.6)	(-15.6, 156.9)	(-141.3, 30.7)
P-value	0.23	0.90	0.40	0.05	0.18	0.29
Outpatient Non-ER						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	371	379	375	370	746	749
Difference	$141.14	$29.52	186.82*	- $0.94	326.32**	$28.15
90% Confidence Interval	(-6.9, 289.2)	(-124.8, 183.9)	(24.9, 348.8)	(-150.3, 148.4)	(90.0, 562.7)	(-241.1, 297.4)
P-value	0.12	0.75	0.06	0.99	0.02	0.86
Physician and Ancillary						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	551	555	555	559	1,106	1,114
Difference	$14.70	- $16.89	$204.03	$119.76	$217.93	$103.16
90% Confidence Interval	(-139.3, 168.7)	(-185.6, 151.9)	(-63.6, 471.7)	(-55.4, 294.9)	(-118.8, 554.7)	(-195.1, 401.5)
P-value	0.88	0.87	0.21	0.26	0.29	0.57
Durable Medical Equipment						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	132	112	138	117	270	229
Difference	$7.11	- $13.05	$27.59	- $9.99	$34.53	- $23.21
90% Confidence Interval	(-41.4, 55.6)	(-31.5, 5.4)	(-28.3, 83.5)	(-30.1, 10.1)	(-51.8, 120.8)	(-57.5, 11)
P-value	0.81	0.25	0.42	0.41	0.51	0.27
Home Health						
Total Participant Observations in the Post-Intervention Period	584	583	578	574	1,162	1,157
Nonzero Participant Observations in the Post-Intervention Period	28	17	28	28	56	45
Difference	$26.27	- $9.06	- $11.37	$21.14	$16.18	$12.31
90% Confidence Interval	(-56.3, 108.9)	(-82.7, 64.6)	(-82.5, 59.7)	(-53, 95.3)	(-113.2, 145.5)	(-112.8, 137.4)
P-value	0.60	0.84	0.79	0.64	0.84	0.87

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and

reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.16: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (Average Proportion of Days Covered), PANO Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	114	88	128	94	242	182
Difference	3.63**	3.38*	3.26*	4.31***	3.46**	3.85**
90% Confidence Interval	(0.9, 6.4)	(0.5, 6.3)	(0.5, 6.0)	(1.6, 7.0)	(1.0, 5.9)	(1.4, 6.3)
P-value	0.03	0.05	0.05	0.01	0.02	0.01
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	91	65	85	74	176	139
Difference	4.81***	0.68	3.26	4.50***	4.03**	2.75*
90% Confidence Interval	(1.9, 7.8)	(-2.6, 3.9)	(-0.1, 6.6)	(1.9, 7.1)	(1.2, 6.9)	(0.2, 5.3)
P-value	0.01	0.73	0.11	< 0.01	0.02	0.08
Diabetes Medication						
Total Participant Observations in the Post-Intervention Period	59	51	61	47	120	98
Difference	1.83	2.84	-0.26	6.39***	0.89	4.6**
90% Confidence Interval	(-1.9, 5.6)	(-1.2, 6.9)	(-4.4, 3.9)	(3.2, 9.6)	(-2.7, 4.5)	(1.3, 7.9)
P-value	0.42	0.25	0.92	< 0.01	0.69	0.02
RAS Antagonists						
Total Participant Observations in the Post-Intervention Period	157	134	154	140	311	274
Difference	5.37***	4.04***	6.03***	3.9***	5.69***	4.00***
90% Confidence Interval	(3.2, 7.5)	(1.9, 6.1)	(3.9, 8.2)	(1.8, 6.0)	(3.7, 7.7)	(2.1, 5.9)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Statins						
Total Participant Observations in the Post-Intervention Period	167	143	159	159	326	302
Difference	6.5***	7.17***	6.47***	8.02***	6.51***	7.63***
90% Confidence Interval	(4.2, 8.8)	(4.6, 9.8)	(4.3, 8.7)	(5.5, 10.5)	(4.4, 8.6)	(5.3, 9.9)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.17: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (PDC ≥ 80%), PANO Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	114	88	128	94	242	182
Difference	4.68	10.26**	1.20	10.51**	3.00	10.26**
90% Confidence Interval	(-2.2, 11.6)	(2.0, 18.5)	(-6.1, 8.5)	(2.6, 18.5)	(-3.2, 9.2)	(2.9, 17.6)
P-value	0.27	0.04	0.79	0.03	0.43	0.02
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	91	65	85	74	176	139
Difference	6.84	0.36	2.52	10.04**	4.70	5.58
90% Confidence Interval	(-1.2, 14.9)	(-8.5, 9.3)	(-6.1, 11.2)	(3.3, 16.8)	(-2.6, 12.0)	(-1.4, 12.5)
P-value	0.16	0.95	0.63	0.02	0.29	0.19
Diabetes Medication						
Total Participant Observations in the Post-Intervention Period	59	51	61	47	120	98
Difference	4.30	7.36	-6.28	15.41***	-0.73	11.44**
90% Confidence Interval	(-3.0, 11.6)	(-3.0, 17.7)	(-15.8, 3.2)	(6.1, 24.7)	(-8.1, 6.6)	(2.2, 20.7)
P-value	0.33	0.25	0.28	0.01	0.87	0.04
RAS Antagonists						
Total Participant Observations in the Post-Intervention Period	157	134	154	140	311	274
Difference	9.31***	6.60*	11.49***	6.05*	10.37***	6.44**
90% Confidence Interval	(3.7, 15)	(0.9, 12.3)	(6.2, 16.8)	(0.4, 11.7)	(5.3, 15.4)	(1.3, 11.6)
P-value	0.01	0.06	< 0.01	0.08	< 0.01	0.04
Statins						
Total Participant Observations in the Post-Intervention Period	167	143	159	159	326	302
Difference	11.09***	15.85***	9.28***	17.62***	10.20***	16.83***
90% Confidence Interval	(5.8, 16.4)	(9.0, 22.7)	(3.8, 14.8)	(11.1, 24.1)	(5.3, 15.1)	(10.8, 22.9)
P-value	< 0.01	< 0.01	0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.18: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Rate of ER Visits, Rate of Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 beneficiaries, FP Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
ER Visits						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Nonzero Participant Observations in the Post-Intervention Period	180	180	182	198	362	378
Difference	21.75	-3.01	42.8*	38.75	63.99*	35.06
90% Confidence Interval	(-13.6, 57.1)	(-38.0, 32.0)	(4.8, 80.8)	(-2.1, 79.6)	(4.8, 123.2)	(-25.9, 96.0)
P-value	0.31	0.89	0.06	0.12	0.08	0.34
Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	119	106	117	114	236	220
Difference	38.61**	20.62	36.33**	22.82	75.21***	43.74
90% Confidence Interval	(11.3, 66.0)	(-6.8, 48.0)	(8.9, 63.7)	(-3.4, 49.1)	(32.4, 118)	(-0.1, 87.6)
P-value	0.02	0.22	0.03	0.15	0.00	0.10
Unplanned Inpatient Admissions						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	96	93	98	97	194	190
Difference	37.57**	22.70	39.33***	21.54	76.98***	44.56*
90% Confidence Interval	(12.8, 62.3)	(-1.7, 47.0)	(15, 63.7)	(-1.7, 44.8)	(38.9, 115.1)	(5.7, 83.4)
P-value	0.01	0.13	0.01	0.13	0.00	0.06
Length of Stay						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	117	104	116	114	233	218
Difference	240.85**	97.18	274.18**	182.51*	515.08***	280.34*
90% Confidence Interval	(84.4, 397.3)	(-52.2, 246.5)	(96.9, 451.5)	(17.3, 347.7)	(263.1, 767.1)	(32, 528.7)
P-value	0.01	0.29	0.01	0.07	0.00	0.06
Falls/Fractures						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	186	160	231	216	417	376
Difference	13.00	-12.23	56.63***	43.45***	34.46**	15.16
90% Confidence Interval	(-12.0, 38.0)	(-36.6, 12.1)	(30.2, 83.0)	(17.1, 69.8)	(12.6, 56.4)	(-6.6, 36.9)
P-value	0.39	0.41	< 0.01	0.01	0.01	0.25

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates, except Falls/Fractures, are regression-adjusted for these covariates.

Appendix Table B.19: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Healthcare Expenditures per Beneficiary, FP Programs

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Part D						
Total Participant Observations in the Post-Intervention Period	795	784	779	761	1574	1545
Nonzero Participant Observations in the Post-Intervention Period	773	768	756	743	1529	1511
Difference	- $13.45	$71.98	- $125.84	$79.52	- $137.59	$150.55
90% Confidence Interval	(-332,.0 305.1)	(-206.7, 350.7)	(-368.0, 116.3)	(-217.8, 376.8)	(-625.5, 350.3)	(-315.5, 616.6)
P-value	0.95	0.67	0.39	0.66	0.64	0.60
Total Medical						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	1,096	1,106	1,088	1,075	2,184	2,181
Difference	$320.81	$526.55	849.15***	$442.38	1167.91**	977.72**
90% Confidence Interval	(-144.6, 786.2)	(-30.7, 1083.8)	(321.6, 1376.7)	(-20, 904.7)	(372.9, 1962.9)	(165.7, 1789.8)
P-value	0.26	0.12	0.01	0.12	0.02	0.05
Inpatient						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	117	105	117	111	234	216
Difference	254.2*	$154.37	460.47***	$245.82	713.84***	402.37*
90% Confidence Interval	(12.6, 495.8)	(-78.5, 387.2)	(170, 750.9)	(-13.9, 505.5)	(298.7, 1129)	(8.8, 795.9)
P-value	0.08	0.28	0.01	0.12	0.01	0.09
Outpatient ER						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	180	169	182	186	362	355
Difference	$31.61	- $5.99	$18.14	$18.21	$49.78	$11.99
90% Confidence Interval	(-10.0, 73.2)	(-32.6, 20.6)	(-9.3, 45.5)	(-12.3, 48.7)	(-5.3, 104.9)	(-34.9, 58.9)
P-value	0.21	0.71	0.28	0.33	0.14	0.67
Outpatient Non-ER						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	749	787	736	713	1,485	1,500
Difference	$6.96	$101.17	$111.64	$27.09	$117.58	$129.61
90% Confidence Interval	(-94.0, 107.9)	(-8.8, 211.1)	(-3.2, 226.5)	(-86.1, 140.3)	(-56.7, 291.9)	(-47.0, 306.2)
P-value	0.91	0.13	0.11	0.69	0.27	0.23
Physician and Ancillary						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	1,089	1,101	1,083	1,070	2,172	2,171
Difference	$0.48	$32.34	- $6.32	$77.40	- $4.57	$110.17
90% Confidence Interval	(-140.1, 141.1)	(-110.2, 174.9)	(-138.1, 125.5)	(-53.0, 207.8)	(-240.4, 231.2)	(-115.6, 336)
P-value	1.00	0.71	0.94	0.33	0.98	0.42
Durable Medical Equipment						

Measures (2011 USD)	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	281	296	282	285	563	581
Difference	$21.10	$0.94	$9.34	- $23.70	$30.61	- $22.36
90% Confidence Interval	(-14.8, 57.0)	(-97.4, 99.3)	(-22.0, 40.7)	(-95.9, 48.5)	(-22.6, 83.9)	(-178.2, 133.4)
P-value	0.33	0.99	0.62	0.59	0.34	0.81
Home Health						
Total Participant Observations in the Post-Intervention Period	1,136	1,133	1,114	1,097	2,250	2,230
Nonzero Participant Observations in the Post-Intervention Period	72	58	86	61	158	119
Difference	$28.90	- $51.80	80.63*	- $11.81	$108.47	- $64.08
90% Confidence Interval	(-37.9, 95.7)	(-111.0, 7.4)	(8.2, 153.1)	(-77.5, 53.8)	(-5.8, 222.8)	(-169.7, 41.6)
P-value	0.48	0.15	0.07	0.77	0.12	0.32

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.20: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (Average Proportion of Days Covered), FP Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	243	254	241	251	484	505
Difference	4.32***	2.36**	3.89***	3.69***	4.11***	3.05***
90% Confidence Interval	(2.5, 6.1)	(0.5, 4.3)	(2.0, 5.7)	(1.9, 5.4)	(2.4, 5.8)	(1.4, 4.7)
P-value	< 0.01	0.04	< 0.01	< 0.01	< 0.01	< 0.01
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	163	180	162	170	325	350
Difference	5.26***	4.70***	4.60***	4.67***	4.94***	4.70***
90% Confidence Interval	(3.1, 7.4)	(2.6, 6.8)	(2.4, 6.8)	(2.6, 6.7)	(3.0, 6.9)	(2.8, 6.6)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Diabetes Medication						
Total Participant Observations in the Post-Intervention Period	116	95	107	89	223	184
Difference	2.26	6.64***	3.14*	5.28***	2.72	6.09***
90% Confidence Interval	(-0.8, 5.3)	(3.9, 9.4)	(0.1, 6.2)	(2.4, 8.1)	(0.0, 5.4)	(3.5, 8.7)
P-value	0.22	< 0.01	0.09	< 0.01	0.10	< 0.01
RAS Antagonists						

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	349	329	322	313	671	642
Difference	3.27***	4.29***	3.70***	5.17***	3.50***	4.75***
90% Confidence Interval	(1.8, 4.7)	(2.7, 5.9)	(2.1, 5.3)	(3.7, 6.6)	(2.2, 4.8)	(3.4, 6.1)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Statins						
Total Participant Observations in the Post-Intervention Period	326	347	312	322	638	669
Difference	5.69***	5.75***	5.01***	5.73***	5.37***	5.78***
90% Confidence Interval	(4.1, 7.3)	(4.0, 7.5)	(3.4, 6.6)	(4.0, 7.4)	(3.9, 6.8)	(4.2, 7.3)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

