Impacts of Workplace Health Promotion and Wellness Programs on Health Care Utilization and Costs

Results From an Academic Workplace

John M. Dement, PhD, Carol Epling, MD, MSPH, Julie Joyner, BA, and Kyle Cavanaugh, MEd, MBA

Objective: This study evaluated the impacts of a long-standing workplace health promotion (HP) program on health care utilization and costs and estimated return on investment (ROI). **Methods:** Analyses used a retrospective, observational cohort design based on 7 years (2005 to 2011) of health claims and HP program participation data for 3829 HP participants and 6617 controls. Inverse propensity score-weighted mixed-model regression methods were used to balance employee demographics and comorbidities by study arm. **Results:** Mean monthly health care costs were $35 less for HP participants compared with controls, and results were robust based on sensitivity analyses. ROI was estimated to be $2.53 for every dollar spent on the HP program. **Conclusions:** Results support a positive impact of HP program participation with regard to reduced health care utilization and costs and a positive ROI.

The Centers for Disease Control and Prevention (CDC) defines a workplace health program as ''a coordinated and comprehensive set of health promotion and protection strategies implemented at the worksite that includes programs, policies, benefits, environmental supports, and links to the surrounding community designed to encourage the health and safety of all employees.''[1,2] The workplace can be an important setting for wellness programs because of the concentration of individuals by time and place and the provision of health benefits through employer-based health plans. In addition, worksites generally function as systems for communication, education, and social support, offering repeated contacts with program participants, and a social environment that can influence norms and expectations. Worksite health programs, however, are extremely variable with regard to scope and employee participation, and prior studies have demonstrated variable but generally positive impacts of these programs as measured by changes in biometric measures, health care costs, and return on investment (ROI).[3–8]

Dating back to the mid-1980s, analyses have been conducted on the impact of worksite health promotion (HP) programs. The analyses have tended to focus on health behavior such as smoking or exercise, health outcomes such as body mass index (BMI), and financial impacts such as reduced health care costs or absenteeism. Baxter et al[4] have recently conducted a meta-analysis across worksite HP programs, and assessed the relationship between ROI and the quality of the study methodology. The authors appropriately point out the complexities involved in issues ranging from the lack of control groups, defined intervention, and common financial analysis. Currently, there is no agreement on best analytic methods and outcomes. Researchers do not agree on a standard approach for calculation of ROI. Studies performed with higher methodologic quality tended to show smaller positive ROI. In addition, studies including indirect costs tended to show lower positive ROI. In their review of published US studies from 2000 to 2010, Lerner et al[3] noted that only 2 of 10 studies that reported a positive ROI for HP programs incorporated relatively high-quality methods. Thus, Lerner called for more rigorous assessment of HP outcomes.

Several studies have evaluated shorter-term outcomes related to newly implemented HP programs. Grossmeier et al[9] found a positive ROI for a newly implemented comprehensive HP program including financial incentives tied to health care benefit premiums. Nyman et al[10–12] found no savings associated with HP programs during the first 2 years. Naydeck et al[13] and Liu et al[14] identified savings associated with 3 years of health claims following implementation of HP programs. In 2015, Musich et al[15] examined outcomes related to HP programs in place for 7 years demonstrating significant improvements in health and reduced health care costs. Parkinson et al[16] demonstrated risk reduction and reduced health care costs among heath care workers participating in a comprehensive wellness, prevention, and chronic disease management program established in 2005.

The challenges and limitations of this field of inquiry notwithstanding, the efforts of this study are aimed at contributing to the discussion of the impact of worksite HP programs. This study evaluated the effectiveness of a workplace HP in place for greater than 20 years using a retrospective, observational cohort design. Health and pharmacy claims were used in longitudinal analyses (2005 to 2011) to evaluate the impact of HP programs with regard to health care utilization and costs. Robust statistical methods were used to account for potential bias related to employees who choose not to participate in HP.

METHODS

Description of the Workforce and Workplace HP Program

The worksite engaged during this study is a large university that includes an academic medical center as well as two regional hospitals in surrounding counties. The workforce engages in research, instruction, administration, and direct patient care. The overall employment base during the study period was approximately 36,000 consisting of 35% men and 65% women. The average age of the workforce was 44.9 years with an average service time of 9.9 years. The population distribution by race/ethnicity was: white 69%; black or African American 22%; Asian 8%; and approximately 1% American Indian or Hawaiian/Pacific Islander. Approximately 3% of the workforce self-identifies as Hispanic.

The university is self-insured with four health plans that employees can choose, and approximately 32,000 of the total workforce of 36,000 are benefits eligible. In the fall of 1988, an HP program was initiated and was available to all benefits-eligible employees. The HP program is a part of Employee Occupational

From the Division of Occupational and Environmental Medicine (Dr Dement, Dr Epling), Duke University Medical Center; Duke Occupational Health and Wellness (Ms Joyner); and Duke Human Resources (Mr Cavanaugh), Durham, North Carolina.

This work was partially funded by the National Institute for Occupational Safety and Health (R01-OH099468).

The authors report no conflicts of interest.

Address correspondence to: John M. Dement, PhD, Professor, Division of Occupational and Environmental Medicine, Department of Community and Family Medicine, Duke University Medical Center, 2200 W. Main Street, Suite 400, Durham, NC 27705 (John.Dement@Duke.edu).

DOI: 10.1097/JOM.0000000000000555

Health and Wellness (EOHW), which encompasses Employee Occupational Health, the HP program, and the Employee Assistance Program. This organizational structure allows for cross-referrals between the units and more coordinated employee services. For example, the Health Risk Assessment (HRA) is part of pre-employment, and new employees are connected to the appropriate wellness resources at the start of their employment. The HP program is an employee benefit, so there is a strong integration with the Benefits Department and Human Resources. HP is integrated into the Benefit Plan Design and HR Policy. For example, 3 months of HP Health Coaching are a prerequisite for bariatric surgery to be covered by the health benefit plans. In addition, participants can earn a portion of their bariatric copay back for continued participation in HP Health Coaching after surgery. During insurance plan open enrollment, self-identified smokers are accessed a $30 premium surcharge that requires successful completion of the HP Tobacco Cessation Program, for the surcharge to be removed. Referrals to HP programs and services are built into the separate disease management program. Participants in disease management can earn pharmacy and copay incentives for completing care management and participating in HP services/programs, such as fitness/nutrition consultations or health coaching. The HP program can also refer a participant for disease management services, who may not have been identified through the claims process yet, but could benefit from services. During the course of clinical care for job-related injuries or fitness assessments, EOHW may refer employees to HP programs such as smoking cessation, fitness, and nutrition consultations. The HP program is integrated in these areas as part of the organization's overall population health strategy.

The HP program is voluntary, free of charge (with the exception of gym membership and chair massage), and delivered by a dedicated on-site staff. Programs and services are offered in-person in multiple and varied work locations, with web and telephonic options available, in an effort to remove participation barriers. Participants can earn ''wellness dollars'' and/or token incentives for participation. Wellness dollars can be used to purchase health and fitness items. HP incentives are not linked to health benefit plan premiums. The HP program is branded to fit the organizational culture and to be easily identified by participants across all work locations. HP program marketing and communication are multi-faceted, and the HP program works closely with internal communications in the development of marketing materials and web content. Many communications are targeted to the broad population through print materials, web, bulletin boards, listserves, health plan publications, and organizational meetings/forums. Other communications are targeted to specific work groups that are receiving HP services customized for their department or unit. Individuals may also elect to receive individualized communications based on their specific health risk or health interests.

