Effects of a Workplace Wellness Program on Employee Health, Health Beliefs, and Medical Utilization: A Randomized Clinical Trial[†]

April 2020

Julian Reif, PhD, University of Illinois and NBER

David Chan, MD, PhD, Stanford University and NBER

Damon Jones, PhD, University of Chicago and NBER

Laura Payne, PhD, University of Illinois

David Molitor, PhD, University of Illinois and NBER

Word count: 2689

[†] Corresponding author: David Molitor, Gies College of Business, University of Illinois, 340 Wohlers Hall, 1206 S. Sixth St., Champaign, IL 61820. Phone: 217-244-0504. Email: dmolitor@illinois.edu.

Key Points

Question: How did a comprehensive workplace wellness program affect health, health beliefs, and medical utilization after 24 months?

Findings: In a two-year randomized controlled trial of 4,834 employees at a large US university, employees invited to join a wellness program showed no significant differences in biometrics, medical diagnoses, or medical utilization relative to the control group. The intervention increased self-reports of having a primary care physician and improved a set of employee health beliefs among the treatment group.

Meaning: Workplace wellness changed health beliefs and increased self-reports of having a primary care physician but did not significantly affect clinical outcomes.

Abstract

Importance: Many employers use workplace wellness programs to improve employee health and reduce medical costs, but randomized evaluations of their efficacy are rare.

Objective: To evaluate the effect of a comprehensive workplace wellness program on employee health, health beliefs, and medical utilization after 12 and 24 months.

Design: This study was a randomized controlled trial of 4,834 employees. Members of the treatment group (n=3300) received incentives to participate in the workplace wellness program. Members of the control group (n=1534) did not participate in the wellness program.

Setting: Large US university.

Participants: Diverse set of employees.

Interventions: The two-year program included financial incentives and paid time off for annual on-site biometric screenings, annual health risk assessments, and ongoing wellness activities (e.g., physical activity, smoking cessation, disease management).

Main Outcomes and Measures: Measures taken at 12 and 24 months included clinician-collected biometrics (16 outcomes); administrative claims related to medical diagnoses (diabetes, hypertension, hyperlipidemia) and utilization (office visits, inpatient visits, emergency room visits); and self-reported health behaviors and health beliefs (14 outcomes).

Results: We found no significant effects on biometrics, medical diagnoses, or medical utilization after 12 or 24 months. A significantly higher proportion of employees in the treatment group reported having a primary care physician after 24 months (92.2% vs. 86.1%; adjusted $P < 0.01$). The intervention significantly improved a set of employee health beliefs on average, although effects on individual belief measures were not significant.

Conclusions and Relevance: This randomized controlled trial showed that a comprehensive workplace wellness program had no significant effects on measured physical health outcomes, rates of medical diagnoses, or the use of health care services after 24 months, but it increased the proportion of employees reporting they have a primary care physician and improved employee beliefs about their own health.

Trial Registration: American Economic Association Randomized Controlled Trial Registry number AEARCTR-0001368 (http://www.socialscienceregistry.org/trials/1368).

Introduction

Employers increasingly offer workplace wellness programs to reduce health care costs and improve employee health. Among large US firms offering health benefits in 2019, 84% also offered a wellness program.[1] The wellness industry has grown rapidly since the passage of the 2010 Affordable Care Act, which encouraged firms to adopt wellness programs by raising the maximum limit on financial incentives offered to program participants.

However, evidence of the causal effects of workplace wellness programs is limited. Observational studies that compare participants to nonparticipants are susceptible to selection bias.[2] Randomized trials frequently evaluate narrow wellness interventions with only one or two program components and examine only a few outcomes.[3-8] Reviews of the literature have yielded mixed results and raised concerns about publication bias.[9,10] A recently published randomized controlled trial (RCT) with 160 randomized units reported outcomes at 18 months post-intervention.[11] Another recently published RCT with 4,834 randomized units reported effects on medical spending and employee productivity, but not clinical outcomes.[2] Neither study investigated the effect of workplace wellness programs on employees' beliefs about their own health. Measuring these beliefs sheds light on employees' perceptions about the effectiveness of participating in wellness programs. These beliefs may also shape how much value and effort individuals place on health behaviors, a channel emphasized by social cognitive theory.[12,13]

In this study, individual employees were randomly assigned to a treatment group, which was eligible to participate in a two-year comprehensive workplace wellness program, or a control group, which was not eligible. We evaluated the effects of this program on health beliefs, self-

reported health behaviors, clinician-collected biometrics, and claims-based medical diagnoses and medical utilization, over the 24 months following initial randomization into the program.