Appendix Table B.21: Difference in Six-Month, Twelve-Month, and Cumulative Means from Baseline Means for Medication Adherence (PDC ≥ 80%), FP Programs

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Beta Blockers						
Total Participant Observations in the Post-Intervention Period	243	254	241	251	484	505
Difference	6.64**	5.12*	7.40***	8.16***	7.03***	6.67***
90% Confidence Interval	(2.0, 11.3)	(0.3, 10.0)	(2.8, 12.1)	(3.6, 12.7)	(2.9, 11.2)	(2.5, 10.8)
P-value	0.02	0.08	0.01	< 0.01	0.01	0.01
Calcium Channel Blockers						
Total Participant Observations in the Post-Intervention Period	163	180	162	170	325	350
Difference	11.27***	10.83***	9.52***	10.34***	10.47***	10.62***
90% Confidence Interval	(5.7, 16.8)	(5.7, 16)	(3.7, 15.3)	(5.2, 15.5)	(5.4, 15.6)	(5.9, 15.3)
P-value	< 0.01	< 0.01	0.01	< 0.01	< 0.01	< 0.01
Diabetes Medication						
Total Participant Observations in the Post-Intervention Period	116	95	107	89	223	184
Difference	0.80	11.81***	2.96	14.11***	1.84	13.16***
90% Confidence Interval	(-5.8, 7.4)	(5.2, 18.5)	(-3.4, 9.3)	(7.9, 20.4)	(-3.8, 7.5)	(7.1, 19.2)
P-value	0.84	< 0.01	0.45	< 0.01	0.59	< 0.01
RAS Antagonists						
Total Participant Observations in the Post-Intervention Period	349	329	322	313	671	642
Difference	6.55***	7.19***	8.39***	9.50***	7.45***	8.37***
90% Confidence Interval	(2.7, 10.4)	(3.5, 10.9)	(4.6, 12.2)	(6.0, 13.0)	(4.0, 10.9)	(5.1, 11.7)
P-value	0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Statins						

Measures	Interim Estimates: 0-6 Months		Interim Estimates: 7-12 Months		Cumulative Estimates	
	Comparison	Participants	Comparison	Participants	Comparison	Participants
Total Participant Observations in the Post-Intervention Period	326	347	312	322	638	669
Difference	10.14***	11.80***	8.56***	12.14***	9.40***	12.09***
90% Confidence Interval	(6.2, 14.1)	(7.7, 15.9)	(4.4, 12.7)	(8.0, 16.3)	(5.8, 13.0)	(8.4, 15.8)
P-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no change over time, the observed single difference could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported difference estimate. The unit of observation is beneficiary-half-years. Six-Month Analysis: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Twelve-Month Analysis: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Cumulative Outcomes: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Single difference models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates are regression-adjusted for these covariates.

APPENDIX C – INTENTION-TO-TREAT (ITT) ANALYSIS TABLES

Appendix C presents additional intention-to-treat (ITT) measure summary statistics and results not reported in the body of the report. Section C.1 presents the weighted characteristics of the survey-based ITT samples. Section C.2 presents summary statistics for all healthcare utilization, expenditure, and medication adherence measures, across priority areas. Section C.3 presents additional healthcare utilization, expenditure, and adherence analyses results across priority areas.

C.1 Survey-Based ITT Analysis Additional Summary Statistics

Appendix Table C.1 presents the weighted survey-based characteristics.

Appendix Table C.1: Weighted Characteristics of the Survey-Based ITT Samples

Characteristic (measured at baseline)	ACA Priority Area					
	CDM		PANO		FP	
	Part. N=529	Comp. N=585	Part. N=656	Comp. N=693	Part. N=1,252	Comp. N=1,339
Average Age[a]	75.6	75.5	74.7	75.2	77.6	77.7
% Female[a]	78.3	79.2	83.6	83.4	77.0	76.3
Race/Ethnicity[a]						
% White	72.1	74.6***	82.1	80.5	89.5	92.0***
% Black/African American	23.7	24.0	14.3	16.2	6.3	5.5
% Hispanic	2.9	0.6	1.2	0.9	2.7	1.0
% Asian	0.3	0.2	0.4	1.2	0.6	0.3
% Native American	0.0	0.0	0.2	0.1	0.2	0.1
% Other	0.9	0.5	1.8	1.1	0.9	1.1
% Urban[a]	70.6	78.1***	85.3	77.2***	71.2	76.4***
% Dual[a]	16.6	16.6	7.1	7.8	11.9	9.2***
Income[b]						
% less than $20,000	57.5	55.0	42.7	45.8	47.6	46.7
% $20,000-$49,999	24.5	26.3	29.3	28.9	30.1	30.1
% $50,000-$99,999	14.5	14.5	22.1	21.0	17.9	18.1
% $100,000 or more	3.6	4.3	5.9	4.3	4.5	5.2
Educational attainment[b]						
% less than high school	19.7	19.2	11.0	15.6**	11.3	10.4
% high school graduate	26.9	29.4	24.6	24.8	29.7	31.1
% some college/2 year degree	41.8	38.2	43.1	40.2	39.9	40.0
% 4 year college graduate or higher	11.7	13.2	21.3	19.3	19.2	18.5

[a] Characteristics are identified through Medicare enrollment data.
[b] Characteristics are identified through baseline national and participant survey data.
Notes: Part.: Wellness program participants. Comp.: Comparison group. *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01. The p-value is the probability that, if there are no differences in characteristics between participants and the comparison group in each priority area, the observed differences could have occurred by chance in the data. Missing data are included in the lowest income and education categories, and among those of "other" race.

C.2 Claims-Based ITT Analysis Summary Statistics

Appendix Table C.2 through Appendix Table C.4 present summary statistics for all measures, by priority area. In the tables below, "index" is defined as the start of the post-intervention period.

Appendix Table C.2: Health Services Utilization and Incidence of Falls and Fractures by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries	529	528	514	517	506	502	593	593	583	584	574	578	1,160	1,161	1,133	1,136	1,097	1,114
Number of Beneficiaries with Nonzero ER Visits	146	147	98	97	106	69	117	128	84	74	60	89	317	294	169	180	187	182
Number of Beneficiaries with Nonzero IP Admissions	86	84	59	45	54	43	66	65	31	42	39	42	174	183	106	119	114	117
Number of Beneficiaries with Nonzero Unplanned IP Admissions	64	63	49	33	47	35	47	44	25	28	26	37	141	142	93	96	97	98
Number of Beneficiaries with Nonzero Lengths of Stay	85	84	58	45	52	43	66	65	30	42	39	42	169	183	104	117	114	116
Number of Beneficiaries with Nonzero Falls/Fractures	62	65	48	72	82	78	42	57	51	68	81	94	178	175	160	186	216	231
Mean Number of Events per 1,000 Beneficiaries																		
ER Visits	451.8	450.8	276.3	232.1	278.7	167.3	312.0	323.8	181.8	172.9	120.2	197.2	405.2	383.3	199.5	213.0	242.5	235.2
All Inpatient Admissions	232.5	208.3	163.4	125.7	140.3	115.5	188.9	163.6	66.9	85.6	94.1	100.3	219.0	208.4	129.7	142.6	133.1	141.8
Unplanned Inpatient Admissions	164.5	155.3	130.4	96.7	124.5	93.6	138.3	102.9	48.0	58.2	67.9	88.2	174.1	155.9	109.4	115.3	109.4	118.5
Length of Stay	1,047.3	1,092.8	723.7	553.2	681.8	639.4	905.6	785.8	274.4	426.4	440.8	510.4	921.6	813.1	555.2	646.1	645.4	686.7
Falls/Fractures	117.2	123.1	93.4	139.3	162.1	155.4	70.8	96.1	87.5	116.4	141.1	162.6	153.4	150.7	141.2	163.7	196.9	207.4

Note: PP = Program Participant; NSR = National Survey Respondent; IP = Inpatient; ER = Emergency Room

Appendix Table C.3: Expenditures by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries	529	528	514	517	506	502	593	593	583	584	574	578	1,160	1,161	1,133	1,136	1,097	1,114
Number of Beneficiaries with Nonzero Part D	372	363	361	362	354	352	344	390	360	398	355	393	760	775	768	773	743	756
Number of Beneficiaries with Nonzero Parts A and B Expenditures	524	520	503	499	492	488	344	390	331	381	326	376	1,145	1,149	1,106	1,096	1,075	1,088
Number of Beneficiaries with Nonzero IP Expenditures	85	84	58	44	53	42	64	63	31	43	39	41	169	180	105	117	111	117

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries with Nonzero OP ER Expenditures	146	145	94	97	103	69	117	127	84	74	60	88	317	293	169	180	186	182
Number of Beneficiaries with Nonzero OP Non-ER Expenditures	425	428	349	347	341	331	454	467	379	371	370	375	905	937	787	749	713	736
Number of Beneficiaries with Nonzero Physician and Ancillary Expenditures	522	520	501	494	490	485	577	580	555	551	559	555	1,143	1,148	1,101	1,089	1,070	1,083
Number of Beneficiaries with Nonzero DME Expenditures	268	233	209	183	213	170	150	179	112	132	117	138	373	396	296	281	285	282
Number of Beneficiaries with Nonzero HH Expenditures	50	50	40	30	35	28	33	41	17	28	28	28	122	107	58	72	61	86
Total Part D Expenditures																		
Mean	$4,042	$4,570	$2,140	$2,290	$2,311	$1,948	$2,495	$2,667	$1,238	$1,232	$1,059	$1,288	$2,951	$2,809	$1,547	$1,386	$1,548	$1,275
Median	$1,996	$1,757	$898	$778	$896	$572	$716	$1,051	$292	$412	$334	$431	$1,156	$1,226	$503	$501	$465	$513
Total Parts A and B Expenditures																		
Mean	$7,633	$7,422	$4,670	$3,670	$4,175	$3,662	$5,331	$5,303	$2,700	$2,926	$2,971	$3,313	$6,613	$6,632	$3,831	$3,633	$3,762	$4,191
Median	$2,960	$3,011	$1,507	$1,155	$1,575	$1,100	$1,953	$2,122	$933	$835	$990	$935	$2,830	$2,997	$1,210	$1,140	$1,177	$1,205
Inpatient Expenditures																		
Mean	$1,786	$1,822	$1,231	$881	$1,078	$1,132	$1,380	$1,295	$603	$699	$702	$737	$1,566	$1,503	$935	$1,004	$1,030	$1,220
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Outpatient ER Expenditures																		
Mean	$275	$264	$173	$122	$153	$92	$208	$157	$101	$129	$52	$97	$226	$202	$108	$132	$132	$120
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Outpatient Non-ER Expenditures																		
Mean	$1,305	$1,259	$611	$699	$725	$601	$893	$878	$475	$580	$444	$622	$1,138	$1,125	$671	$570	$596	$674
Median	$282	$319	$110	$107	$108	$100	$202	$225	$75	$88	$82	$90	$328	$292	$115	$106	$98	$104
Physician and Ancillary Expenditures																		
Mean	$2,967	$2,881	$1,614	$1,385	$1,490	$1,272	$2,286	$2,124	$1,124	$1,077	$1,264	$1,275	$2,544	$2,607	$1,306	$1,303	$1,355	$1,305
Median	$1,838	$1,768	$917	$787	$876	$634	$1,333	$1,349	$596	$573	$664	$630	$1,574	$1,715	$717	$715	$704	$711
Durable Medical Equipment Expenditures																		
Mean	$417	$280	$198	$143	$169	$133	$127	$176	$50	$95	$52	$114	$287	$184	$144	$112	$118	$101
Median	$14	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Home Health Expenditures																		
Mean	$465	$393	$315	$223	$252	$175	$265	$316	$124	$184	$158	$156	$468	$423	$181	$239	$223	$295
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0

Note: PP = Program Participant; NSR = National Survey Respondent; IP = Inpatient; ER = Emergency Room; HH = Home Health; DME = Durable Medical Equipment

Appendix Table C.4: Medication Adherence by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Beta Blockers																		
Number of Eligible Beneficiaries	155	132	126	112	122	111	98	135	88	114	94	128	300	278	254	243	251	241
Median	96.94	95.92	99.71	99.29	100.00	98.28	95.94	95.89	97.75	99.24	98.81	98.95	97.39	95.92	99.09	99.41	98.75	98.36
Rate (PDC ≥ 80)	0.83	0.86	0.93	0.89	0.93	0.91	0.81	0.84	0.91	0.89	0.91	0.85	0.83	0.84	0.88	0.91	0.92	0.91
Calcium Channel Blockers																		
Number of Eligible Beneficiaries	122	104	96	85	94	82	78	103	65	91	74	85	215	197	180	163	170	162
Median	95.65	95.29	100.00	97.53	100.00	97.61	96.89	95.29	98.80	98.34	99.71	99.03	97.42	95.79	99.33	99.38	99.39	99.03
Rate (PDC ≥ 80)	0.75	0.84	0.96	0.85	0.94	0.93	0.87	0.82	0.88	0.89	0.97	0.85	0.83	0.82	0.94	0.93	0.94	0.91
Diabetes Medication																		
Number of Eligible Beneficiaries	98	105	88	97	80	94	50	61	51	59	47	61	115	122	95	116	89	107
Median	97.00	96.32	99.34	99.39	100.00	100.00	95.43	97.54	100.00	97.67	100.00	98.94	97.70	97.44	100.00	99.30	100.00	99.37
Rate (PDC ≥ 80)	0.84	0.85	0.86	0.93	0.94	0.93	0.82	0.90	0.90	0.95	0.98	0.85	0.83	0.89	0.95	0.90	0.98	0.92
RAS Antagonists																		
Number of Eligible Beneficiaries	215	201	164	174	177	165	157	180	134	157	140	154	390	380	329	349	313	322
Median	96.92	96.90	100.00	99.41	99.41	99.17	96.93	96.33	99.40	99.33	99.30	100.00	97.15	96.95	99.45	99.16	99.43	99.39
Rate (PDC ≥ 80)	0.86	0.86	0.90	0.91	0.94	0.94	0.86	0.83	0.93	0.93	0.93	0.95	0.85	0.85	0.93	0.91	0.95	0.93
Statins																		
Number of Eligible Beneficiaries	209	204	167	158	159	153	178	204	143	167	159	159	418	404	347	326	322	312
Median	94.96	94.11	99.12	99.33	100.00	97.77	92.56	95.00	97.98	98.94	98.31	98.66	95.53	95.06	99.22	99.09	98.81	98.55
Rate (PDC ≥ 80)	0.77	0.73	0.88	0.92	0.91	0.90	0.73	0.83	0.89	0.94	0.91	0.92	0.78	0.82	0.91	0.92	0.91	0.90

Note: PP = Program Participant; NSR = National Survey Respondent

C.3 Claims-Based ITT Analysis Additional Results

Appendix Table C.5 through Appendix Table C.13 present additional ITT health services utilization, expenditure, and adherence results by priority area.