The HP program was launched with the administration of an on-site HRA and biometric screening for cholesterol, height, weight, BMI, and blood pressure to get a baseline of the lifestyle related risks of the workforce. Initially, the main activities were general health education classes and health awareness events targeted at obtaining broad program visibility and participation. After establishing a core participant base over the years, the program direction shifted toward more targeted high-risk interventions while continuing to maintain broad-based participation and program visibility. Programs and services also shifted to a delivery model that included providing programs and services at the employees' work locations, and the program added telephonic and online modes of interaction in an effort to reduce barriers to participation and increase program accessibility. The program framework incorporates interventions that target key health indicators, including smoking, blood pressure, cholesterol, weight/nutrition, physical activity, and stress. HP program elements include:

- HRA: The HRA takes place during work hours, requires approximately 30 minutes for completion, and is facilitated by a registered nurse. Employees can complete an HRA each year via three mechanisms: at new employee orientation, voluntarily through a mobile system at worksites, or via appointment in the employee health clinic. Participants may also complete a web-based HRA with self-reported biometrics.
- Health Coaching: During the HRA anyone with a measured BMI 30 or more, a blood pressure of at least 140/90, a cholesterol at least 240, or who self-reports they are a tobacco user are offered health coaching. Programs are approximately 12 months long, but differ in intensity and focus, depending on how much support the participant desires. In the lighter touch coaching program, participants receive one face-to-face meeting with a counselor in the first month, after monthly mailings of health education materials and electronic recommendations for various programs and wellness resources. They have two telephone counseling sessions (at months 6 and 12) which are coupled with biometric feedback. In the higher touch coaching program, participants receive monthly coaching sessions (face-to-face in months 1, 3, 6, 9, and 12, and the rest via telephone), they receive quarterly biometric feedback, targeted health education materials, and information and links to upcoming programs and wellness resources. Both programs share similar theoretical foundations.[17]
- In addition to the more intensive Health Coaching interventions, many other programs and services are offered through a comprehensive HP program. These include:
    - Weight Management Group Classes: Five 10-week classes annually in varying work locations.
    - Fitness and Nutrition Consultations: In-person or telephonic at three employee health clinics and at the worksite as requested.
    - Running and Walking Clubs: Two 12-week clubs at five locations, annually.
    - Health Campaigns to increase physical activity: Quarterly, across all locations.
    - On-site Farmers Markets: Weekly, April to September at three locations.
    - Mobile Farmers Markets: Year-round with home and worksite delivery.
    - Web-based resources: Access to informational materials as well as individual health status, exercise, nutrition, and biometric trackers.
    - Lending Library: Available at two employee health clinic locations.
    - Participation incentive program whereby employee earns health and wellness related items for personal use or donation to pediatric patients. No incentive is linked to health benefit premiums.
    - Chair Massage: Weekly at two locations and at worksites as requested.
    - Fitness Club: Discounted, payroll deducted memberships to over 30 fitness facilities across seven counties.

Data Sources and Cohort Definition

A matched cohort design was considered for this study, but proved impractical as the employer HP program has been in place for many years and has engaged a large fraction of employees, thus limiting the pool of potential controls for individual matching. Inability to match HP participants to a control would result in discarding a substantial number of subjects, leading to a loss in statistical power.[18] Instead, inverse propensity score-weighted multivariate regression methods were used to balance employee demographics and comorbidities by study arm in the primary analyses. A comprehensive longitudinal data system provided data for this study, and has been previously described and used for other

longitudinal analyses involving health claims. [19,20] Expenditures for employee copayments and deductibles were not included because these expenditures were not directly relevant to the employer ROI.

The HP study population included employees less than 65 years old, completing at least one HRA between January 1, 2006 and December 31, 2010, and having no evidence of participation in any HP programs in the 2 years before the study start date of January 1, 2006. The HP study population was further restricted to those who also participated in one of the employer-provided health insurance plans for a minimum of 12 months before their first study-eligible HRA and for a minimum of 12 months following their entry HRA. The start date of January 2006 was chosen, as this is the starting date for the HP health risk appraisal in the current format. The requirement for at least 12 months of health plan participation before and after the entry HRA assured sufficient data to investigate changes in health care costs and utilization related to HP participation. Employees participating in the employer-based disease management program were eliminated from these analyses to better define the impact of HP programs separate from disease management.

Controls were defined as employees less than 65 years of age who did not complete an HRA during the study period of 2006 to 2011 and did not participate in the disease management program in the 2 years before study start date or during the study period. Entry dates for controls were assigned by random selection of year and month between available dates while assuring a minimum of 12 months of health insurance plan participation before and after cohort entry. In addition, the distribution of cohort entry year for the pool of controls was matched as closely as possible to that of the study population using the range of available start dates.

HP participants and controls diagnosed with certain conditions based on health claims were excluded. These exclusions included employees undergoing transplants (heart, liver, lung, renal, and pancreas), end stage renal disease, those diagnosed with HIV/AIDS or hemophilia, or employees undergoing bariatric surgery. Members otherwise eligible were eliminated from the HP or control study arms based on diagnoses or procedures identified in baseline or follow-up health claims. DRG codes, Elixhauser disease groups, and Healthcare Cost and Utilization Project (HCUP) disease groupings were used to identify conditions for exclusions.[21,22]

Claims-Level Exclusions, Cost Data Adjustments, and High-Cost Individuals

To better compare HP participants and controls, claims related to certain health conditions unlikely to benefit from HP participation during the study time frame were excluded. These included claims related to fertility treatments, trauma, and burns. In addition, the potential impact of excluding claims related to cancers and pregnancy was assessed in sensitivity analyses. All claims-level exclusions were based on HCUP Clinical Classification System diagnosis groupings.[22]

The health insurance and pharmacy claims cost data span 7 years (2005 to 2011). To compare costs incurred during different years and express them in current dollars, we adjusted for inflation. The national Consumer Price Index (CPI) for medical care was used to adjust medical and pharmacy costs as well as HP program costs to calendar year 2011.

Individuals with $100,000 or more in CPI-adjusted health and pharmacy claims costs in a given year were identified. We conducted sensitivity analyses to evaluate the impact of high-cost individuals on primary study results.

Prevalent Disease Status and Comorbidities

Medical and pharmacy claims during the 12 months of insurance coverage before the index date (HRA date or cohort entry date for controls) were used to calculate several measures of disease prevalence and comorbidity. Individual prevalent health conditions at baseline (coronary artery disease, congestive heart failure, diabetes, hypertension, asthma, chronic obstructive pulmonary disease [COPD], depression, obesity, and peripheral vascular disease) were identified using the Agency for Healthcare Research and Quality, HCUP Clinical Classification System,[22] Elixhauser comorbid conditions,[21] or the National Committee for Quality Assurance Healthcare Effectiveness Data and Information Set classification system.[23,24] Dichotomous variables were generated to indicate presence or absence of these claims-based comorbidities.

Two baseline summary measures of comorbidity were calculated based on published methods. The first comorbidity summary was calculated using the combined Chronic Illness and Disability Payment System (CDPS) and the Medicaid Rx system (MRX).[25,26] CDPS is a diagnosis-based risk adjustment model that uses ICD-9 codes to assess risk, and MRX is a pharmacy-based model that uses National Drug Codes (NDC) to assess risk. CDPS-Rx is a combined diagnosis and pharmacy-based model that uses both ICD-9 and NDC codes, and has been used in other costs analyses.[9,27] CDPS-Rx Version 5.3 was used for this study, and utilized both primary and secondary ICD-9 diagnoses codes. A summary CDPS-Rx risk score for the baseline period was calculated for each HP participant or control using weights specified for adults in a commercial heath care benefit program.