Methods

Study Design

We conducted an RCT of a workplace wellness program among employees of the University of Illinois at Urbana-Champaign (UIUC). Among the study population, 3,300 were randomly assigned to be eligible for program participation (the treatment group). The other 1,534 study participants were ineligible for the program (the control group). Randomization was stratified by employee class, sex, age, salary, and race (Section 1.2 of Supplement 2). We specified the research design, subgroup analysis, and the health belief, biometric, and medical utilization outcomes prior to analysis. Self-reported health behavior and medical diagnosis outcomes were specified post hoc. Our pre-analysis protocol was publicly archived and is available in Supplement 1. The UIUC, University of Chicago, and National Bureau of Economic Research institutional review boards approved the study.

Study Participants

A total of 12,459 benefits-eligible UIUC employees were invited in July 2016 to enroll in the study and complete a survey (Figure). Employees were informed that they might be selected for further participation in the study, but no other details about the intervention were disclosed prior to enrollment. Invitations were sent by mail and email (Figures S4-S6 in Supplement 2). Our study population included 4,834 employees who enrolled in the study over a three-week enrollment period. All study participants provided written informed consent. Random assignment

of study participants to treatment and control groups occurred in August 2016, after study enrollment had closed.

Intervention

A comprehensive workplace wellness program named iThrive was introduced at UIUC and ran for two years, August 2016–April 2018. The program, designed to be representative of typical comprehensive wellness programs offered by employers, included three annual components: an on-site biometric screening and survey, an online health risk assessment (HRA), and a choice of wellness activities.[14] Employees in the treatment group were eligible to participate in all three intervention components using paid time off and received randomly assigned cash awards that ranged from $0 to $200 per year for completing the annual screening and HRA. Treatment group participants who completed the biometric screening and HRA were then eligible to register for one wellness activity class per semester, for a total of two activities per year. Classes ranged from 6 to 12 weeks in length and addressed numerous topics (e.g., physical activity, nutrition, stress management) (eTables 2–3, 5–6 in Supplement 2). Upon completion of a wellness activity, participants earned $0 to $75 as a cash reward or Amazon.com gift card. The on-site biometric screenings and surveys were administered by local clinicians. The HRA was designed by Wellsource, an established wellness vendor. The wellness activities were selected and implemented by UIUC's director of Campus Wellbeing Services. Details are provided in Supplement 2.

Employees in the control group were invited to complete the on-site biometric screening and survey in August 2017 (12 months following randomization) and in August 2018 (24 months following randomization) in order to serve as a comparison group. Control group employees

were not eligible to participate in the first on-site biometric screening and survey in August 2016 and were never eligible to participate in any of the HRAs or wellness activities offered throughout the two-year iThrive program. Although the research team never informed the control group about the intervention, some may have learned about it from coworkers. To assess how often this occurred, a 2017 online survey asked study participants whether they ever communicated about iThrive with coworkers. Only 3% of the control group responded affirmatively, compared to 44% of the treatment group.[2]

Outcome Measures

Health beliefs, self-reported health behaviors, and biometrics were collected on-site by clinicians. Study participants were asked to report their height and weight. They also reported, on a scale from 0 to 100, their expected chances (subjective probabilities) of having high cholesterol, high blood pressure, impaired fasting glucose, and a body mass index (BMI) over 30 (Figure S17 in Supplement 2). We interpret self-reported height and weight and these expected chances as measures of participants' health beliefs.[12,15,16] Study participants were then directed to a station where a clinician measured their height, weight, waist circumference, and blood pressure. The clinician also measured their cholesterol (total, HDL, and LDL), triglycerides, and glucose levels using a CardioChek Plus Analyzer, and recorded their answers to questions about tobacco use, physical activity, mood, and having a primary care physician (Figure S18 in Supplement 2).

Administrative health claims data were obtained for employees enrolled in UIUC's Health Alliance insurance plan, which covers 67% of employees in our sample. These data include all

inpatient, outpatient, and prescription drug claims with a date of service between October 1, 2015, and July 31, 2018.

Statistical Analysis

We performed power calculations for all outcomes by calculating ex-post minimum detectable effects.[17] The results are provided in Supplement 3.