Appendix Table C.5: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, CDM Programs

Measures	CDM	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	96/1,020	110/1,020
Difference-in-Difference	53.41	251.36
P-value	0.29	0.52
90% Confidence Interval	(-28.7, 135.5)	(-390.3, 893)
Baseline Participant Mean	168.43	808.00
Intervention Period Participant Mean	257.43	1,163.54
Baseline Comparison Mean	148.99	782.64
Intervention Period Comparison Mean	184.61	887.84
Relative Difference	31.7%	31.1%
Interim Estimates: 0-6 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	49/514	58/514
Difference-in-Difference	28.59	192.26
P-value	0.39	0.39
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	47/506	52/506
Difference-in-Difference	24.74	60.42
P-value	0.41	0.81

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.6: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, CDM Programs

Measures (2011 USD)	CDM		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	715/727	422/1,020	75/1,020
Difference-in-Difference	$666.48	- $45.36	$92.74

Measures (2011 USD)	CDM		
	Total Part D	DME	Home Health
P-value	0.54	0.62	0.51
90% Confidence Interval	(-1,110.8, 2,443.7)	(-197.3, 106.6)	(-140.7, 326.2)
Baseline Participant Mean	$2,972.34	$129.89	$784.40
Intervention Period Participant Mean	$3,411.95	$80.32	$885.08
Baseline Comparison Mean	$3,357.41	- $1.29	$704.20
Intervention Period Comparison Mean	$3,130.55	- $5.51	$712.14
Relative Difference	22.4%	-34.9%	11.8%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	361/368	209/514	40/514
Difference-in-Difference	$74.33	- $14.30	$57.37
P-value	0.91	0.79	0.53
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	354/359	202/506	33/506
Difference-in-Difference	$591.53	- $31.15	$35.78
P-value	0.34	0.55	0.65

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.7: Medication Adherence (Proportion of Days Covered ≥ 80%), CDM Programs

Measures (PDC ≥ 80%)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	248	190	168	341	326
Difference-in-Difference	5.83	13.76**	-1.63	-0.45	-4.85
P-value	0.23	0.02	0.78	0.91	0.32
90% Confidence Interval	(-2.2, 13.8)	(3.9, 23.7)	(-11.2, 7.9)	(-6.9, 6.0)	(-12.8, 3.1)
Baseline Participant Mean	83.27	79.38	90.07	86.44	68.32
Intervention Period Participant Mean	92.69	98.35	96.19	92.07	80.53
Baseline Comparison Mean	87.10	87.53	91.69	85.89	64.61
Intervention Period Comparison Mean	90.72	92.80	99.26	91.98	81.61
Relative Difference	7.0%	17.3%	-1.8%	-0.5%	-7.1%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	126	96	88	164	167
Difference-in-Difference	7.28	18.63***	-5.07	-0.53	-7.06
P-value	0.19	0.01	0.45	0.91	0.19

Measures (PDC ≥ 80%)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	122	94	80	177	159
Difference-in-Difference	4.70	8.77	2.66	-0.33	-2.42
P-value	0.39	0.17	0.67	0.94	0.66

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.8: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, PANO Programs

Measures	PANO	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	51/1,157	69/1,157
Difference-in-Difference	-66.71	-342.89
P-value	0.22	0.30
90% Confidence Interval	(-155.5, 22.1)	(-888.7, 203.0)
Baseline Participant Mean	314.04	1,984.15
Intervention Period Participant Mean	291.69	1,791.69
Baseline Comparison Mean	281.06	1,880.55
Intervention Period Comparison Mean	325.49	2,030.28
Relative Difference	-21.2%	-17.3%
Interim Estimates: 0-6 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	25/583	30/583
Difference-in-Difference	-28.85	-214.63
P-value	0.29	0.22
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	26/574	39/574
Difference-in-Difference	-38.15	-127.25
P-value	0.27	0.56

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline

period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.9: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, PANO Programs

Measures (2011 USD)	PANO		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	715/755	229/1,157	45/1,157
Difference-in-Difference	- $99.35	- $57.74	- $3.87
P-value	0.89	0.31	0.97
90% Confidence Interval	(-1267.7, 1069)	(-150.4, 34.9)	(-183.7, 176.0)
Baseline Participant Mean	$1,896.25	- $57.40	$349.60
Intervention Period Participant Mean	$1,680.48	- $80.61	$361.91
Baseline Comparison Mean	$2,035.49	- $6.25	$376.50
Intervention Period Comparison Mean	$1,919.06	$28.28	$392.68
Relative Difference	-5.2%	100.6%	-1.1%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	360/381	122/583	17/583
Difference-in-Difference	$65.51	- $20.16	- $35.33
P-value	0.88	0.52	0.60
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	355/374	177/574	28/574
Difference-in-Difference	- $165.94	- $37.58	$32.51
P-value	0.69	0.30	0.60

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.10: Medication Adherence (Proportion of Days Covered ≥ 80%), PANO Programs

Measures (PDC ≥ 80%)	PANO				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					

Measures (PDC ≥ 80%)	PANO				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Total Participant Observations in the Post-Intervention Period	182	139	98	274	302
Difference-in-Difference	7.25	0.88	12.17*	-3.92	6.63
P-value	0.22	0.89	0.09	0.37	0.16
90% Confidence Interval	(-2.4, 16.9)	(-9.2, 11.0)	(0.4, 23.9)	(-11.1, 3.2)	(-1.2, 14.4)
Baseline Participant Mean	73.90	82.48	63.27	77.32	78.19
Intervention Period Participant Mean	84.15	87.76	74.87	83.77	94.98
Baseline Comparison Mean	76.68	76.30	74.91	74.72	87.91
Intervention Period Comparison Mean	79.74	81.14	73.97	85.06	98.16
Relative Difference	9.8%	1.1%	19.2%	-5.1%	8.5%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	88	65	51	134	143
Difference-in-Difference	5.58	-6.48	3.06	-2.71	4.76
P-value	0.39	0.38	0.69	0.58	0.37
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	94	74	47	140	159
Difference-in-Difference	9.31	7.51	21.69***	-5.44	8.33
P-value	0.16	0.26	0.01	0.25	0.11

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.11: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, FP Programs

Measures	FP	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	190/2,230	218/2,230
Difference-in-Difference	-32.42	-234.74
P-value	0.33	0.28
90% Confidence Interval	(-86.8, 22.0)	(-588.5, 119.0)
Baseline Participant Mean	209.13	863.20
Intervention Period Participant Mean	253.68	1,144.95
Baseline Comparison Mean	193.61	764.20
Intervention Period Comparison Mean	270.62	1,279.84
Relative Difference	-15.5%	-27.2%
Interim Estimates: 0-6 Months		

Measures	FP	
	Unplanned Inpatient Admissions	Length of Stay
Nonzero/Total Participant Observations in the Post-Intervention Period	93/1,133	104/1,133
Difference-in-Difference	-14.87	-143.67
P-value	0.48	0.28
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	97/1,097	114/1,097
Difference-in-Difference	-17.80	-91.67
P-value	0.39	0.53

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.12: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, FP Programs

Measures (2011 USD)	FP		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	1,511/1,545	581/2,230	119/2,230
Difference-in-Difference	$288.14	- $52.97	- $172.55*
P-value	0.48	0.60	0.07
90% Confidence Interval	(-387.6, 963.8)	(-217.8, 111.8)	(-328.4, -16.7)
Baseline Participant Mean	$4,895.13	$124.22	$775.98
Intervention Period Participant Mean	$5,045.68	$101.86	$711.90
Baseline Comparison Mean	$4,802.79	$16.12	$753.96
Intervention Period Comparison Mean	$4,665.20	$46.73	$862.44
Relative Difference	5.9%	-42.6%	-22.2%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	768/784	296/1,133	58/1,133
Difference-in-Difference	$85.43	- $20.16	- $80.73
P-value	0.74	0.75	0.14
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	743/761	285/1,097	61/1,097
Difference-in-Difference	$205.36	- $33.04	- $92.44
P-value	0.38	0.49	0.12

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates:

comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table C.13: Medication Adherence (Proportion of Days Covered ≥ 80%), FP Programs

Measures (PDC ≥ 80%)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	505	350	184	642	669
Difference-in-Difference	-0.35	0.15	11.32**	0.93	2.69
P-value	0.92	0.97	0.03	0.75	0.393
90% Confidence Interval	(-6.2, 5.5)	(-6.8, 7.1)	(3, 19.6)	(-3.8, 5.7)	(-2.5, 7.9)
Baseline Participant Mean	79.86	76.59	74.52	78.61	70.18
Intervention Period Participant Mean	86.53	87.20	87.68	86.98	82.28
Baseline Comparison Mean	80.17	75.30	80.60	77.60	73.01
Intervention Period Comparison Mean	87.20	85.76	82.44	85.04	82.42
Relative Difference	-0.4%	0.2%	15.2%	1.2%	3.8%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	254	180	95	329	347
Difference-in-Difference	-1.52	-0.44	11.01*	0.64	1.66
P-value	0.71	0.92	0.05	0.84	0.63
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	251	170	89	313	322
Difference-in-Difference	0.76	0.82	11.16**	1.11	3.57
P-value	0.85	0.86	0.04	0.72	0.32

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. DiD models include covariates for urban/rural status, dual eligibility status, gender, race, age, education, and income. Estimates and reported means are regression-adjusted for these covariates. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

APPENDIX D – AVERAGE TREATMENT EFFECT AMONG THE TREATED (ATT) ANALYSIS TABLES

Appendix D presents the survey and claims-based average treatment effect among the treated (ATT) analysis results across priority areas. Section D.1 presents survey-based ATT results. Section D.2 presents claims-based ATT summary statistics for all measures, for beneficiaries enrolled in FFS. Section D.3 presents claims-based ATT analyses for beneficiaries enrolled in FFS on utilization, expenditure, and adherence measures across priority areas.

D.1 Survey-Based ATT Analysis Tables

Appendix Table D.1 through Appendix Table D.10 present the survey-based ATT analysis results by priority area.

Appendix Table D.1: Weighted Characteristics of the Survey-Based ATT Samples

Characteristic (measured at baseline)	ACA Priority Area					
	CDM		PANO		FP	
	Part. N=453	Comp. N=487	Part. N=489	Comp. N=494	Part. N=1,102	Comp. N=1,146
Average Age[a]	74.6	74.9	74.2	74.5	77.0	77.0
% Female[a]	79.0	79.5	80.2	83.0	77.9	76.4
Race/Ethnicity[a]						
% White	76.6	74.1*	85.7	83.8	92.3	92.7
% Black/African American	20.1	24.9	10.8	13.0	4.9	5.0
% Hispanic	1.8	0.4	0.4	0.8	1.5	0.9
% Asian	0.4	0.2	0.6	1.2	0.5	0.4
% Native American	0.0	0.0	0.2	0.2	0.1	0.1
% Other	1.1	0.4	2.3	1.0	0.6	1.1
% Urban[a]	69.8	78.6***	85.3	77.7***	70.4	76.2***
% Dual[a]	13.0	16.6	3.7	5.9	8.3	8.0
Income[b]						
% less than $20,000	52.3	54.6	34.0	41.9**	46.7	44.1
% $20,000-$49,999	25.8	26.3	31.5	30.0	31.9	31.2
% $50,000-$99,999	17.4	14.6	27.4	23.1	19.3	19.2
% $100,000 or more	4.4	4.5	7.2	5.1	5.1	5.5
Educational attainment[b]						
% less than high school	15.5	17.9	7.0	13.4***	8.1	8.7
% high school graduate	24.9	27.5	22.7	24.3	29.5	31.0
% some college/2 year degree	46.1	41.1	43.4	41.9	41.7	41.5
% 4 year college graduate or higher	13.5	13.6	27.0	20.5	20.8	18.9

[a] Characteristics are identified through Medicare enrollment data.
[b] Characteristics are identified through baseline national and participant survey data.
Notes: Part.: Wellness program participants. Comp.: Comparison group. *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01. The p-value is the probability that, if there are no differences in characteristics between participants and the comparison group in each priority area, the observed differences could have occurred by chance in the data. Missing data are included in the lowest income and education categories, and among those of "other" race.