In addition to the CDPS-Rx risk score, Elixhauser comorbidity measures were calculated using the baseline claims data.[21] The Elixhauser comorbidity system uses claims data to assign binomial indicator variables representing the presence or absence of 31 different diagnoses. A Elixhauser comorbidity index was developed by summing binomial variables for the baseline year in the same manner as used by Dominick et al.[28]

Covariate Balance Using Propensity Scores

A randomized control trial (RCT) was impractical to address the specific aims of this study; however, robust methods are available for nonrandomized studies.[29] Inverse propensity score-weighted regression was used to balance risk factors and covariates between study arms to mimic what happens in an RCT.[30]

A propensity score is the individual probability of exposure to a treatment (HP participation) given an employee's demographics, comorbidities, and other covariates. Propensity scores for HP participation were calculated using a logistic regression model based on individual covariates in the baseline year. Our objective was to construct the best predictive propensity score model as opposed to a parsimonious model as suggested by Shah et a.[31] Covariates included in the final propensity score logistic model are shown in Table 1.

The output of the logistic model was an individual propensity score, and this was used to calculate inverse propensity score weights used in the primary analyses. Propensity scores were classified into quintiles for some analyses. The use of five groups is common, and has been shown to remove 90% of bias from a confounding variable.[32] Covariate balance achieved using the propensity score quintiles was assessed by evaluating the HP and control group propensity score distribution by propensity score quintile and by calculating standardized differences for each covariate. Standardized differences were calculated as the difference in means for continuous variables or proportions (prevalence) for dichotomous variables between the HP and control groups divided by a measure of the standard deviation of the variable. For categorical variables with multiple levels, the standardized difference was calculated for each level.[33] The standardized difference thus represents the number of standard deviations by which the two groups differ and is not influenced by sample size. The SAS macros developed by Suarez and Faries, based on the methods developed by Austin, were used to calculate standardized differences.[32,34]

TABLE 1. Characteristics of HP Participants and Controls at Baseline

Covariate	HP Participants ($N = 3,829$)	Controls ($N = 6,603$)	P Value Comparing Study Arms*
Mean age (STD)	41.8 (10.3)	41.4 (11.5)	0.0666
% Women	66.8	52.9	<0.0001
Race/ethnicity (%)			<0.0001
White	62.5	70.4	
Black	26.7	17.6	
Other	9.8	12.0	
Comorbidity prevalence (%)			
Diabetes	5.9	4.8	0.0148
Hypertension	10.5	8.6	0.0008
Coronary artery disease	1.2	1.2	0.7767
Congestive heart failure	0.3	0.3	0.7766
COPD	0.4	0.7	0.0207
Asthma	1.9	2.2	0.3229
Peripheral vascular disease	0.4	0.5	0.5554
Depression	7.0	6.5	0.2840
Obesity	1.9	1.4	0.1021
Mean CDPS-Rx risk score (STD)	1.79 (0.23)	1.80 (0.32)	0.0925
Mean number of Elixhauser comorbidities (STD)	0.50 (0.83)	0.45 (0.82)	0.0030
Health plan			<0.0001
A	10.5	18.8	
B	5.9	18.9	
C	7.4	6.6	
D	76.2	55.7	
Mean baseline monthly paid costs (STD)	290 (806)	308 (1,688)	<0.0001
Mean months of follow-up after entry (STD)	41.28 (0.26)	35.46 (0.24)	

CDPS, Chronic Illness and Disability Payment System; HP, health promotion; STD, standard deviation.
*P value comparing HP participants and controls using a chi-square statistic for categorical variables, and Analysis of Variance (ANOVA) procedures using SAS Proc GLM for continuous variables.

Although there are no firm criteria for evaluating standardized differences, absolute standardized differences less than 0.10 are generally considered an indication that adequate covariate balance has been achieved.[34,35]

Levels of HP Program Participation

As previously described, employees are provided with a wide range of HP activities to choose from, ranging from informational materials to in-person health coaching and physical activity programs. The data system recorded employee participation in 148 different HP program activities for each month of the study period. For the current study, HP participants were assigned categorical variables to define the level of program participation based on intensity of the intervention in a manner similar to that used by Neydeck et al[13] and Ozminkowski et al. [36] Senior HP program staff classified all recorded HP program activity types into the following summary categories for statistical analyses: (1) Health Risk Appraisal, (2) Health Checks with biometric measurements, (3) Health Improvement–Informational Materials, (4) Health Improvement—In-Person Counseling, (5) Physical Activity—Cardio, and (6) Physical Activity—Non-Cardio. Time-varying summary measures of cumulative HP program engagement were developed based on the cumulative sum of individual program encounters by these activity categories for each month of follow-up. An overall time-varying cumulative measure of HP program engagement was developed as the sum of all activity categories.

We conducted several preliminary analyses of the individual HP participation data, and found that most employees in the HP study arm participated in multiple offered programs and only 10.3% of employees in the HP study arm had only an HRA. In addition, most employees participating in in-person counseling and health coaching also had evidence of participating in other activities. Given the high degree of overlap and correlation within the HP participation data, we dichotomized HP participation into the following categories by intensity:

HP participation group	Included activities
Low to medium intensity	Health Risk Appraisal, Health Improvement—Informational Materials, Physical Activity—Cardio, Physical Activity—Non-Cardio, Health Checks, Health Improvement—In-Person Counseling
High intensity	12-month duration Health Coaching programs

For our longitudinal analyses, level of HP participation was treated as a time-varying covariate. After the entry HRA, employees may begin HP program participation at any time during the follow-up period, and our analytical framework took this into consideration. For the analyses of impacts by HP program-intensity categories described above, we classified employees by highest intensity achieved, and once an employee participated in one of the high-intensity programs, their status was classified as high intensity for the remainder of their follow-up.

Outcomes Analyses

Descriptive analyses were used to compare baseline characteristics of HP participants and controls. Continuous variables were compared by Analysis of Variance procedures using SAS Proc generalized linear model (GLM). The Wilcoxon rank-sum test using SAS Proc NPARIWAY was used for comparing continuous variables that departed significantly from a normal distribution. Categorical variables were compared using the chi-square test of general association.

Health care utilization and costs for HP participants and controls were compared using inverse propensity score-weighted regression in a mixed-model context. Inverse propensity score weighting was used to eliminate most of the statistical differences between HP participants and nonparticipants, and any remaining differences in demographic, health plan choice, and baseline comorbidities were adjusted for in the final mixed models.[15,35,37] The primary study outcome was mean monthly plan paid costs for the sum of inpatient, outpatient, and pharmacy services. Subanalyses investigated mean monthly health plan paid costs separately for inpatient, outpatient, and pharmacy services as well as mean monthly utilization rates (events/1000 months of health insurance coverage) for emergency department (ED) visits, hospital admissions, and hospital stay days.

HP program effects were estimated by comparing changes in health care costs and utilization between the baseline year and follow-up years for HP participants and controls. Mixed effects linear models that included individual-level random effects were used to account for repeated measures by individuals and individual factors affecting HP program engagement in a manner similar to other studies.[14,38] Inverse propensity score-weighted mixed effects models were developed for each study outcome using the following model form:

$$Y_{it} = \beta_0 + \beta_1\,(\text{Baseline}_i) + \beta_2(\text{X}_{it}) + \beta_3\,(\text{Follow-up Period}) + \beta_4(\text{Study Arm}) + \text{B}_6\,(\text{Follow-up Period}) * (\text{Study Arm}) + \xi_{it} + \mu_{it},$$

where i is the index for individuals; t the index for follow-up period (years) after cohort entry (1 to 2 and 3 or more); Y_{it} the monthly cost or utilization for member i in follow-up time period t; X_{it} the vector of individual-level covariates for age, sex, race/ethnicity, plan type, CDPS-Rx risk score, and dichotomous variables for coronary artery disease, congestive heart failure, diabetes, hypertension, asthma, COPD, depression, obesity, and peripheral vascular disease using claims data in the baseline year; Baseline_i the mean monthly cost or utilization for member i in the baseline year; ξ_i the each study member's mixed-model random effects; and μ_{it} is an error term.