We estimated the effect of being invited to participate in the iThrive wellness program in the available population. Some employees in our sample ceased employment with the university over the course of the 24-month study. For administrative health claims outcomes, we restricted comparisons to employees enrolled in Health Alliance. For all other outcomes, we compared participants in the treatment group who completed the follow-up (2017 or 2018) on-site screening and survey to all employees in the control group who completed the follow-up (2017 or 2018) on-site screening and survey (Figure). Baseline characteristics between treatment and control were compared to evaluate the potential for bias due to missing data (Section 2 of Supplement 3).[18]

For each outcome, we estimated an individual-level linear model with a binary indicator for treatment assignment as the key independent variable. For biometric and self-reported outcomes, we included all study participants who completed the on-site follow-up screening and survey in 2017 ($n = 2004$) or 2018 ($n = 1761$). For medical diagnoses and medical utilization outcomes, we included all study participants ($n = 4834$) and weighted each individual by the number of months with Health Alliance insurance coverage. We included baseline values of the outcome (when available) and stratification variables as controls in our linear model to improve precision. Analyses were performed using Stata, version 15.[19] We calculated standard errors that are robust

to arbitrary heteroscedasticity and employed two-tailed tests with a significance level of $P = 0.05$.

To account for less-than-universal participation among the treatment group, we used an instrumental-variable approach to estimate the local average treatment effect of participating in the program, instrumenting participation with assignment into the treatment group.[11,20,21] Participation was defined as completing the first (2016) screening component, which was offered only to members of the treatment group (Figure). The results are provided in Supplement 3.

Because we estimated our model for many outcomes, the probability that we incorrectly reject at least one null hypothesis is greater than the significance level used for each individual hypothesis test. We accounted for this multiple testing concern in two ways. First, we calculated a standardized treatment effect for a "family" of outcomes by dividing the estimate for each individual outcome by its standard deviation and then averaging across all the outcomes within the family.[11,22] This method gives equal weight to each outcome in the family, which may be undesirable. Therefore, we also used resampling to calculate an adjusted *P* value for each outcome that corrects for the number of hypothesis tests within a family of outcomes.[2,23] We consider effects to be statistically significant if the adjusted *P* value or the standardized treatment effect *P* value is less than 0.05.

Results

Baseline Characteristics and Program Participation

Table 1 reports baseline characteristics for the treatment ($N = 3300$) and control ($N = 1534$) groups. Among all 4,834 study participants, the mean (SD) age was 44 (11), 57% were female,

16% were non-white, 20% were faculty, and 24% earned less than $40,000 per year. About 67% of the sample was enrolled in Health Alliance insurance coverage during the 10-month pre-intervention period October 2015–July 2016. Among this subsample and during this time, study participants had 2.5 outpatient visits on average and had medical claims with diagnoses codes related to three common chronic conditions in the following proportions: diabetes (5%), hypertension (14%), and hyperlipidemia (16%). Inpatient and emergency room visits were uncommon in this sample. Overall, baseline participant characteristics were well-balanced across both study arms.

Of the 3,300 participants in the treatment group, 56% completed both the biometric screening and online HRA in the first year, and 31% completed the biometric screening, online HRA, and at least one wellness activity in the first year. Over the course of the two-year program, 64% of the treatment group completed at least one component of the iThrive wellness program. These completion rates are similar to those reported for other comprehensive wellness programs.[11,14]

Effects of the Intervention

Table 2 reports effects of the intervention on health beliefs and self-reported health behaviors. When combined into a standardized treatment effect, participant beliefs about their chance of having a BMI > 30, high cholesterol, high blood pressure, and impaired glucose jointly decreased by 0.07 standard deviations (95% CI, -0.12 to -0.01; $P = 0.02$). While these health beliefs changed significantly as a group, changes in specific measures of health beliefs were less precise and thus not individually significant.

Self-reports of having a primary care physician significantly increased by 6.13 percentage points (95% CI, 3.04 to 9.22; adjusted $P < 0.01$) after 24 months. There were no significant effects on self-reported tobacco use, physical activity intensity, or mood after 12 or 24 months.

Table 3 reports that the intervention had no significant effects on height, weight, waist circumference, body mass index, blood pressure, cholesterol, or glucose.