Appendix Table D.2: DiD Statistics for Physical Health Measures in Chronic Disease Management Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	434/478	453/480	432/480	434/484	453/487
Difference-in-Difference	-0.14	0.29	0.04	-0.19	-0.13
P-value	0.79	0.60	0.94	0.76	0.82
90% Confidence Interval	(-1.0,0.7)	(-0.6,1.2)	(-0.8,0.9)	(-1.3,0.9)	(-1.1,0.8)
Baseline Participant Mean	41.8	40.1	42.3	44.4	48.0
Twelve-Month Participant Mean	41.4	39.8	41.8	44.6	47.1
Baseline Comparison Mean	40.9	39.9	41.4	44.4	48.4
Twelve-Month Comparison Mean	40.6	39.3	40.9	44.8	47.7
Relative Difference	-0.3%	0.7%	0.1%	-0.4%	-0.3%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	437/481	453/481	436/481	436/487	453/487
Difference-in-Difference	-0.03	0.54	0.35	-0.30	0.53
P-value	0.95	0.31	0.54	0.61	0.34
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	419/473	453/474	418/474	420/484	453/487
Difference-in-Difference	0.12	-0.27	0.00	0.16	-0.66
P-value	0.79	0.59	1.00	0.75	0.25

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.3: DiD Statistics for Mental Health Measures in Chronic Disease Management Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	434/478	453/485	453/485	432/477	453/485
Difference-in-Difference	0.05	0.13	-0.55	0.58	0.17
P-value	0.93	0.82	0.43	0.44	0.79
90% Confidence Interval	(-1.0,1.1)	(-0.8,1.0)	(-1.7,0.6)	(-0.6,1.8)	(-0.9,1.2)
Baseline Participant Mean	51.0	49.2	47.2	44.4	50.8
Twelve-Month Participant Mean	50.5	48.3	47.0	44.2	50.5

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Baseline Comparison Mean	51.8	49.6	46.6	45.4	51.8
Twelve-Month Comparison Mean	51.2	48.6	47.0	44.6	51.3
Relative Difference	0.1%	0.3%	-1.2%	1.3%	0.3%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	437/481	452/487	453/487	436/478	452/487
Difference-in-Difference	0.68	-0.16	-0.62	1.46**	1.06*
P-value	0.27	0.80	0.36	0.05	0.05
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	419/473	452/485	453/485	418/472	452/485
Difference-in-Difference	-0.54	0.25	0.06	-0.76	-0.90
P-value	0.39	0.63	0.93	0.30	0.14

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.4: DiD Statistics for Activity, Balance, and Medication Adherence Measures in Chronic Disease Management Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	440/468	420/452	412/446	300/325	392/438
Difference-in-Difference	-0.13	0.01	0.05	4.59***	-0.00
P-value	0.38	0.84	0.18	0.01	0.99
90% Confidence Interval	(-0.4,0.1)	(-0.1,0.1)	(-0.0,0.1)	(1.7,7.5)	(-0.2,0.1)
Baseline Participant Mean	4.9	0.6	0.2	52.2	3.1
Twelve-Month Participant Mean	4.6	0.5	0.2	53.1	3.2
Baseline Comparison Mean	4.4	0.5	0.2	58.4	3.0
Twelve-Month Comparison Mean	4.4	0.5	0.2	54.7	3.2
Relative Difference	-2.7%	1.8%	25.3%	8.8%	0.0%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	444/466	420/456	408/446	295/322	400/430
Difference-in-Difference	-0.16	0.00	-0.01	2.97*	-0.07
P-value	0.26	0.97	0.78	0.08	0.42
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	443/477	425/467	416/455	318/328	399/438
Difference-in-Difference	-0.02	0.01	0.04	1.31	0.08

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
P-value	0.88	0.88	0.25	0.39	0.32

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

Appendix Table D.5: DiD Statistics for Physical Health Measures in Physical Activity, Nutrition, and Obesity Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	465/479	489/483	465/482	464/489	489/494
Difference-in-Difference	0.02	0.60	0.25	0.54	-0.39
P-value	0.96	0.17	0.65	0.37	0.40
90% Confidence Interval	(-0.7,0.7)	(-0.1,1.3)	(-0.7,1.2)	(-0.4,1.5)	(-1.2,0.4)
Baseline Participant Mean	46.0	44.8	46.0	47.5	52.9
Twelve-Month Participant Mean	45.6	44.7	46.2	48.1	51.8
Baseline Comparison Mean	46.0	45.0	45.6	47.6	52.1
Twelve-Month Comparison Mean	45.6	44.3	45.5	47.7	51.4
Relative Difference	0.0%	1.3%	0.5%	1.1%	-0.7%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	472/484	489/487	472/487	472/489	489/494
Difference-in-Difference	0.31	0.38	0.83*	0.54	-0.42
P-value	0.45	0.40	0.08	0.37	0.36
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	449/470	489/476	449/475	448/484	489/494
Difference-in-Difference	-0.22	0.31	-0.45	-0.04	0.03
P-value	0.62	0.46	0.39	0.95	0.94

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.6: DiD Statistics for Mental Health Measures in Physical Activity, Nutrition, and Obesity Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	465/479	489/490	489/490	465/483	489/490
Difference-in-Difference	1.02*	0.73	1.33**	0.97	0.70
P-value	0.09	0.14	0.03	0.16	0.22
90% Confidence Interval	(0.0,2.0)	(-0.1,1.5)	(0.3,2.3)	(-0.2,2.1)	(-0.2,1.6)
Baseline Participant Mean	53.3	52.2	50.3	47.9	53.3
Twelve-Month Participant Mean	53.7	51.7	50.5	49.0	53.6
Baseline Comparison Mean	52.8	52.6	50.3	47.5	52.8
Twelve-Month Comparison Mean	52.2	51.4	49.1	47.6	52.3
Relative Difference	1.9%	1.4%	2.6%	2.0%	1.3%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	472/484	489/491	488/491	471/487	489/491
Difference-in-Difference	0.91	0.48	1.65***	1.05*	0.15
P-value	0.10	0.29	0.00	0.09	0.76
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	449/470	489/487	488/487	448/476	489/487
Difference-in-Difference	0.14	0.23	-0.38	-0.13	0.54
P-value	0.80	0.62	0.52	0.81	0.30

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.7: DiD Statistics for Activity, Balance, and Medication Adherence Measures in Physical Activity, Nutrition, and Obesity Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	485/485	446/464	453/471	335/346	404/417
Difference-in-Difference	0.32***	0.16***	-0.01	1.78	0.06
P-value	0.01	0.00	0.60	0.26	0.39
90% Confidence Interval	(0.1,0.5)	(0.1,0.2)	(-0.1,0.0)	(-0.8,4.4)	(-0.1,0.2)
Baseline Participant Mean	5.1	0.7	0.2	65.8	3.1
Twelve-Month Participant Mean	5.1	0.8	0.2	64.9	3.3
Baseline Comparison Mean	5.1	0.7	0.2	68.1	3.2

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Twelve-Month Comparison Mean	4.8	0.6	0.2	65.5	3.3
Relative Difference	6.3%	22.6%	-5.7%	2.7%	1.9%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	483/484	447/465	452/473	346/329	410/412
Difference-in-Difference	0.42***	0.16***	-0.01	0.58	-0.07
P-value	0.00	0.00	0.78	0.71	0.30
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	487/481	449/465	455/470	352/346	405/419
Difference-in-Difference	-0.12	-0.02	-0.01	0.87	0.10
P-value	0.22	0.48	0.73	0.48	0.12

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

Appendix Table D.8: DiD Statistics for Physical Health Measures in Falls Prevention Programs

Measures	Physical Components Summary Score	Physical Functioning Subscale	Role Physical Subscale	Bodily Pain Subscale	General Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	1,044/1,118	1,100/1,127	1,044/1,126	1,046/1,134	1,097/1,146
Difference-in-Difference	0.18	0.29	0.72**	0.40	-0.11
P-value	0.58	0.41	0.04	0.21	0.67
90% Confidence Interval	(-0.4,0.7)	(-0.3,0.9)	(0.2,1.3)	(-0.1,0.9)	(-0.5,0.3)
Baseline Participant Mean	42.2	40.6	42.4	45.3	49.7
Twelve-Month Participant Mean	41.6	40.0	42.5	45.3	48.8
Baseline Comparison Mean	43.2	41.4	43.2	46.2	50.1
Twelve-Month Comparison Mean	42.4	40.5	42.5	45.8	49.2
Relative Difference	0.4%	0.7%	1.7%	0.9%	-0.2%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,060/1,116	1,102/1,124	1,060/1,123	1,061/1,133	1,100/1,145
Difference-in-Difference	0.12	0.35	0.91***	0.61*	-0.21
P-value	0.68	0.24	0.01	0.08	0.42
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,007/1,090	1,100/1,105	1,007/1,103	1,011/1,128	1,095/1,145
Difference-in-Difference	0.11	-0.05	-0.17	-0.11	0.08
P-value	0.68	0.84	0.61	0.76	0.75

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10%

level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Physical Functioning" assesses performance of physical activities such as self-care and walking. "Bodily Pain" assesses level of pain and limitations due to pain. "Role Physical" assesses limitations to performing work and other activities. "General Health" assesses respondents' evaluation of their overall health. The "Physical Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.9: DiD Statistics for Mental Health Measures in Falls Prevention Programs

Measures	Mental Components Summary Score	Vitality Subscale	Social Functioning Subscale	Role Emotional Subscale	Mental Health Subscale
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	1,044/1,118	1,101/1,137	1,101/1,137	1,042/1,125	1,101/1,137
Difference-in-Difference	0.69**	0.18	0.43	1.24***	0.32
P-value	0.05	0.55	0.22	0.00	0.34
90% Confidence Interval	(0.1,1.3)	(-0.3,0.7)	(-0.2,1.0)	(0.5,1.9)	(-0.2,0.9)
Baseline Participant Mean	52.0	49.8	48.3	45.5	51.9
Twelve-Month Participant Mean	52.3	49.4	48.2	46.0	52.0
Baseline Comparison Mean	52.0	50.2	48.1	46.1	52.0
Twelve-Month Comparison Mean	51.6	49.6	47.7	45.4	51.7
Relative Difference	1.3%	0.4%	0.9%	2.7%	0.6%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,060/1,116	1,100/1,137	1,097/1,136	1,052/1,121	1,101/1,137
Difference-in-Difference	1.03***	0.14	0.71*	1.68***	0.68**
P-value	0.01	0.66	0.09	0.00	0.04
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,007/1,090	1,099/1,129	1,096/1,128	1,002/1,102	1,100/1,129
Difference-in-Difference	-0.51	0.01	-0.29	-0.51	-0.36
P-value	0.12	0.96	0.45	0.23	0.24

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.
Measures: "Vitality" assesses a person's feelings of energy. "Social Functioning" assesses whether mental health problems interfere with social activities. "Role Emotional" assesses role limitations related to mental health. The "Mental Components Summary Score" is a composite consisting of these four areas.

Appendix Table D.10: DiD Statistics for Activity, Balance, and Medication Adherence Measures in Falls Preventions Programs

Measures	Aerobic Activity	Strength and Flexibility	Any Falls in Past 6 Months	Confidence in Balance Scale	Medication Adherence
Cumulative Estimates					
Number of Beneficiaries (Participants/Comparators)	1,074/1,107	1,000/1,038	996/1,053	788/774	931/1,000
Difference-in-Difference	-0.16**	0.04*	0.02	3.83***	-0.02
P-value	0.04	0.08	0.41	0.00	0.60
90% Confidence Interval	(-0.3,-0.0)	(0.0,0.1)	(-0.0,0.1)	(2.2,5.5)	(-0.1,0.0)
Baseline Participant Mean	4.8	0.6	0.3	51.4	3.2
Twelve-Month Participant Mean	4.5	0.6	0.3	52.0	3.2
Baseline Comparison Mean	4.6	0.5	0.3	56.2	3.3
Twelve-Month Comparison Mean	4.5	0.5	0.2	53.0	3.3
Relative Difference	-3.4%	6.9%	6.1%	7.5%	-0.6%
Interim Estimates: Baseline-6 Months					
Number of Beneficiaries (Participants/Comparators)	1,076/1,094	1,012/1,036	1,013/1,047	816/792	945/1,005
Difference-in-Difference	-0.23***	0.03	0.02	2.61**	0.01
P-value	0.00	0.11	0.32	0.01	0.85
Interim Estimates: 6-12 Months					
Number of Beneficiaries (Participants/Comparators)	1,077/1,117	1,014/1,089	1,017/1,077	806/815	948/1,027
Difference-in-Difference	0.06	0.00	-0.01	0.90	-0.02
P-value	0.44	0.88	0.75	0.34	0.69

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. DiD estimates and reported means are adjusted for age, gender, race, urban residence, dual status, income, and education. Relative difference is calculated as the DiD estimate divided by the participant mean at the beginning of the time frame and is expressed as a percentage. The number of beneficiaries reported for interim estimates and cumulative estimates vary slightly due to missing data for each pairwise comparison.

D.2 Claims-Based ATT Analysis Summary Statistics

Appendix Table D.11 through Appendix Table D.14 present FFS cohort summary statistics for all measures, by priority area.