An interaction between time period and study arm was included to allow for differential effects of the HP program as a function of time since first HP program engagement and cohort entry. All individual characteristics were fixed at baseline except age which varied over the follow-up period. Analyses that investigated effects of varying levels of HP participation also included the measures of cumulative HP participation as a time-varying effect. HP programs were only allowed to have an effect on the current or following years in the analysis. SAS Proc Mixed was used for model-building using restricted maximum likelihood. Least squares means estimators from the regression models were used to estimate HP program effects while adjusting for model covariates.

ROI

ROI calculation was based on a benefit–cost ratio, and was calculated as the total savings associated with HP program participation divided by total costs associated with the HP program. The savings were derived based on the mixed-model estimates of mean monthly health care cost saving associated with HP program participation multiplied by the total number of HP program participants in each year of follow-up to arrive at total estimated program cost savings. HP program management and administration functions are purchased from an external vendor, and detailed budgets for the HP program are maintained by the university; therefore, HP program costs are well defined. Institutional staff and administrative costs associated with leadership and management of the HP program were included.[7] Program costs were based on total annual costs for the entire HP program for all employees (HP program participants and nonparticipants) adjusted to 2011 dollars. ROI was estimated as the total savings divided by program total costs for 2006 to 2011.

Sensitivity Analyses

In addition to our inverse propensity score mixed effects regression modeling of costs and utilization, we conducted several additional sensitivity analyses using different analytical approaches and study designs to evaluate robustness of findings for the primary study outcome (total plan paid costs). Propensity score adjustment as opposed to inverse propensity score weighting is another method used in statistical models for bias control.[32] Sensitivity analyses were conducted based on propensity score adjustment using the same mixed-model structure and covariates included in the primary study model. Propensity score quintile categories were included in this model for control of potential bias.

An alternative modeling approach used for several recent studies used inverse propensity score weighting or propensity score adjustment in exponential conditional mean regression models.[15,32,37,39] We conducted similar analyses of our data. For these analyses, inverse propensity score weights were used in a general linear model with a log link function, a negative binomial distribution function, and included the same covariates for age, sex, race/ethnicity, health plan, and baseline comorbidities incorporated in the primary study model. An identical model was developed based on propensity score adjustment rather than weighting, and the final choice of conditional mean regression model was based on a comparison of model fit statistics. A repeated measures generalized estimating equations (GEE) model was used to control for correlated repeated monthly measures of health care utilization and associated costs. Impact of HP participation was calculated using difference-in-difference estimates from the final regression model.

Our sensitivity analyses also used several non-model-based methods. One non-model-based analytical approach commonly used is direct calculation of mean differences in treatment effects within propensity score strata, with an overall effect calculated as a weighted average across all strata.[40,41] Strata effects were weighted by the inverse variance of the stratum-specific mean effect to arrive at an overall HP program effect. This approach is based on the assumption that HP program exposure is essentially at random for individuals with the same propensity quintile, with the expectation that mean comparisons within propensity score groups are unbiased.

Although we considered individual matching of HP participants to controls to be potentially inefficient because of the possibility of omitting nonmatching cases and subsequent loss of statistical power, we performed a 1-to-1 propensity score-matched analysis as our final check of robustness of study findings. We matched HP participants with a single observation in the control group (1:1 matching without replacement) using the "greedy" best-match sampling algorithm developed by Lanehart et al.[40] A matching caliper of 0.15 was used, and if no control observations were available within the range specified by the caliper for an HP participant, then no matched pair is created and the HP participant was dropped. The difference between the outcome means in the HP and control group was tested using a correlated means t test.

RESULTS

Cohort Characteristics and Propensity Score Diagnostics

The final study cohort for the primary analyses included 3829 HP participants and 6603 controls, who contributed 517,431

TABLE 2. Propensity Score Covariates

Covariate	Data Type	Description
Age	Continuous	Age at beginning of baseline year
Sex	Dichotomous	Men, women
Race/ethnicity	Categorical	White, black, other
Cohort entry year	Categorical	Year of cohort entry
Health plan type	Categorical	Employer health plan
CDPS-Rx risk score	Continuous	CDPS-Rx risk score using baseline claims
Elixhauser comorbidity index	Continuous	Sum of Elixhauser disease risks (0–31)
Baseline health claim costs	Continuous	Log of the total paid costs (inpatient, outpatient, and pharmacy) in the baseline year
Prevalent disease: • Coronary artery disease • Congestive heart failure • Diabetes • Hypertension • Asthma • COPD • Depression • Obesity • Peripheral vascular disease	Dichotomous for each disease or health condition	Diabetes and hypertension was based on the HEDIS definition using both health and pharmacy claims.[21] Other conditions were based on HCUP and Elixhauser classifications

CDPS, Chronic Illness and Disability Payment System; HCUP, Healthcare Cost and Utilization Project; HEDIS, Healthcare Effectiveness Data and Information Set.

person-months of health plan coverage during the baseline and follow-up time periods. Table 1 provides a comparison of baseline covariates for the cohort by study arm before propensity score stratification to balance covariates. Compared with controls, HP participants were more represented by women and blacks, had a higher prevalence of diabetes and hypertension, and had a higher mean Elixhauser comorbidity index score. COPD prevalence was slightly higher among controls.

The distribution of propensity scores for HP participants and controls demonstrated a significant degree of overlap (Fig. 1), thus providing common support for propensity score methods. Stratification of propensity scores into quintiles resulted in acceptable overall balance within each stratum as shown in Figure 2. Lastly, results of the standardized differences analyses demonstrated adequate overall covariate adjustment as shown in Table 3. Although good overall covariate balance was achieved with the propensity scores, baseline covariates were included in the statistical models as an additional measure of control for any remaining differences in demographic, health plan choice, and baseline comorbidities.

Overall Results

Primary study results are shown in Table 4. HP participants were found to have significantly lower mean monthly health care costs and utilization following initiation of HP program participation compared with controls. On average, total mean monthly health care costs were approximately $35.1 less for HP participants compared with controls. The reduction in total monthly health care costs was driven by reductions in outpatient costs (−$22.9) and inpatient costs (−$17.9). HP participants experienced a slight increase in average monthly pharmacy costs (+$3.8), although this increase was not statistically significant ($P = 0.4375$). HP participants also experienced significant reductions in ED visit rates (−1.33 visits/1000 insurance months), rates of hospital admissions (−1.10 admissions/1000 insurance months), and days in the hospital (−4.50 days/1000 insurance months) compared with controls.

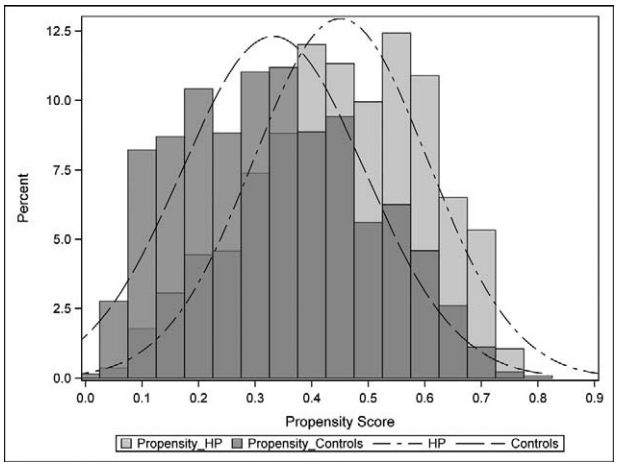

FIGURE 1. Propensity scores by study arm with fitted normal distributions.

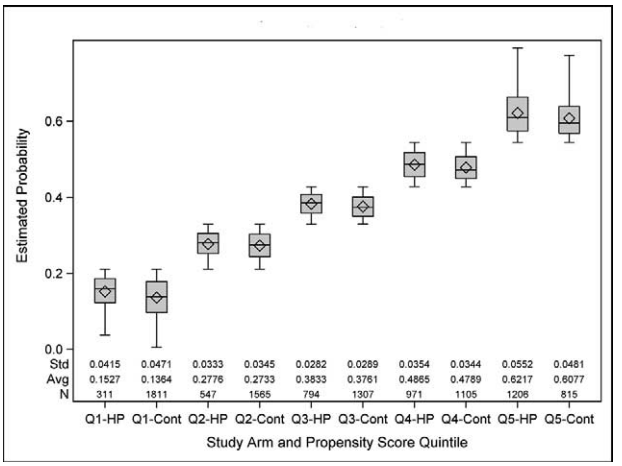

FIGURE 2. Propensity score quintiles for HP participants and controls.

TABLE 3. Standardized Differences Before and After Propensity Score Adjustment

Propensity Score Covariate	Mean	Adjusted Mean Standardized Differences	Standardized Differences by Propensity Score Quintile				
			Quintile #1	Quintile #2	Quintile #3	Quintile #4	Quintile #5
Male sex	−0.2848	−0.0259	−0.0862	0.1784	−0.0627	−0.1690	0.0188
White race	−0.1471	−0.0269	−0.0722	−0.0423	−0.0163	0.0247	−0.0372
Black race	0.2197	0.0345	0.1424	−0.0063	−0.0775	0.0047	0.1185
Other race	−0.0694	−0.0051	−0.0293	0.0594	0.1028	−0.0496	−0.1596
Age	0.0374	−0.0028	−0.0058	0.0662	0.0744	−0.0953	−0.0408
Diabetes	0.0483	−0.0188	−0.0335	−0.0077	−0.0338	−0.0088	−0.0097
Hypertension	0.0670	−0.0022	0.0140	−0.0421	−0.0127	−0.0019	0.0217
Coronary artery disease	−0.0029	−0.0069	−0.0140	−0.0192	−0.0285	0.0246	0.0028
Congestive heart failure	0.0014	−0.0018	−0.0076	0.0051	0.0103	−0.0077	−0.0091
Asthma	−0.0131	−0.0060	−0.0234	−0.0074	0.0185	−0.0097	−0.0082
COPD	−0.0166	−0.0026	−0.0069	0.0126	−0.0015	−0.0119	−0.0056
Depression	0.0217	−0.0062	−0.0284	0.0136	−0.0371	0.0254	−0.0102
Obesity	0.0191	0.0038	0.0332	0.0064	−0.0031	−0.0288	0.0111
Peripheral vascular disease	−0.0038	−0.0032	−0.0253	0.0072	0.0043	−0.0155	0.0134
Health plan A	−0.2376	0.0270	0.2319	−0.1743	0.0598	−0.0331	0.0000
Health plan B	−0.4015	−0.0657	−0.2236	−0.0035	0.0237	0.0047	0.0000
Health plan C	0.0337	0.0150	0.0112	0.0949	0.0182	−0.0183	−0.0458
Health plan D	0.4614	0.0280	0.0885	0.1026	−0.0726	0.0361	0.0458
Log baseline paid costs	1.9058	−0.0981	−0.6547	−1.0966	1.4078	−0.6506	0.5034
Elixhauser comorbidities	0.0507	−0.0142	0.0204	−0.0358	−0.0063	−0.0558	0.0042
CDPS-Rx risk score	−0.0356	−0.0162	0.0019	0.0060	−0.0555	−0.0084	−0.0724

CDPS, Chronic Illness and Disability Payment System; COPD, chronic obstructive pulmonary disease.

In addition to overall estimates of HP program impact, our statistical model allowed assessment of impacts by time period following initial HP program engagement. Figure 3 provides estimates of mean cost differences for the first 2 years after HP engagement and for years 3 to 5. In the first 2 years following HP engagement, the mean monthly total cost savings was $37.7 compared with $32.4 in years 3 to 5. Outpatient and inpatient costs observed a different pattern with outpatient costs decreasing more in the first 2 years, whereas greater inpatient cost reductions were observed after 3 or more years following HP program engagement.

TABLE 4. Mean Health Care Utilization and Costs in the Follow-Up Time Period*

Outcome Measure	Difference† (Controls−HP)	Difference 95% CI	Effect *P* Value
Mean plan paid monthly health care cost ($)			
Total paid	−35.1	−69.6 to −0.5	0.0465
Outpatient paid	−22.9	−42.5 to −3.23	0.0225
Inpatient paid	−17.9	−35.0 to −0.8	0.0400
Pharmacy paid	+3.8	−5.8 to 13.3	0.4375
Mean utilization, events per 1,000 insurance months			
Emergency department visits	−1.33	−2.63 to −0.04	0.0437
Hospital admits	−1.10	−1.71 to −0.50	0.0004
Hospital days	−4.50	−8.13 to −0.87	0.0151

CI, confidence interval; HP, health promotion.

*Inverse propensity score-weighted mixed models, including covariates for age, sex, race/ethnicity, health plan, CDPS-Rx comorbidity risk score, and dichotomous covariates for diabetes, hypertension, CAD, CHF, asthma, COPD, depression, and PVD. Individual-level random effects were included in these models.

†Difference in per member per month costs or utilization for HP participants compared with controls (Control−HP participants) in the post study entry follow-up period.

Analyses based on the level of HP program participation demonstrated a broad, general effect of HP programs. Only 10.3% of employees in the HP study arm had an HRA only, and a high degree of overlap and correlation was found for the categorical classifications for HP program participation. We observed a statistically significant association between the cumulative measure of HP program participation and reduced total mean monthly health care costs ($P < 0.05$). Based on the dichotomous classification of participation intensity, a greater impact was observed among employees participating in more intense 12-month coaching programs as shown in Table 5. Although the overall cost reduction among all HP participants was $35.1 per member per month, participants in the two high-intensity programs experienced a mean reduction in plan paid costs of $88.2 per member per month.

Sensitivity Analyses

We conducted a number of sensitivity analyses to evaluate robustness of our primary effect estimates, and these results are shown in Table 6. These additional analyses were based on total plan paid costs. Mixed models based on propensity score adjustment rather than inverse propensity score weighting found a mean monthly cost savings of $39.7 ($P = 0.0363$) for HP program participants compared with controls. The propensity score-adjusted exponential conditional means GEE model found a comparable mean monthly reduction of $39.5 ($P = 0.0071$).

Analyses based on propensity score stratification and direct calculation of mean differences in treatment effects within propensity strata resulted in an estimated mean monthly total cost reduction of $22 ($P = 0.0100$) among HP participants compared with controls. Analyses based on individual matching resulted in 2951 HP participants matched to 2951 controls. Although the matched analyses had reduced statistical power because of exclusion of 878 HP participants and 3666 controls, HP program participants were found to have a mean monthly total paid cost savings of $33.0 ($P = 0.1540$). The effect point estimate from the matched analyses is thus comparable to the effect seen with models using statistical adjustment.