Appendix Table D.11: Baseline Demographic Summary Statistics, ATT Analysis

Characteristic	FFS Beneficiaries Included in Claims-based Analysis					
	CDM		PANO		FP	
	PP N=309	NSR N=347	PP N=327	NSR N=370	PP N=746	NSR N=812
Average Age	74.9	75.5	74.6	74.8	77.4	77.2
% Female	76.4	77.5	81.7	81.4	78.2	76.1
Race						
% White	86.1*	82.7	85.3	84.6	93.0	92.6
% Black	11.7*	16.4	12.2	13.0	4.2	5.3
% Other	2.3*	0.9	2.4	2.4	2.8	2.1
% Dual Eligible	7.4**	12.7	3.7	4.9	7.1	8.0
% Urban	59.5***	74.4	86.2***	75.1	69.0	72.8
Evaluation and Management (E&M) Visits						
% E&M Visits: 0	2.3	2.0	4.6	4.6	2.7	3.0
% E&M Visits: 1-10	54.4	56.5	67.9	70.0	62.1	59.7
% E&M Visits: 11+	43.4	41.5	27.5	25.4	35.3	37.3
IP Stays						
% 0 IP Stays (Prior Year)	86.7*	83.3	91.1	89.2	86.5	85.3
% 1 IP Stay (Prior Year)	9.4*	14.4	6.4	6.2	10.6	11.2
% 2+ IP Stays (Prior Year)	3.9*	2.3	2.4	4.6	2.9	3.4
ER Visits						
% ER Visits: 0	76.1	73.5	83.2	80.5	73.6	76.4
% ER Visits: 1	15.9	15.0	12.8	14.1	18.1	16.7
% ER Visits: 2+	8.1	11.5	4.0	5.4	8.3	6.9
Total Parts A and B Cost per Beneficiary	$8,159	$8,886	$5,420	$6,320	$6,997	$7,322
IP Cost per Beneficiary	$1,810	$1,954	$1,093	$1,676	$1,536	$1,576
Part D Cost per Beneficiary	$3,325*	$5,191	$2,216	$2,573	$2,436	$3,056

Notes: Part.: Program participants. Comp.: Comparison group. IP: Inpatient; ER: Emergency Room. *p-value< 0.10; ** p-value< 0.05; ***p-value< 0.01. The p-value is the probability that, if there are no differences in characteristics between participants and the comparison group in each priority area, the observed differences could have occurred by chance in the data. Part D cost per beneficiary only accounts for beneficiaries who have Part D coverage. E&M visits do not include annual wellness visits or visits to FQHCs.

Appendix Table D.12: Health Services Utilization by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries	309	347	302	339	299	333	327	370	321	368	319	365	746	812	735	802	720	789
Number of Beneficiaries with Nonzero ER Visits	72	92	51	61	57	48	52	72	38	46	26	60	192	192	105	124	119	131
Number of Beneficiaries with Nonzero IP Admissions	41	58	32	22	30	26	29	40	17	28	25	25	101	119	50	69	56	80
Number of Beneficiaries with Nonzero Unplanned IP Admissions	28	40	28	13	24	21	19	28	13	18	16	22	75	95	43	57	45	67
Number of Beneficiaries with Nonzero Lengths of Stay	40	58	32	22	28	26	29	40	17	28	25	25	98	119	49	68	56	79
Number of Beneficiaries with Nonzero Falls/Fractures	26	45	27	47	56	55	16	35	25	43	40	62	98	126	89	126	130	165
Mean Number of Events per 1,000 Beneficiaries																		
ER Visits	385.1	466.9	235.1	200.6	257.5	174.2	204.9	310.8	158.9	187.5	90.9	213.7	257.4	236.5	142.9	154.6	165.3	166.0
All Inpatient Admissions	207.1	193.1	142.4	82.6	120.4	93.1	113.1	170.3	62.3	84.2	90.9	109.6	179.6	192.1	84.4	112.2	104.2	134.3
Unplanned Inpatient Admissions	142.4	138.3	119.2	50.1	97.0	75.1	67.3	108.1	43.6	51.6	56.4	95.9	132.7	149.0	69.4	89.8	80.6	112.8
Length of Stay	828.5	899.1	447.0	348.1	478.3	378.4	452.6	864.9	186.9	385.9	341.7	504.1	705.1	703.2	296.6	487.5	495.8	588.1
Falls/Fractures	84.1	129.7	89.4	138.6	187.3	165.2	48.9	94.6	77.9	116.8	125.4	169.9	131.4	155.2	121.1	157.1	180.6	209.1

Note: PP = Program Participant; NSR = National Survey Respondent; IP = Inpatient; ER = Emergency Room

Appendix Table D.13: Expenditures by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries	309	347	302	339	299	333	327	370	321	368	319	365	746	812	735	802	720	789
Number of Beneficiaries with Nonzero Part D	216	244	214	240	213	235	179	240	189	248	188	244	489	546	497	543	485	533
Number of Beneficiaries with Nonzero Parts A and B Expenditures	308	341	298	329	293	324	322	364	311	349	311	353	739	805	719	777	706	770
Number of Beneficiaries with Nonzero IP Expenditures	41	58	31	22	30	26	28	38	17	29	25	24	99	116	50	67	53	80
Number of Beneficiaries with Nonzero OP ER Expenditures	72	91	51	61	57	48	52	71	38	46	26	60	192	191	105	124	118	131
Number of Beneficiaries with Nonzero OP Non-ER Expenditures	260	283	206	228	210	221	248	289	202	236	203	233	579	667	504	531	467	526
Number of Beneficiaries with Nonzero Physician and Ancillary Expenditures	307	341	297	325	292	321	321	363	308	345	310	350	738	804	717	770	703	767

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Beneficiaries with Nonzero DME Expenditures	158	156	124	125	133	119	88	113	64	86	76	89	242	278	184	202	188	196
Number of Beneficiaries with Nonzero HH Expenditures	18	33	17	15	13	17	10	25	6	19	14	16	59	64	24	43	32	57
Total Part D Expenditures																		
Mean	$3,325	$5,191	$2,201	$2,594	$2,309	$2,154	$2,216	$2,573	$896	$1,122	$961	$1,107	$2,436	$3,056	$1,410	$1,508	$1,367	$1,366
Median	$1,798	$1,757	$835	$733	$850	$528	$729	$1,139	$271	$448	$358	$471	$1,004	$1,222	$424	$531	$382	$531
Total Parts A and B Expenditures																		
Mean	$6,811	$7,431	$3,778	$3,057	$3,851	$3,046	$4,524	$5,273	$2,535	$2,897	$2,990	$3,356	$5,838	$6,112	$3,112	$3,212	$3,243	$3,943
Median	$2,960	$3,178	$1,520	$1,225	$1,631	$1,161	$1,833	$2,061	$854	$852	$990	$938	$2,651	$3,023	$1,131	$1,111	$1,130	$1,208
Inpatient Expenditures																		
Mean	$1,511	$1,644	$802	$591	$800	$724	$913	$1,399	$541	$719	$748	$747	$1,284	$1,318	$594	$762	$763	$1,043
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Outpatient ER Expenditures																		
Mean	$209	$302	$183	$81	$130	$106	$122	$159	$102	$144	$36	$114	$229	$187	$90	$99	$132	$128
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Outpatient Non-ER Expenditures																		
Mean	$1,186	$1,257	$528	$685	$800	$507	$771	$749	$533	$547	$395	$709	$1,124	$1,174	$628	$620	$546	$739
Median	$295	$417	$113	$106	$148	$108	$175	$216	$66	$74	$82	$91	$314	$304	$108	$103	$94	$107
Physician and Ancillary Expenditures																		
Mean	$2,967	$3,090	$1,550	$1,369	$1,517	$1,245	$2,368	$2,160	$1,181	$1,109	$1,329	$1,338	$2,433	$2,501	$1,169	$1,224	$1,292	$1,308
Median	$1,872	$1,888	$950	$817	$1,007	$651	$1,279	$1,427	$554	$581	$692	$656	$1,506	$1,688	$671	$710	$653	$725
Durable Medical Equipment Expenditures																		
Mean	$354	$277	$166	$115	$159	$118	$136	$145	$58	$79	$54	$108	$194	$161	$93	$81	$118	$86
Median	$14	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0
Home Health Expenditures																		
Mean	$278	$362	$210	$156	$186	$154	$99	$297	$65	$199	$118	$130	$305	$328	$98	$193	$172	$258
Median	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0	$0

Note: PP = Program Participant; NSR = National Survey Respondent; IP = Inpatient; ER = Emergency Room; HH = Home Health; DME = Durable Medical Equipment

Appendix Table D.14: Average Medication Adherence by Priority Area

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Beta Blockers																		

Measures	CDM						PANO						FP					
	Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months		Baseline		0 - 6 Months		7 - 12 Months	
	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR	PP	NSR
Number of Eligible Beneficiaries	95	86	83	74	77	70	49	84	45	69	50	82	183	196	159	171	163	174
Mean	89.29	89.45	95.37	93.93	94.37	94.09	91.06	88.82	95.14	93.98	94.27	93.70	91.02	90.93	94.28	94.08	94.95	94.43
Median	96.71	94.61	100.00	99.26	100.00	99.40	96.80	95.25	99.18	99.40	99.27	99.14	97.85	96.23	100.00	99.41	100.00	98.95
25th percentile	86.96	87.17	95.97	93.15	94.34	92.25	89.10	85.90	96.07	93.33	93.44	93.94	89.97	88.22	94.06	93.15	94.12	93.38
75th percentile	99.64	99.43	100.00	100.00	100.00	100.00	99.13	99.52	100.00	100.00	100.00	100.00	99.69	99.44	100.00	100.00	100.00	100.00
Rate (PDC ≥ 80)	0.83	0.85	0.92	0.88	0.91	0.93	0.84	0.82	0.96	0.88	0.90	0.89	0.86	0.86	0.91	0.89	0.94	0.93
Calcium Channel Blockers																		
Number of Eligible Beneficiaries	74	71	59	57	59	54	39	66	34	57	35	55	134	147	113	123	108	121
Mean	85.15	89.59	96.33	91.08	95.88	94.10	92.31	87.72	93.49	95.51	97.87	93.39	89.76	89.58	95.07	95.63	96.17	94.44
Median	95.19	95.88	100.00	98.32	100.00	97.19	96.99	95.47	98.97	99.38	100.00	98.99	96.05	95.58	99.39	99.41	99.39	98.82
25th percentile	75.84	87.25	96.88	92.00	97.58	92.35	90.79	87.21	91.00	94.58	97.81	93.30	89.50	87.25	94.74	95.00	95.22	94.00
75th percentile	99.19	99.44	100.00	100.00	100.00	100.00	99.45	99.20	100.00	100.00	100.00	100.00	99.41	99.13	100.00	100.00	100.00	100.00
Rate (PDC ≥ 80)	0.70	0.87	0.97	0.86	0.95	0.93	0.87	0.80	0.91	0.91	1.00	0.89	0.82	0.83	0.94	0.95	0.96	0.92
Diabetes Medication																		
Number of Eligible Beneficiaries	53	70	49	67	45	65	26	38	25	37	26	39	79	92	66	87	64	78
Mean	90.25	92.07	93.85	95.05	93.85	94.92	92.13	90.25	94.65	92.74	97.12	91.35	91.00	91.16	95.84	94.40	95.38	94.67
Median	97.60	96.92	98.97	99.32	100.00	100.00	95.43	95.65	100.00	98.20	100.00	99.05	98.03	97.32	100.00	99.40	100.00	99.38
25th percentile	90.97	89.74	91.80	92.90	93.75	96.00	86.87	88.28	97.93	87.88	96.59	93.41	90.03	89.91	96.69	93.67	91.53	94.89
75th percentile	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.41	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Rate (PDC ≥ 80)	0.85	0.86	0.90	0.94	0.91	0.91	0.81	0.87	0.92	0.92	1.00	0.85	0.84	0.90	0.92	0.92	0.98	0.94
RAS Antagonists																		
Number of Eligible Beneficiaries	127	134	95	113	105	111	87	115	68	100	73	98	246	272	216	254	210	231
Mean	91.69	91.30	95.22	93.90	95.59	95.00	89.72	90.40	95.13	94.98	95.84	95.67	91.12	91.73	94.95	94.07	95.74	94.80
Median	97.51	97.07	100.00	99.39	100.00	99.07	95.51	96.34	98.56	99.72	100.00	99.20	97.27	97.20	100.00	99.22	100.00	100.00
25th percentile	90.16	90.09	96.09	95.27	96.15	94.67	86.34	86.76	94.47	93.81	97.02	95.29	89.77	88.53	96.87	93.25	95.45	94.58
75th percentile	100.00	100.00	100.00	100.00	100.00	100.00	100.00	99.72	100.00	100.00	100.00	100.00	99.68	100.00	100.00	100.00	100.00	100.00
Rate (PDC ≥ 80)	0.90	0.87	0.92	0.88	0.92	0.93	0.82	0.84	0.93	0.93	0.95	0.95	0.87	0.86	0.92	0.91	0.95	0.94
Statins																		
Number of Eligible Beneficiaries	124	138	98	106	99	98	97	129	77	109	89	109	271	287	234	233	213	231
Mean	88.26	84.67	94.79	93.59	94.18	93.66	86.25	88.74	93.30	94.12	93.56	95.39	87.49	89.09	93.64	94.89	93.83	94.49
Median	95.81	92.55	99.71	99.38	100.00	98.02	92.46	95.54	97.66	98.73	97.94	99.40	95.36	95.06	99.40	99.27	98.80	98.83
25th percentile	82.74	75.18	95.38	93.14	93.43	92.25	78.16	86.30	89.44	92.49	91.57	93.75	83.29	87.18	93.75	94.74	92.54	93.84
75th percentile	99.29	98.62	100.00	100.00	100.00	100.00	96.90	99.16	100.00	100.00	100.00	100.00	99.09	99.31	100.00	100.00	100.00	100.00
Rate (PDC ≥ 80)	0.77	0.71	0.91	0.92	0.89	0.89	0.72	0.84	0.91	0.93	0.93	0.94	0.79	0.82	0.89	0.93	0.91	0.92

Note: PP = Program Participant; NSR = National Survey Respondent

D.3 Claims-Based ATT Analysis Results

Appendix Table D.15 through Appendix Table D.32 present ATT analytic results on health services utilization, expenditure, and adherence outcomes by priority area for the sample of beneficiaries enrolled in FFS. These reported estimates correspond to DiD models without covariates; DiD models with covariates were not feasible due to low sample sizes.