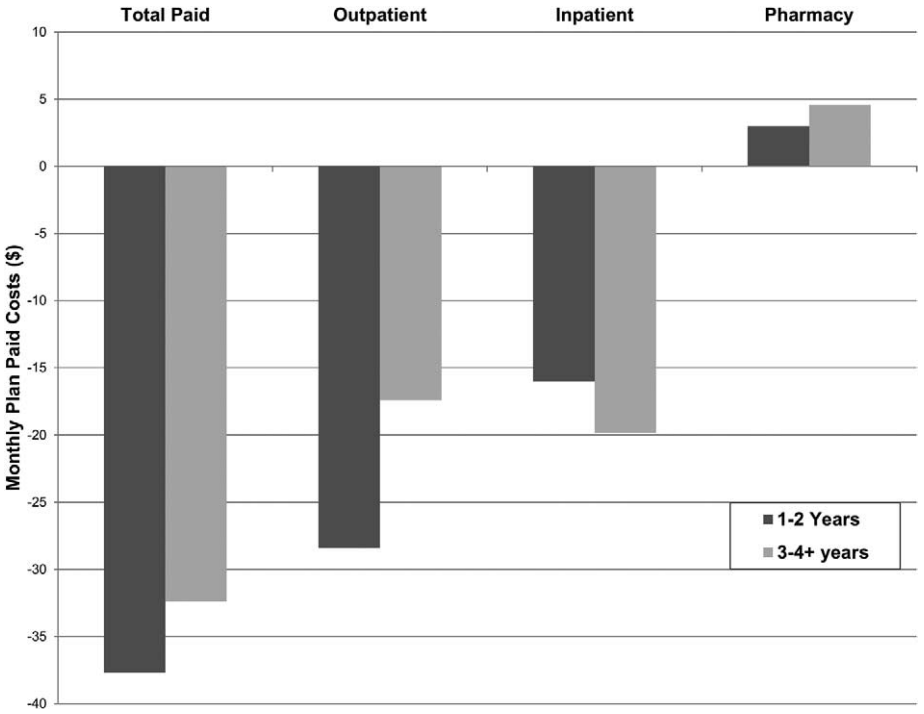

FIGURE 3. Cost differences by category and time since program engagement.

We assessed the potential impact of excluding cancer and pregnancy claims in our primary model for monthly plan paid costs. Exclusion of cancer claims reduced the total paid cost savings by approximately 10% (to $31.4) and exclusion of pregnancy-related claims reduced the monthly savings less than 3% (to $34.3). Support for inclusion of cancer and pregnancy claims is provided in the literature. Both diet and physical activity have been shown to reduce risk of pregnancy-related conditions such as gestational diabetes[42] and to decrease inflammatory markers such as C-reactive protein, associated with risk of some cancers.[43–46]

Lastly, we ran an additional model for total paid costs that included a dichotomous covariate to indicate individuals (HP participants and controls) with total paid costs in excess of $100,000 in any study year. After model adjustment for high-cost employees, HP participants experienced an estimated mean monthly total paid cost savings of $26.2 ($P = 0.0898$) compared with controls. Adjustment for high-cost individuals likely represents an overly conservative approach, as these individuals are generally those of interest with regard to cost savings and those who may benefit most from HP program participation. Taken as a whole, these additional sensitivity analyses suggest that cost reductions observed in the primary analyses are reasonably robust with regard to study design, analytical approach, inclusion or exclusion of cancers and pregnancy claims, and high-cost outliers.

Return on Investment

Results of the ROI analyses that included data for 2006 to 2011 are shown in Table 7. The ROI estimate is based on our study best estimate of $35 per member per month health care cost savings associated with HP program participation and HP program participation data. We estimate a health care cost savings of $2.53 for each dollar invested in the HP program.

DISCUSSION

This study found that employees participating in HP programs experienced significantly reduced plan paid health care costs and health care utilization compared with controls, after adjustment for differences in demographics, disease prevalence, and measures of comorbidity. A much earlier study of this employee population found HP program participation to reduce health risk factors.[47] The primary study results found a mean monthly reduction in total costs of $35.1, and sensitivity analyses provided additional support for this result. Given that only 10.3% of the HP study arm had only an HRA, these results should be considered to apply to employees participating in HP programs beyond just completing an HRA. A prior review by the Task Force on Community Preventive Services concluded that there was insufficient evidence to determine the effectiveness of interventions such as an HRA with feedback when implemented by itself as a primary intervention.[48] HP participation

TABLE 5. Changes in Mean Monthly Health Care Costs by HP Participation Level*

Maximum HP program participation intensity	Difference ($)† (Controls–HP)	Difference 95% CI	Effect *P* Value
Low to medium intensity	−27.3	−62.7 to 8.02	0.1297
High intensity‡	−88.2	−151.1 to −25.3	0.0060

CI, confidence interval; HP, health promotion.

*Inverse propensity score-weighted mixed models, including covariates for age, sex, race/ethnicity, health plan, CDPS-Rx comorbidity risk score, and dichotomous covariates for diabetes, hypertension, CAD, CHF, asthma, COPD, depression, and PVD. Individual-level random effects were included in these models.

†Difference in per member per month costs or utilization for HP participants compared with controls.

‡High intensity programs include Health Coaching. Employees were assigned to the highest level of HP program participation achieved during the follow-up period.

TABLE 6. Sensitivity Analyses Results for Primary Study Outcome

Statistical Method for Bias Adjustment	HP Program Estimated Monthly Cost Saving ($)	*P* Value for Estimated Cost Savings
Model-based estimates		
Propensity score-weighted mixed model*	−35.1	0.0465
Propensity score-adjusted mixed model†	−39.7	0.0363
Propensity score-adjusted GLM model‡	−39.5	0.0071
Propensity score matching or stratification estimates		
Propensity score-matched analyses§	−33.0	0.1540
Propensity score-stratified Analyses‖	−22.0	0.0100

GLM, generalized linear model; HP, health promotion.

*Primary study results based on an inverse propensity score-weighted mixed model with adjustment for mean monthly costs at baseline, time period, age, sex, race/ethnicity, plan type, and baseline comorbidities.

†Propensity score adjusted mixed model with adjustment for propensity score quintile, mean monthly costs at baseline, time period, age, sex, race/ethnicity, plan type, and baseline comorbidities.

‡Propensity score adjusted difference-in-difference GEE exponential conditional mean model with adjustment for propensity score quintile, time period, age, sex, race/ethnicity, plan type, and baseline comorbidities.

§1-to-1 propensity score-matched analysis with difference between the outcome means in the HP and control group tested using a correlated means t-test. These analyses based on 2951 HP participants matched to 2951 controls, resulting in reduced statistical power compared with the study primary analyses.

‖Overall estimates of effect of HP participation estimated based on inverse variance-weighted propensity score quintile stratum-specific mean effects.

was complex and highly correlated; however, these data provide some support for a greater impact of more intense programs such as 12-month duration health coaching.

It is likely that certain program design elements contributed to the higher levels of participant engagement for this employee population compared with some other HP programs with ROI estimates. The HRA being administered by an HP program staff registered nurse as part of pre-employment and offered at various worksites across the organization on an ongoing basis made it easily accessible. The HRA also included a feedback and goal setting session and immediate referral to coaching and other appropriate HP program and organizational programs/services. In addition, the structure of the coaching program provided a combination of in-person and telephonic coaching monthly, with biometric feedback on a quarterly basis. Overall, the HP program offered a wide variety of program options delivered in-person, online, and telephonically providing the participant multiple engagement options.