Appendix Table D.15: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, CDM Programs

Measures	CDM		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	108/601	62/601	83/601
Difference-in-Difference	199.36**	73.22	31.81
P-value	0.04	0.25	0.31
90% Confidence Interval	(41.3, 357.4)	(-31.2, 177.6)	(-19.2, 82.8)
Baseline Participant Mean	385.11	207.12	84.14
Intervention Period Participant Mean	492.51	262.79	138.10
Baseline Comparison Mean	466.86	193.08	129.68
Intervention Period Comparison Mean	374.90	175.69	151.84
Relative Difference	51.8%	35.4%	37.8%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observation in the Post-Intervention Period s	51/302	32/302	27/302
Difference-in-Difference	75.38	52.77	-3.70
P-value	0.18	0.19	0.92
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	57/299	30/299	56/299
Difference-in-Difference	124.22**	20.29	67.67*
P-value	0.03	0.58	0.08

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.16: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, CDM Programs

Measures	CDM	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	52/601	60/601
Difference-in-Difference	87.14	269.47
P-value	0.12	0.41
90% Confidence Interval	(-5.5, 179.8)	(-272.3, 811.2)
Baseline Participant Mean	142.39	828.48
Intervention Period Participant Mean	216.20	925.28
Baseline Comparison Mean	138.33	899.14
Intervention Period Comparison Mean	125.22	726.46
Relative Difference	61.2%	32.5%
Interim Estimates: 0-6 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	28/302	32/302
Difference-in-Difference	67.02*	134.27
P-value	0.06	0.50
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	24/299	28/299
Difference-in-Difference	19.88	135.21
P-value	0.55	0.48

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.17: Medicare Expenditures per Beneficiary, CDM Programs

Measures (2011 USD)	CDM				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	591/601	61/601	108/601	416/601	589/601
Difference-in-Difference	$2,145.17*	$420.69	$219.42**	$203.34	$575.36
P-value	0.06	0.41	0.02	0.52	0.18
90% Confidence Interval	(251.2, 4039.1)	(-420.7, 1262)	(64.1, 374.7)	(-315.2, 721.9)	(-136.7, 1287.4)
Baseline Participant Mean	$6,810.74	$1,511.36	$208.62	$1,186.19	$2,966.70
Intervention Period Participant Mean	$7,628.28	$1,602.71	$313.37	$1,325.73	$3,066.75
Baseline Comparison Mean	$7,431.43	$1,644.17	$301.75	$1,256.85	$3,090.23
Intervention Period Comparison Mean	$6,103.79	$1,314.84	$187.09	$1,193.05	$2,614.92
Relative Difference	31.5%	27.8%	105.2%	17.1%	19.4%

Measures (2011 USD)	CDM				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	298/302	31/302	51/302	206/302	297/302
Difference-in-Difference	$1,030.45	$277.34	$148.6**	- $122.55	$241.85
P-value	0.13	0.35	0.02	0.57	0.34
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	293/299	30/299	57/299	210/299	292/299
Difference-in-Difference	$1,115.14	$142.67	$70.43	328.14*	$333.97
P-value	0.10	0.65	0.14	0.06	0.16

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.18: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, CDM Programs

Measures (2011 USD)	CDM		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	427/433	257/601	30/601
Difference-in-Difference	$1,554.26	$15.57	$171.38
P-value	0.27	0.85	0.27
90% Confidence Interval	(-779.3, 3887.8)	(-122.2, 153.4)	(-82.2, 425)
Baseline Participant Mean	$2,208.84	$353.97	$277.68
Intervention Period Participant Mean	$3,451.44	$325.11	$396.47
Baseline Comparison Mean	$3,690.38	$277.45	$362.48
Intervention Period Comparison Mean	$3,378.71	$233.01	$309.88
Relative Difference	70.4%	4.4%	61.7%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	214/218	124/302	17/302
Difference-in-Difference	$481.12	$12.68	$96.55
P-value	0.59	0.79	0.31
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	213/215	133/299	13/299
Difference-in-Difference	$1,053.00	$2.84	$74.72
P-value	0.20	0.95	0.41

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and

reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.19: Medication Adherence (Average Proportion of Days Covered), CDM Programs

Measures (Average PDC)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	160	118	94	200	197
Difference-in-Difference	1.04	7.96**	0.69	0.55	- 2.74
P-value	0.70	0.02	0.82	0.79	0.25
90% Confidence Interval	(-3.5, 5.5)	(2.4, 13.5)	(-4.3, 5.7)	(-2.8, 3.9)	(-6.6, 1.2)
Baseline Participant Mean	89.29	85.15	90.25	91.69	88.26
Intervention Period Participant Mean	94.87	96.10	93.85	95.41	94.49
Baseline Comparison Mean	89.45	89.59	92.07	91.30	84.67
Intervention Period Comparison Mean	94.01	92.59	94.98	94.45	93.63
Relative Difference	1.2%	9.3%	0.8%	0.6%	-3.1%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	83	59	49	95	98
Difference-in-Difference	1.60	9.69**	0.63	0.93	- 2.40
P-value	0.59	0.01	0.85	0.70	0.36
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	77	59	45	105	99
Difference-in-Difference	0.44	6.23*	0.75	0.20	- 3.07
P-value	0.88	0.07	0.83	0.93	0.24

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.20: Medication Adherence (Proportion of Days Covered ≥ 80%), CDM Programs

Measures (PDC ≥ 80%)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	160	118	94	200	197

Measures (PDC ≥ 80%)	CDM				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Difference-in-Difference	2.72	23.54***	- 1.23	- 1.95	- 6.69
P-value	0.67	0.00	0.87	0.69	0.28
90% Confidence Interval	(-7.8, 13.2)	(11.3, 35.8)	(-13.6, 11.1)	(-9.9, 6.0)	(-16.8, 3.4)
Baseline Participant Mean	83.16	70.27	84.91	89.76	77.42
Intervention Period Participant Mean	91.24	95.76	90.45	91.98	89.85
Baseline Comparison Mean	84.88	87.32	85.71	86.57	71.01
Intervention Period Comparison Mean	90.35	89.28	92.40	90.64	90.14
Relative Difference	3.3%	33.5%	-1.4%	-2.2%	-8.6%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	83	59	49	95	98
Difference-in-Difference	5.45	27.70***	- 3.43	- 0.11	- 7.10
P-value	0.46	0.00	0.68	0.98	0.29
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	77	59	45	105	99
Difference-in-Difference	- 0.22	19.38**	1.15	- 3.61	- 6.29
P-value	0.98	0.02	0.89	0.50	0.37

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.21: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, PANO Programs

Measures	PANO		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	64/640	42/640	65/640
Difference-in-Difference	-45.20	16.50	3.95
P-value	0.55	0.76	0.88
90% Confidence Interval	(-168.7, 78.3)	(-71.9, 104.9)	(-39, 46.9)
Baseline Participant Mean	204.89	113.15	48.93
Intervention Period Participant Mean	250.00	153.21	101.56
Baseline Comparison Mean	310.81	170.27	94.59
Intervention Period Comparison Mean	401.12	193.83	143.27
Relative Difference	-22.1%	14.6%	8.1%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	38/321	17/321	25/321
Difference-in-Difference	24.34	6.63	6.70
P-value	0.62	0.82	0.82

Measures	PANO		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	26/319	25/319	40/319
Difference-in-Difference	-69.83	9.88	1.19
P-value	0.10	0.79	0.97

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.22: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, PANO Programs

Measures	PANO	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	29/640	42/640
Difference-in-Difference	- 6.55	50.78
P-value	0.88	0.87
90% Confidence Interval	(-76.8, 63.7)	(-464.5, 566.1)
Baseline Participant Mean	67.28	452.60
Intervention Period Participant Mean	100.04	528.61
Baseline Comparison Mean	108.11	864.86
Intervention Period Comparison Mean	147.52	889.98
Relative Difference	-9.7%	11.2%
Interim Estimates: 0-6 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	13/321	17/321
Difference-in-Difference	12.40	7.18
P-value	0.57	0.97
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	16/319	25/319
Difference-in-Difference	- 19.05	43.72
P-value	0.54	0.84

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the

post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.23: Medicare Expenditures per Beneficiary, PANO Programs

Measures (2011 USD)	PANO				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	622/640	42/640	64/640	405/640	618/640
Difference-in-Difference	$21.44	$308.87	- $83.94	- $349.20	- $144.20
P-value	0.98	0.55	0.34	0.16	0.72
90% Confidence Interval	(-1607, 1649.9)	(-529.7, 1147.5)	(-229.3, 61.4)	(-754.1, 55.7)	(-794.2, 505.8)
Baseline Participant Mean	$4,524.38	$912.72	$122.45	$770.63	$2,368.01
Intervention Period Participant Mean	$5,523.25	$1,287.82	$138.21	$927.83	$2,510.15
Baseline Comparison Mean	$5,273.49	$1,399.17	$158.50	$748.97	$2,159.76
Intervention Period Comparison Mean	$6,250.92	$1,465.40	$258.20	$1,255.38	$2,446.10
Relative Difference	0.5%	33.8%	-68.6%	-45.3%	-6.1%
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	311/321	17/321	38/321	202/321	308/321
Difference-in-Difference	$12.48	$65.49	- $23.65	- $24.64	- $31.57
P-value	0.98	0.82	0.75	0.88	0.88
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	311/319	25/319	26/319	203/319	310/319
Difference-in-Difference	$8.95	$243.94	-60.4*	-325.5**	- $112.88
P-value	0.99	0.47	0.09	0.05	0.68

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.24: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, PANO Programs

Measures (2011 USD)	PANO		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	377/392	140/640	20/640
Difference-in-Difference	- $23.20	- $64.98	$51.46
P-value	0.97	0.18	0.66
90% Confidence Interval	(-1,009.4, 963.0)	(-143.9, 13.9)	(-139.5, 242.4)
Baseline Participant Mean	$2,732.47	$135.66	$98.82
Intervention Period Participant Mean	$2,374.70	$112.29	$183.07
Baseline Comparison Mean	$3,163.10	$145.30	$296.52
Intervention Period Comparison Mean	$2,828.53	$186.91	$329.31

Measures (2011 USD)	PANO		
	Total Part D	DME	Home Health
Relative Difference	-0.8%	-47.9%	52.1%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	189/197	64/321	6/321
Difference-in-Difference	- $50.63	- $15.80	- $34.83
P-value	0.87	0.51	0.65
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	188/195	76/319	14/319
Difference-in-Difference	$28.11	- $49.29	$86.67
P-value	0.93	0.16	0.18

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.25: Medication Adherence (Average Proportion of Days Covered), PANO Programs

Measures (Average PDC)	PANO				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	95	69	51	141	166
Difference-in-Difference	- 1.40	- 3.31	2.00	0.84	1.13
P-value	0.62	0.31	0.59	0.70	0.63
90% Confidence Interval	(-6.1, 3.3)	(-8.6, 2.0)	(-4.1, 8.1)	(-2.8, 4.5)	(-2.7, 5.0)
Baseline Participant Mean	91.06	92.31	92.13	89.72	86.25
Intervention Period Participant Mean	94.70	95.68	95.88	95.49	93.43
Baseline Comparison Mean	88.82	87.72	90.25	90.40	88.74
Intervention Period Comparison Mean	93.84	94.45	92.04	95.32	94.76
Relative Difference	-1.5%	-3.6%	2.2%	0.9%	1.3%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	45	34	25	68	77
Difference-in-Difference	- 1.09	-6.61*	0.04	0.83	1.68
P-value	0.73	0.07	0.99	0.73	0.51
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	50	35	26	73	89
Difference-in-Difference	- 1.68	- 0.11	3.89	0.86	0.66
P-value	0.60	0.97	0.35	0.72	0.79

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10%

level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.26: Medication Adherence (Proportion of Days Covered ≥ 80%), PANO Programs

Measures (PDC ≥ 80%)	PANO				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	95	69	51	141	166
Difference-in-Difference	2.37	- 1.37	13.97	2.37	11.35*
P-value	0.76	0.87	0.19	0.69	0.07
90% Confidence Interval	(-10.3, 15.0)	(-14.8, 12)	(-3.3, 31.3)	(-7.5, 12.2)	(1.2, 21.5)
Baseline Participant Mean	83.67	87.18	80.77	81.61	72.16
Intervention Period Participant Mean	92.78	95.59	96.00	93.58	92.08
Baseline Comparison Mean	82.14	80.30	86.84	84.35	84.50
Intervention Period Comparison Mean	88.72	90.16	88.25	93.95	93.12
Relative Difference	2.8%	-1.6%	17.3%	2.9%	15.7%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	45	34	25	68	77
Difference-in-Difference	5.62	- 6.93	6.18	2.39	10.58
P-value	0.50	0.47	0.60	0.72	0.13
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	50	35	26	73	89
Difference-in-Difference	- 0.56	4.03	21.46*	2.36	12.01*
P-value	0.95	0.63	0.06	0.71	0.07

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.27: ER Visits, Inpatient Admissions, and Incidence of Falls and Fractures per 1,000 Beneficiaries, FP Programs

Measures	FP		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Cumulative Estimates			