Results of published studies concerning the economic impact of HP and disease prevention programs have widely varied. In a recent review, Lerner et al[3] identified 13 published studies with 10 reporting favorable results. Unfortunately, various deficiencies were noted in most of the studies reporting favorable results, and only two studies were judged to exhibit relatively high-quality research methods. The high-quality studies included the publication by Naydeck et al,[13] which evaluated the Highmark, Inc programs and found an average annual savings of $176 for program participants over the 3-year study period following program initiation in 2002. Their study used risk matching to control for potential confounders, but did not exclude employees who engaged in disease management. In contrast to our findings, Naydeck et al showed that the largest savings were associated with inpatient expenses. This study also found an overall significant effect regardless of whether employees only completed an HRA, participated in coaching, online, group or individual programs, or visited the fitness center along with engaging in other wellness programs. The only statistically significant program-specific effect was for participants who used an HRA and the fitness center.

A recently published study provides additional evidence to support positive impacts of HP and disease prevention programs.[14] This study evaluated the impact of PepsiCo's health and wellness program on medical cost and utilization. Liu et al found that health and wellness program participants experienced an average reduction in health care costs of $38 per member per month over the 3-year study period. Significant decreases were also observed in ED visits and hospital admissions. Unlike our analysis this study included participants in disease management programs. Significant savings were associated with participation in any HP program or in the disease management program, but not the lifestyle management programs.

Several additional recent financial analytic studies of HP programs found positive savings and ROI. In 2015, Musich et al[15] published an evaluation of the Well at Dell Health Management Program. This well-designed study demonstrated cost savings of $40 per member per month for participants in the HRA, year-round HP, and wellness campaigns. No significant savings were found for lifestyle management or disease management participation. Our ROI of 2.53 is similar to the ROI of 2.48 reported by Musich et al.

In 2013, Grossmeier et al[9] evaluated financial impact of the newly initiated BP America population health management program and found a $60 per member per month savings for participants HP and an ROI of 3.00. Our study is similar to Grossmeier et al and Musich et al, using robust statistical methods to account for participation bias. Program elements, however, differed between these studies and ours. BP included the employee assistance program and disease management. The ''Well at Dell'' study also

TABLE 7. Estimated HP Program ROI

Year	Eligible Population	HP Program Costs Adjusted to 2011 Dollars	Estimated HP Program Participants	Estimated HP Program Health Care Costs Savings*
2006	24,940	1,486,646	11,741	4,931,220
2007	25,186	1,497,753	12,863	5,402,460
2008	26,125	1,473,964	13,748	5,774,160
2009	28,034	1,479,179	13,957	5,861,940
2010	28,859	1,484,987	10,518	4,417,560
2011	30,000	1,476,629	11,927	5,009,340
Total		8,899,157	74,754	31,396,680
Estimated ROI (savings/costs) = 2.53				

HP, health promotion; ROI, return on investment.

*Cost savings based on study best estimate of $35 per member per month health care cost savings associated with HP program participation.

included disease management. Both the Well at Dell and BP programs included premium incentives for participants, a feature not adopted by our administration.

Parkinson et al[16] evaluated the impact of a health insurance incentivized wellness, prevention, and chronic disease management program among health care workers established in 2005, and observed over a 5-year time period. Significant improvements in health risk levels attributable to the programs were observed, and significantly lower health care costs were observed among members who reduced risk levels relative to a propensity score-matched comparison group. Members who reduced risks from moderate to low risk experienced a reduction of $73 per member per month in total health care costs, whereas those moving from high to low risk experienced a much larger reduction of $246 per member per month. This study did not provide an ROI estimate.

Findings from recent well-designed studies differ regarding the specific elements of HP programs associated with savings. We found greater savings associated with higher intensity participation. Other studies have not reported findings from similar analyses. To examine effects of the HP program separate from disease management, we excluded any disease management program participants, thus our ROI might be less compared with studies that investigated these programs in combination.

The scope of our evaluation was limited to impacts of the HP program on direct employer-paid health care costs. Workplace wellness programs generate a range of non-health related benefits as well as indirect benefits that provide substantial value to employers not captured in ROI analyses.[1] Indirect benefits associated with HP programs may include increased attraction and retention of top talent, public recognition of the sponsoring organization as an employer of choice, increased employee engagement with their employer, and increased employee productivity.[9]

STRENGTHS AND LIMITATIONS

This study has several strengths including a reasonably large study population of HP program participants and controls observed for up to 5 years after initial HP program participation. In addition, the HP program was very mature at the time of this evaluation having been initiated in 1988, assuring greater overall integration into the workplace culture and minimizing potential startup effects. The study design used several methods to control for potential bias related to HP program participation, and robustness of primary study findings was supported through sensitivity analyses using other designs and analytical approaches.

Although we made every effort to control for potential covariate imbalance and bias, concerns may remain over the generalizability of estimates obtained and whether use of propensity scores within the statistical models removed bias. Apparent balance between compared groups on measured variables could leave differences in nonobservable or unmeasured variables not included in the propensity score unbalanced and lead to confounded effect estimates. Some additional assurance of adequate covariate balance and removal of bias is provided through the sensitivity analyses, which supported the primary study results. An additional limitation is a possible measurement bias in the categorization of participants into the various HP program categories. There may have been individuals placed in the control group who were actually physically active or actively pursuing wellness activities outside the employer HP programs.

CONCLUSIONS

This study supports a positive impact of HP program participation with regard to reduced health care costs and health care utilization. The plausible range of total cost savings per member per month was $22 to $40 with a study best estimate of approximately $35. ROI was estimated to be $2.53 for every dollar spent on HP. These results are comparable to results observed in well-designed studies of the impacts of comprehensive and well-integrated HP programs. HP programs may have numerous other indirect benefits that are difficult to quantitate, but are none-the-less important with regard to employer decisions to invest in these programs.

Our results may not be generalizable to smaller employers implementing new HP programs. Our long-standing HP program has become ingrained into the workplace culture. The program is comprehensive as defined in Healthy People 2010, and provides health education in a supportive social and physical environment. Free services are provided to employees who volunteer to participate. The program is proactive visiting worksites to collect HRAs and provide special programming. Participants are linked to related programs including EAP and disease management. Owing to the long-standing availability and acceptance of HP programs, managers regularly utilize HP programs when seeking to improve work culture. These features along with efficient programming costs likely contribute to our findings of a positive ROI.

REFERENCES

1. Pronk NP. Placing workplace wellness in proper context: value beyond money. *Prev Chronic Dis*. 2014;11:E119.
2. Centers for Disease Control and Prevention. Workplace Health Program Definition and Description. Centers for Disease Control and Prevention. Available at: http://www.cdc.gov/workplacehealthpromotion/pdfs/Workplace_Health_Program_Definition_and_Description.pdf. Accessed March 11, 2015.
3. Lerner D, Rodday AM, Cohen JT, Rogers WH. A systematic review of the evidence concerning the economic impact of employee-focused health promotion and wellness programs. *J Occup Environ Med*. 2013; 55:209–222.
4. Baxter S, Sanderson K, Venn AJ, Blizzard CL, Palmer AJ. The relationship between return on investment and quality of study methodology in workplace health promotion programs. *Am J Health Promot*. 2014;28:347–363.
5. Osilla KC, Van Busum K, Schnyer C, Larkin JW, Eibner C, Mattke S. Systematic review of the impact of worksite wellness programs. *Am J Manag Care*. 2012;18:e68–e81.
6. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Aff (Millwood)*. 2010;29:304–311.
7. Goetzel RZ, Henke RM, Tabrizi M, et al. Do workplace health promotion (wellness) programs work? *J Occup Environ Med*. 2014;56:927–934.
8. Goetzel RZ, Tabrizi M, Henke RM, et al. Estimating the return on investment from a health risk management program offered to small Colorado-based employers. *J Occup Environ Med*. 2014;56:554–560.
9. Grossmeier J, Seaverson EL, Mangen DJ, et al. Impact of a comprehensive population health management program on health care costs. *J Occup Environ Med*. 2013;55:634–643.
10. Nyman JA, Barleen NA, Abraham JM. The effectiveness of health promotion at the University of Minnesota: expenditures, absenteeism, and participation in specific programs. *J Occup Environ Med*. 2010;52:269–280.
11. Nyman JA, Barleen NA, Dowd BE. A return-on-investment analysis of the health promotion program at the University of Minnesota. *J Occup Environ Med*. 2009;51:54–65.
12. Nyman JA, Jeffery MM, Abraham JM, Jutkowitz E, Dowd BE. The 3-year disease management effect: understanding the positive return on investment. *J Occup Environ Med*. 2013;55:1356–1364.
13. Naydeck BL, Pearson JA, Ozminkowski RJ, Day BT, Goetzel RZ. The impact of the highmark employee wellness programs on 4-year healthcare costs. *J Occup Environ Med*. 2008;50:146–156.
14. Liu H, Harris KM, Weinberger S, Serxner S, Mattke S, Exum E. Effect of an employer-sponsored health and wellness program on medical cost and utilization. *Popul Health Manag*. 2013;16:1–6.
15. Musich S, McCalister T, Wang S, Hawkins K. An evaluation of the Well at Dell health management program: health risk change and financial return on investment. *Am J Health Promot*. 2015;29:147–157.
16. Parkinson MD, Peele PB, Keyser DJ, Liu Y, Doyle S. UPMC MyHealth: managing the health and costs of U.S. healthcare workers. *Am J Prev Med*. 2014;47:403–410.
17. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. *Am J Health Promot*. 1997;12:38–48.
18. Williamson E, Morley R, Lucas A, Carpenter J. Propensity scores: from naive enthusiasm to intuitive understanding. *Stat Methods Med Res*. 2012;21: 273–293.

19. Dement JM, Pompeii LA, Ostbye T, et al. An integrated comprehensive occupational surveillance system for health care workers. *Am J Ind Med*. 2004;45:528–538.
20. Dement JM, Lipscomb HJ, Schoenfisch AL, Pompeii LA. Impact of hospital Type II violent events: use of psychotropic drugs and mental health services. *Am J Ind Med*. 2014;57:627–639.
21. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43:1130–1139.
22. Elixhauser A, Steiner C, Palmer L. Clinical Classifications Software (CCS). U.S. Agency for Healthcare Research and Quality; 2012. Available at: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed January 3, 2013.
23. NCQA. *Healthcare Effectiveness Data and Information Set (HEDIS), Antidepressant Medications Management, Table AMM-D*. Washington, DC: National Committee for Quality Assurance; 2012.
24. NCQA. Healthcare Effectiveness Data and Information Set (HEDIS), Comprehensive Diabetes Care and Controlling High Blood Pressure (CBP). Washington, DC: National Committee for Quality Assurance; 2008.
25. Kronick R, Gilmer T, Dreyfus T, Lee L. Improving health-based payment for Medicaid beneficiaries: CDPS. *Health Care Financ Rev*. 2000;21:29–64.
26. Kronick R, Gilmer T, Dreyfus T, Ganiats TG. *The Chronic Illness and Disability Payment System for Medicare Beneficiaries: Final Report to CMS*. San Diego, CA: University of California; 2003.
27. Landers G, Zhou M. The impact of Medicaid peer support utilization on cost. *Medicare Medicaid Res Rev*;. 2014;4.
28. Dominick K, Dudley T, Coffman C, Bosworth H. Comparison of three comorbidity measures for predicting health service use in patients with osteoarthritis. *Arthritis Rheum*. 2005;53:666–672.
29. Linden A. Designing a prospective study when randomization is not feasible. *Eval Health Prof*. 2011;34:164–180.
30. Rosenbaum PRD. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
31. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. *J Clin Epidemiol*. 2005;58:550–559.
32. Suarez D, Faries, DE. Propensity score stratification and regression. In: Faires DE, Leon AC, Haro JM, Obenchain RL, eds. *Analysis of Observational Health Care Data Using SAS*. Cary, NC: SAS Institute Inc; 2010, 23–49.
33. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods*. 2010;15:234–249.
34. Austin PC. Goodness-of-fit diagnostics for the propensity score model when estimating treatment effects using covariate adjustment with the propensity score. *Pharmacoepidemiol Drug Saf*. 2008;17:1202–1217.
35. Musich S, White J, Hartley SK, Bhattarai GR, Hawkins K, Ozminkowski RJ. A more generalizable method to evaluate changes in health care costs with changes in health risks among employers of all sizes. *Popul Health Manag*. 2014;17:297–305.
36. Ozminkowski RJ, Goetzel RZ, Wang F, et al. The savings gained from participation in health promotion programs for Medicare beneficiaries. *J Occup Environ Med*. 2006;48:1125–1132.
37. White J, Hartley SK, Musich S, Hawkins K, Ozminkowski RJ. A more generalizable method to evaluate the association between commonly reported health risks and health care expenditures among employers of all sizes. *J Occup Environ Med*. 2013;55:1179–1185.
38. Liang K, Zeger Z. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13–22.
39. Nyman JA, Abraham JM, Jeffery MM, Barleen NA. The effectiveness of a health promotion program after 3 years: evidence from the University of Minnesota. *Med Care*. 2012;50:772–778.
40. Lanehart RE, Gil PRD, Kim ES, et al. Propensity score analysis and assessment of propensity score apparoches using SAS® procedures. In: *SAS Blobal Forum 2012, Statistics and Data Analysis*. Cary, NC: SAS, Inc; 2012, 1–11.
41. Lunceford JK, Davidian M. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Stat Med*. 2004;23:2937–2960.
42. Simmons D, Jelsma JG, Galjaard S, et al. Results from a European multicenter randomized trial of physical activity and/or healthy eating to reduce the risk of gestational diabetes mellitus (GDM): the DALI lifestyle pilot. *Diabetes Care*;. 2015.
43. Adams SA, Wirth MD, Khan S, et al. The association of C-reactive protein and physical activity among a church-based population of African Americans. *Prev Med*. 2015;77:137–140.
44. Hebert JR, Wirth M, Davis L, et al. C-reactive protein levels in African Americans: a diet and lifestyle randomized community trial. *Am J Prev Med*. 2013;45:430–440.
45. Camhi SM, Stefanick ML, Ridker PM, Young DR. Changes in C-reactive protein from low-fat diet and/or physical activity in men and women with and without metabolic syndrome. *Metabolism*. 2010;59:54–61.
46. Erlinger TP, Platz EA, Rifai N, Helzlsouer KJ. C-reactive protein and the risk of incident colorectal cancer. *JAMA*. 2004;291:585–590.
47. Goetzel RZ, Kahr TY, Aldana SG, Kenny GM. An evaluation of Duke University's Live for Life health promotion program and its impact on employee health. *Am J Health Promot*. 1996;10:340–342.
48. Task Force on Community Preventive Services. Recommendations for worksite-based interventions to improve workers' health. *Am J Prev Med*. 2010: S232–S236.