Measures	FP		
	ER Visits	Inpatient Admissions	Falls/ Fractures
Nonzero/Total Participant Observations in the Post-Intervention Period	224/1,455	106/1,455	219/1,455
Difference-in-Difference	-67.57	- 45.53	-8.53
P-value	0.24	0.26	0.70
90% Confidence Interval	(-161.7, 26.5)	(-111.7, 20.7)	(-45.1, 28.1)
Baseline Participant Mean	367.29	179.62	131.37
Intervention Period Participant Mean	398.63	188.52	150.52
Baseline Comparison Mean	349.75	192.12	155.17
Intervention Period Comparison Mean	448.66	246.57	182.85
Relative Difference	-18.4%	-25.3%	-6.5%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	105/735	50/735	89/735
Difference-in-Difference	-32.87	- 21.62	-12.21
P-value	0.31	0.37	0.63
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	119/720	56/720	130/720
Difference-in-Difference	-34.72	- 23.93	-4.76
P-value	0.37	0.36	0.86

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.28: Unplanned Inpatient Admissions and Length of Stay per 1,000 Beneficiaries, FP Programs

Measures	FP	
	Unplanned Inpatient Admissions	Length of Stay
Cumulative Estimates		
Nonzero/Total Participant Observations in the Post-Intervention Period	88/1,455	105/1,455
Difference-in-Difference	- 36.20	- 286.09
P-value	0.30	0.18
90% Confidence Interval	(-94.2, 21.8)	(-633.5, 61.3)
Baseline Participant Mean	132.71	705.09
Intervention Period Participant Mean	149.94	792.43
Baseline Comparison Mean	149.01	703.20
Intervention Period Comparison Mean	202.58	1,075.62
Relative Difference	-27.3%	-40.6%
Interim Estimates: 0-6 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	43/735	49/735
Difference-in-Difference	- 12.23	- 191.88

Measures	FP	
	Unplanned Inpatient Admissions	Length of Stay
P-value	0.56	0.12
Interim Estimates: 7-12 Months		
Nonzero/Total Participant Observations in the Post-Intervention Period	45/720	56/720
Difference-in-Difference	- 24.09	- 93.20
P-value	0.30	0.53

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.29: Medicare Expenditures per Beneficiary, FP Programs

Measures (2011 USD)	FP				
	Total Parts A and B	Inpatient	Outpatient ER	Outpatient Non-ER	Physician and Ancillary
Cumulative Estimates					
Nonzero/Total Participant Observations in the Post-Intervention Period	1,425/1,455	103/1,455	223/1,455	971/1,455	1,420/1,455
Difference-in-Difference	- $518.96	- $412.86	- $48.46	- $132.76	- $4.50
P-value	0.49	0.24	0.31	0.46	0.98
90% Confidence Interval	(-1,745.7, 707.8)	(-989.6, 163.8)	(-127.4, 30.5)	(-430.2, 164.7)	(-334.5, 325.5)
Baseline Participant Mean	$5,837.92	$1,283.77	$229.38	$1,124.07	$2,433.04
Intervention Period Participant Mean	$6,353.90	$1,354.78	$220.92	$1,175.44	$2,459.33
Baseline Comparison Mean	$6,111.90	$1,318.40	$186.64	$1,174.38	$2,500.61
Intervention Period Comparison Mean	$7,146.84	$1,802.26	$226.63	$1,358.51	$2,531.41
Relative Difference	-8.9%	-32.2%	-21.1%	-11.8%	-0.2%
Interim Estimates: 0-6 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	719/735	50/735	105/735	504/735	717/735
Difference-in-Difference	$37.21	- $150.84	- $30.77	$32.76	- $21.25
P-value	0.94	0.46	0.23	0.78	0.85
Interim Estimates: 7-12 Months					
Nonzero/Total Participant Observations in the Post-Intervention Period	706/720	53/720	118/720	467/720	703/720
Difference-in-Difference	- $562.35	- $263.17	- $17.55	- $167.59	$17.14
P-value	0.21	0.26	0.59	0.13	0.89

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the

post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.30: Part D, Durable Medical Equipment (DME), and Home Health Expenditures per Beneficiary, FP Programs

Measures (2011 USD)	FP		
	Total Part D	DME	Home Health
Cumulative Estimates			
Nonzero/Total Participant Observations in the Post-Intervention Period	982/1,006	372/1,455	56/1,455
Difference-in-Difference	$514.03	$11.78	- $158.03
P-value	0.33	0.77	0.10
90% Confidence Interval	(-348.2, 1376.3)	(-55.7, 79.2)	(-317.8, 1.7)
Baseline Participant Mean	$3,856.18	$193.79	$304.78
Intervention Period Participant Mean	$4,186.86	$211.12	$269.06
Baseline Comparison Mean	$4,434.13	$160.70	$328.03
Intervention Period Comparison Mean	$4,250.78	$166.25	$450.35
Relative Difference	13.3%	6.1%	-51.9%
Interim Estimates: 0-6 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	497/508	184/735	24/735
Difference-in-Difference	$206.63	- $4.22	- $82.89
P-value	0.55	0.85	0.12
Interim Estimates: 7-12 Months			
Nonzero/Total Participant Observations in the Post-Intervention Period	485/498	188/720	32/720
Difference-in-Difference	$306.10	$16.22	- $75.06
P-value	0.29	0.56	0.24

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.31: Medication Adherence (Average Proportion of Days Covered), FP Programs

Measures (Average PDC)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	322	221	130	426	447
Difference-in-Difference	0.26	0.39	1.24	1.52	0.63
P-value	0.88	0.85	0.64	0.27	0.69
90% Confidence Interval	(-2.5, 3.0)	(-2.9, 3.7)	(-3.1, 5.5)	(-0.8, 3.8)	(-2.0, 3.3)
Baseline Participant Mean	91.02	89.76	91.00	91.12	87.49
Intervention Period Participant Mean	94.61	95.62	95.61	95.35	93.73

Measures (Average PDC)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Baseline Comparison Mean	90.93	89.58	91.16	91.73	89.09
Intervention Period Comparison Mean	94.25	95.03	94.54	94.43	94.69
Relative Difference	0.3%	0.4%	1.4%	1.7%	0.7%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	159	113	66	216	234
Difference-in-Difference	0.10	- 0.73	1.60	1.48	0.35
P-value	0.96	0.74	0.57	0.35	0.84
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	163	108	64	210	213
Difference-in-Difference	0.42	1.55	0.86	1.56	0.93
P-value	0.82	0.47	0.76	0.31	0.59

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

Appendix Table D.32: Medication Adherence (Proportion of Days Covered ≥ 80%), FP Programs

Measures (PDC ≥ 80%)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
Cumulative Estimates					
Total Participant Observations in the Post-Intervention Period	322	221	130	426	447
Difference-in-Difference	0.61	2.46	9.30	0.26	0.46
P-value	0.88	0.62	0.11	0.94	0.91
90% Confidence Interval	(-6.2, 7.4)	(-5.8, 10.7)	(-0.3, 18.9)	(-5.4, 5.9)	(-5.9, 6.8)
Baseline Participant Mean	86.34	82.09	83.54	86.99	78.97
Intervention Period Participant Mean	92.53	95.05	95.43	93.45	90.20
Baseline Comparison Mean	85.71	82.99	90.22	86.03	81.53
Intervention Period Comparison Mean	91.29	93.43	92.77	92.23	92.24
Relative Difference	0.7%	3.0%	11.1%	0.3%	0.6%
Interim Estimates: 0-6 Months					
Total Participant Observations in the Post-Intervention Period	159	113	66	216	234
Difference-in-Difference	1.10	- 0.41	7.14	0.22	- 0.82
P-value	0.82	0.94	0.29	0.96	0.85
Interim Estimates: 7-12 Months					
Total Participant Observations in the Post-Intervention Period	163	108	64	210	213
Difference-in-Difference	0.14	5.46	11.52*	0.29	1.87

Measures (PDC ≥ 80%)	FP				
	Beta Blockers	Calcium Channel Blockers	Diabetes Medication	RAS Antagonists	Statins
P-value	0.98	0.32	0.06	0.94	0.66

Notes: *p-value < 0.10; ** p-value < 0.05; *** p-value < 0.01. The p-value is the probability that, if there is no effect of wellness programs, the observed findings could have occurred by chance in the data. The 90% confidence interval represents possible values of the effect of wellness programs that are not statistically different (at the 10% level) from the reported DiD estimate. The unit of observation is beneficiary-half-years. Cumulative Estimates: comparison between the baseline period and the entire twelve-month post-intervention period; estimates and reported means refer to a twelve-month period. Interim Estimates: 0-6 Months: comparison between the baseline period and the first six months of the post-intervention period; estimates and reported means refer to a six-month period. Interim Estimates: 7-12 Months: comparison between the baseline period and the second six months of the post-intervention period; estimates and reported means refer to a six-month period. Relative difference is calculated as the DiD estimate divided by the baseline participant mean and is expressed as a percentage.

APPENDIX E – SURVEY INSTRUMENTS

Appendix E contains two survey instruments. Section E.1 contains the Twelve-Month National Survey and Section E.2 contains the Twelve-Month Participant Survey.

Baseline and Six-Month National Surveys and Baseline and Six-Month Participant Surveys are available in the "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." [77]

[77] "Wellness Prospective Evaluation Report on Six-Month Follow-Up Survey Outcomes and Estimated Operational Costs." Centers for Medicare & Medicaid Services (CMS): Acumen, LLC. November 2017. Available at: https://downloads.cms.gov/files/cmmi/community-basedwellnessrrevention-sixthmnthoutcomes-operationalcostrpt.pdf.

E.1 Twelve-Month National Survey

Start Here

- Please use a black or blue pen to complete this form.
- Mark ☒ to indicate your answer. If you want to change your answer, darken the box ■ and mark the correct answer.

Your Health

These first questions are about your health.

Please mark one answer only. If you are unsure about how to answer a question, please give the best answer you can.

1. In general, would you say your health is

- ☐ Excellent
- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Poor

2. Compared to one year ago, how would you rate your health in general now?

- ☐ Much better than one year ago
- ☐ Somewhat better now than one year ago
- ☐ About the same as one year ago
- ☐ Somewhat worse now than one year ago
- ☐ Much worse now than one year ago

3. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

a. Vigorous activities, such as running, lifting heavy objects, or participating in strenuous sports

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

b. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

c. Lifting or carrying groceries

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

d. Climbing several flights of stairs

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

e. Climbing one flight of stairs

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

f. Bending, kneeling, or stooping

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

g. Walking more than a mile

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

h. Walking several hundred yards

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

17865

1

i. Walking one hundred yards

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

j. Bathing or dressing yourself

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

4. During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

a. Cut down on the amount of time you spent on work or other activities

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. Accomplished less than you would like

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. Were limited in the kind of work or other activities

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

d. Had difficulty performing the work or other activities (for example, it took extra effort)

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

5. During the past 4 weeks, how much of the time have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

a. Cut down on the amount of time you spent on work or other activities

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. Accomplished less than you would like

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. Did work or activities less carefully than usual

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

17865

6. During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

- ☐ Not at all
- ☐ Slightly
- ☐ Moderately
- ☐ Quite a bit
- ☐ Extremely

7. How much bodily pain have you had during the past 4 weeks?

- ☐ None
- ☐ Very mild
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

- ☐ Not at all
- ☐ A little bit
- ☐ Moderately
- ☐ Quite a bit
- ☐ Extremely

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks…

a. Did you feel full of life?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. Have you been very nervous?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. Have you felt so down in the dumps that nothing could cheer you up?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

d. Have you felt calm and peaceful?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

17865

e. Did you have a lot of energy?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

f. Have you felt downhearted and depressed?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

g. Did you feel worn out?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

h. Have you been happy?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

i. Did you feel tired?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

10. During the <u>past 4 weeks</u>, how much of the time has your <u>physical health or emotional problems</u> interfered with your social activities (like visiting friends, relatives, etc.)?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

11. How TRUE or FALSE is <u>each</u> of the following statements for you?

a. I seem to get sick a little easier than other people

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

b. I am as healthy as anybody I know

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

c. I expect my health to get worse

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

17865

d. My health is excellent

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

Physical Activity

Physical activities are activities where you move and increase your heart rate above its resting rate, whether you do them for pleasure, work, or transportation. The following questions ask about the amount and intensity of physical activity you usually do. The intensity of the activity is related to the amount of energy you use to do these activities.

Examples of physical intensity levels:

Intensity Level	Examples
Light activities: Your heart beats slightly faster than normal. You can talk and sing.	Walking leisurely, stretching, or light yard work
Moderate activities: Your heart beats faster than normal. You can talk but not sing.	Fast walking, aerobics class, strength training, swimming gently
Vigorous activities: Your heart rate increases a lot. You can't talk or your talking is broken up by large breaths.	Stair machine, jogging or running, tennis, racquetball, or badminton

12. How physically active are you?
Please mark <u>one</u> answer for each question.

a. I rarely or never do any physical activities.

- ☐ Yes
- ☐ No

b. I do some <u>light</u> or <u>moderate</u> physical activities, but not every week.

- ☐ Yes
- ☐ No

c. I do some <u>light</u> physical activity every week.

- ☐ Yes
- ☐ No

d. I do <u>moderate</u> physical activities every week, but less than 30 minutes a day or 5 days a week.

- ☐ Yes
- ☐ No

e. I do <u>vigorous</u> physical activities every week, but less than 20 minutes a day or 3 days a week.

- ☐ Yes
- ☐ No

f. I do 30 minutes or more a day of <u>moderate</u> physical activities, 5 or more days a week.

- ☐ Yes
- ☐ No

17865

g. I do 20 minutes or more a day of vigorous physical activities, 3 or more days a week.

☐ Yes
☐ No

h. I do activities to increase muscle strength, such as lifting weights or calisthenics, once a week or more.

☐ Yes
☐ No

i. I do activities to improve flexibility, such as stretching or yoga, once a week or more.

☐ Yes
☐ No

Falls

13. A fall is when your body goes to the ground without being pushed. Did you fall in the past 6 months?

☐ Yes ➔ ☐☐ times
☐ No ➔ **SKIP TO 15**

14. How many of these falls caused you to limit your regular activities for at least a day or to see a doctor?

☐☐ Falls limiting activity or requiring medical attention

15. In the past 6 months, have you had a problem with balance or walking?

☐ Yes
☐ No
☐ Limited to a bed or wheelchair
↳ **SKIP TO 18**

16. Are you afraid of falling?

☐ Yes
☐ No

Your Confidence in Balance

The next questions are about keeping your balance in different situations. You may have to imagine yourself in these situations if you have not encountered them recently. For each one, choose any number between 0 (no confidence) and 100 (complete confidence) to say how confident you are that you could keep your balance. If you normally use a cane or walker or hold on to someone, answer as if you had that help.

0	10	20	30	40	50	60	70	80	90	100
No Confidence										**Complete Confidence**

17. How confident are you that you can maintain your balance and remain steady when you…

a. Stand on your tiptoes and reach for something above your head?

☐☐☐

b. Stand on a chair and reach for something?

☐☐☐

c. Are bumped into by people as you walk through the mall?

☐☐☐

17865

d. Step onto or off of an escalator while holding onto a railing?

☐☐☐

e. Step onto or off of an escalator while holding a package so you cannot hold onto the railing?

☐☐☐

f. Walk outside on icy sidewalks?

☐☐☐

Medicines

The next few questions are about medicines.

18. Do you <u>ever forget</u> to take your medicine?

- ☐ I don't take any medicines ➔ **SKIP TO 22**
- ☐ Yes
- ☐ No

19. Do you ever have problems remembering to take your medicine?

- ☐ Yes
- ☐ No

20. When you feel better, do you sometimes stop taking your medicine?

- ☐ Yes
- ☐ No

21. Sometimes if you feel worse when you take your medicine, do you stop taking it?

- ☐ Yes
- ☐ No

22. Have you participated in any program in the past 24 months, either in your community or online, to address any of the following goals?
Please mark all that apply.

- ☐ Eating healthful foods, such as fruits, vegetables, and whole grains
- ☐ Managing your weight
- ☐ Getting regular exercise appropriate for your ability
- ☐ Improving your balance and preventing falls
- ☐ Managing health problems like arthritis, diabetes, high blood pressure, or other conditions
- ☐ None of the above
- ☐ Other, Specify:

23. Date of filling out this survey:

☐☐ / ☐☐ / 2 0 ☐☐

month day year

Thank you for your time.
Please return the survey using the prepaid addressed envelope enclosed.

17865

E.2 Twelve-Month Participant Survey

Start Here

- Please use a black or blue pen to complete this form.
- Mark ☒ to indicate your answer. If you want to change your answer, darken the box ■ and mark the correct answer.

Your Health

These first questions are about your health.

Please mark one answer only. If you are unsure about how to answer a question, please give the best answer you can.

1. In general, would you say your health is

- ☐ Excellent
- ☐ Very good
- ☐ Good
- ☐ Fair
- ☐ Poor

2. Compared to one year ago, how would you rate your health in general now?

- ☐ Much better than one year ago
- ☐ Somewhat better now than one year ago
- ☐ About the same as one year ago
- ☐ Somewhat worse now than one year ago
- ☐ Much worse now than one year ago

3. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

a. Vigorous activities, such as running, lifting heavy objects, or participating in strenuous sports

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

b. Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

c. Lifting or carrying groceries

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

d. Climbing several flights of stairs

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

e. Climbing one flight of stairs

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

f. Bending, kneeling, or stooping

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

g. Walking more than a mile

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

h. Walking several hundred yards

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

17963

1

i. **Walking <u>one hundred yards</u>**

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

j. **Bathing or dressing yourself**

- ☐ Yes, limited a lot
- ☐ Yes, limited a little
- ☐ No, not limited at all

4. **During the <u>past 4 weeks</u>, how much of the time have you had any of the following problems with your work or other regular daily activities <u>as a result of your physical health</u>?**

a. **Cut down on the <u>amount of time</u> you spent on work or other activities**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. **<u>Accomplished less</u> than you would like**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. **Were limited in the <u>kind</u> of work or other activities**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

d. **Had <u>difficulty</u> performing the work or other activities (for example, it took extra effort)**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

5. **During the <u>past 4 weeks</u>, how much of the time have you had any of the following problems with your work or other regular daily activities <u>as a result of any emotional problems</u> (such as feeling depressed or anxious)?**

a. **Cut down on the <u>amount of time</u> you spent on work or other activities**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. **<u>Accomplished less</u> than you would like**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. **Did work or activities <u>less carefully than usual</u>**

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

17963

6. During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbors, or groups?

- ☐ Not at all
- ☐ Slightly
- ☐ Moderately
- ☐ Quite a bit
- ☐ Extremely

7. How much bodily pain have you had during the past 4 weeks?

- ☐ None
- ☐ Very mild
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very severe

8. During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

- ☐ Not at all
- ☐ A little bit
- ☐ Moderately
- ☐ Quite a bit
- ☐ Extremely

9. These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.

How much of the time during the past 4 weeks…

a. Did you feel full of life?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

b. Have you been very nervous?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

c. Have you felt so down in the dumps that nothing could cheer you up?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

d. Have you felt calm and peaceful?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

e. Did you have a lot of energy?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

f. Have you felt downhearted and depressed?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

g. Did you feel worn out?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

h. Have you been happy?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

i. Did you feel tired?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

10. During the <u>past 4 weeks</u>, how much of the time has your <u>physical health or emotional problems</u> interfered with your social activities (like visiting friends, relatives, etc.)?

- ☐ All of the time
- ☐ Most of the time
- ☐ Some of the time
- ☐ A little of the time
- ☐ None of the time

11. How TRUE or FALSE is <u>each</u> of the following statements for you?

a. I seem to get sick a little easier than other people

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

b. I am as healthy as anybody I know

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

c. I expect my health to get worse

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

17963

d. My health is excellent

- ☐ Definitely true
- ☐ Mostly true
- ☐ Don't know
- ☐ Mostly false
- ☐ Definitely false

Physical Activity

Physical activities are activities where you move and increase your heart rate above its resting rate, whether you do them for pleasure, work, or transportation. The following questions ask about the amount and intensity of physical activity you usually do. The intensity of the activity is related to the amount of energy you use to do these activities.

Examples of physical intensity levels:

Intensity Level	Examples
Light activities: Your heart beats slightly faster than normal. You can talk and sing.	Walking leisurely, stretching, or light yard work
Moderate activities: Your heart beats faster than normal. You can talk but not sing.	Fast walking, aerobics class, strength training, swimming gently
Vigorous activities: Your heart rate increases a lot. You can't talk or your talking is broken up by large breaths.	Stair machine, jogging or running, tennis, racquetball, or badminton

12. How physically active are you?
Please mark <u>one</u> answer for each question.

a. I rarely or never do any physical activities.

- ☐ Yes
- ☐ No

b. I do some <u>light</u> or <u>moderate</u> physical activities, but not every week.

- ☐ Yes
- ☐ No

c. I do some <u>light</u> physical activity every week.

- ☐ Yes
- ☐ No

d. I do <u>moderate</u> physical activities every week, but less than 30 minutes a day or 5 days a week.

- ☐ Yes
- ☐ No

e. I do <u>vigorous</u> physical activities every week, but less than 20 minutes a day or 3 days a week.

- ☐ Yes
- ☐ No

f. I do 30 minutes or more a day of <u>moderate</u> physical activities, 5 or more days a week.

- ☐ Yes
- ☐ No

17963

g. I do 20 minutes or more a day of vigorous physical activities, 3 or more days a week.

☐ Yes
☐ No

h. I do activities to increase muscle strength, such as lifting weights or calisthenics, once a week or more.

☐ Yes
☐ No

i. I do activities to improve flexibility, such as stretching or yoga, once a week or more.

☐ Yes
☐ No

Falls

13. A fall is when your body goes to the ground without being pushed. Did you fall in the past 6 months?

☐ Yes ➔ ☐☐ times
☐ No ➔ **SKIP TO 15**

14. How many of these falls caused you to limit your regular activities for at least a day or to see a doctor?

☐☐ Falls limiting activity or requiring medical attention

15. In the past 6 months, have you had a problem with balance or walking?

☐ Yes
☐ No
☐ Limited to a bed or wheelchair
↳ **SKIP TO 18**

16. Are you afraid of falling?

☐ Yes
☐ No

Your Confidence in Balance

The next questions are about keeping your balance in different situations. You may have to imagine yourself in these situations if you have not encountered them recently. For each one, choose any number between 0 (no confidence) and 100 (complete confidence) to say how confident you are that you could keep your balance. If you normally use a cane or walker or hold on to someone, answer as if you had that help.

0	10	20	30	40	50	60	70	80	90	100
No Confidence										**Complete Confidence**

17. How confident are you that you can maintain your balance and remain steady when you…

a. Stand on your tiptoes and reach for something above your head?

☐☐☐

b. Stand on a chair and reach for something?

☐☐☐

c. Are bumped into by people as you walk through the mall?

☐☐☐

d. Step onto or off of an escalator while holding onto a railing?

☐☐☐

e. Step onto or off of an escalator while holding a package so you cannot hold onto the railing?

☐☐☐

f. Walk outside on icy sidewalks?

☐☐☐

Medicines

The next few questions are about medicines.

18. Do you <u>ever forget</u> to take your medicine?

- ☐ I don't take any medicines ➔ **SKIP TO 22**
- ☐ Yes
- ☐ No

19. Do you ever have problems remembering to take your medicine?

- ☐ Yes
- ☐ No

20. When you feel better, do you sometimes stop taking your medicine?

- ☐ Yes
- ☐ No

21. Sometimes if you feel worse when you take your medicine, do you stop taking it?

- ☐ Yes
- ☐ No

Program Participation

Wellness programs *are ongoing, organized group meetings or sessions, done online or in person, where the focus is on improving one's health through knowledge and/or activity. (Do not include diet or fitness programs done on an individual basis.)*

22. Our records show that you started a wellness program in [FILL Month, yyyy]. How many of the program sessions or meetings did you participate in?

- ☐ All sessions or meetings
- ☐ Most of the sessions or meetings
- ☐ Half of the sessions or meetings
- ☐ Fewer than half of the sessions or meetings

23. Are you still participating in this program?

- ☐ Yes ➔ **SKIP TO 33**
- ☐ No ➔ **GO TO 24a**

24a. Besides the above program, have you participated in any other wellness programs, either in your community or online, to improve your health in the <u>past six months</u>?

Please mark all that apply.

- ☐ Yes, in my community ➔ **GO TO 24b**
- ☐ Yes, online ➔ **GO TO 24b**
- ☐ No ➔ **SKIP TO 25**

24b. What other kind of wellness programs did you participate in in the <u>past six months</u>?
Please mark all that apply.

- ☐ Eating healthful foods, such as fruits, vegetables, and whole grains
- ☐ Managing your weight
- ☐ Getting regular exercise appropriate for your ability
- ☐ Improving your balance and preventing falls
- ☐ Managing health problems like arthritis, diabetes, high blood pressure, or other conditions
- ☐ None of the above
- ☐ Other, Specify:

25. How much would you be willing to pay in total for the program that you enrolled in?
Please write a whole dollar amount.

$ ☐☐☐

26. What would you say was the best thing about the program?
Please specify in the space below.

27. Did you stop participating in the program when it was over or before it was over?

- ☐ I stopped participating in the program when it was over ➔ **SKIP TO 33**
- ☐ I stopped participating in the program before it was over ➔ **GO TO 28**

28. Did you decide to leave the program because of your ill health?

- ☐ Yes
- ☐ No

29a. Did you decide to leave the program because it did not meet your health needs?

- ☐ Yes ➔ **GO TO 29b**
- ☐ No ➔ **SKIP TO 30**

29b. In what ways did the program fail to meet your health needs?
Please specify in the space below.

30. Below is a list of possible reasons why someone might leave the program. For <u>each</u>, please select how important it was in YOUR decision to leave the program.

Possible reasons why someone might leave the program	Very important in my decision	Somewhat important in my decision	Not at all important in my decision
a. The instructor was not helpful	☐	☐	☐
b. I did not learn anything new	☐	☐	☐
c. I did not achieve the results I expected	☐	☐	☐
d. Parking was a problem	☐	☐	☐
e. The program location was too far	☐	☐	☐
f. Transportation was a problem	☐	☐	☐
g. The program hours were not convenient to me	☐	☐	☐
h. The program was not offered in my main spoken language	☐	☐	☐
i. Not enough people in the program were the same gender as myself	☐	☐	☐
j. Not enough people in the program were in my age group	☐	☐	☐
k. The instructor was not in my age group	☐	☐	☐
l. The instructor was not the same gender as myself	☐	☐	☐
m. The program cost was too high	☐	☐	☐

31. Please use the space below to describe any other reasons you had for deciding to leave the program.

32. What would it take for you to return to the program? Please list anything that comes to mind when thinking about what it would take for you to return to the program.

33. Date of completing this survey:

☐☐ / ☐☐ / 2 0 ☐☐

month day year

Thank you for your time. Please mail the survey using the prepaid addressed envelope enclosed.

17963