Table 4 reports no significant changes in diagnoses of hypertension, diabetes, or hyperlipidemia after 12 or 24 months. Likewise, no significant effects were found for office visits, inpatient visits, or emergency room visits. Additional analysis also found no significant effects for primary care physician visits (eTable 25 of Supplement 3).

Subgroup analysis

eTables 7–24 of Supplement 3 report effects for prespecified subgroups. Compared to women, men had higher effects on claims-based diabetes diagnoses after 12 months (2.38%; 95% CI, 0.60% to 4.16%; adjusted $P = 0.05$), but not after 24 months (1.53%; 95% CI, −0.58% to 3.65%; adjusted $P = 0.65$) (eTable 9). Compared to younger employees, employees ages 50 and over had lower effects on self-reports of having a primary care physician after 24 months (−9.86%; 95% CI, −15.07% to −4.66%; adjusted $P < 0.01$) (eTable 10). No significant heterogeneity was found with respect to race, employee classification (faculty, civil service, or academic professional), or salary.

Discussion

This individual-level randomized controlled trial of a two-year comprehensive workplace wellness program demonstrated that the program significantly improved employee beliefs about their own health and increased the proportion of employees reporting they have a primary care physician. However, no significant effects were found on biometrics, medical diagnoses, or medical utilization after 24 months. Our study was powered to detect clinically meaningful effects across these three domains. The 95% confidence interval on systolic blood pressure after 24 months rules out a decrease of 1.48 mm Hg, compared to a control group mean of 122.4 mm Hg. The 95% confidence interval on diagnoses of hyperlipidemia after 24 months rules out a decrease of 2.47 percentage points, compared to a control group mean of 26.5%. Likewise, the 95% confidence interval on office visits rules out a decrease of 0.30, compared to a control group mean of 6.67.

These results complement recent RCT evidence that workplace wellness programs affect some self-reported outcomes but have limited effects on clinical or administrative outcomes. Prior findings showed that the iThrive program increased self-reported lifetime health screening rates and improved employee perceptions of management, but did not significantly affect administrative measures of medical spending.[2] A cluster randomized controlled trial of a wellness program at BJ's Wholesale Club found significant effects on self-reports of engaging in regular exercise and actively managing weight but found no significant effects on medical spending or biometric outcomes after 18 months.[11] The similarity in these RCT findings using different randomization designs in different populations increases confidence in their reliability and generalizability.

Our measures of health beliefs, elicited using self-reported subjective probabilities, are a distinct contribution to the literature on wellness interventions. Employees in the treatment group believed they had lower chances of poor biometric health, suggesting that they expected their participation in the wellness program to improve their health. However, there was no significant effect of the program on biometrics or medical utilization, and prior findings showed no significant effects on administratively measured health behaviors.[2] These results demonstrate a mismatch between employee perceptions and physical/administrative measures of health.

Findings from the Illinois Workplace Wellness Study and the BJ's study, both RCTs, differ from many prior studies that found wellness programs improve employee health and reduce medical utilization. Many of these prior studies use observational research designs, which can result in significant selection bias even after controlling for many covariates.[2] Findings from RCTs are less susceptible to selection bias.

Limitations

This study has several limitations. The results may not generalize to other workplace settings with different populations or different wellness programs.[24] Our 95% confidence intervals do not rule out meaningful effects for some outcomes—such as a decrease in ER visits after 24 months of 0.1, compared to a control group mean of 0.28. Also, the outcomes were measured over the first 24 months following randomization. We do not know whether the significant effects on self-reported outcomes persisted beyond 24 months, or whether detectable effects on biometrics, medical diagnoses, or utilization emerged beyond 24 months.

Finally, data were not available for all study participants. Medical diagnoses and utilization outcomes were obtained only for participants enrolled in Health Alliance. Biometric and self-

reported outcomes were obtained only for participants who completed the on-site screening and survey in 2017 or 2018. However, Health Alliance enrollment was well-balanced between treatment and control (Table 1). Baseline characteristics of participants who completed the on-site screenings and surveys were well-balanced between treatment and control (eTables 2–3 in Supplement 3). The balance between treatment and control groups suggests that bias from missing data is unlikely to be substantial.

Conclusions

Among workers of a large employer, a comprehensive workplace wellness program significantly changed a set of beliefs about biometric outcomes and significantly increased self-reports of having a primary care physician, but no significant effects on clinician-measured biometrics, medical diagnoses, or medical utilization were found after 24 months. These findings shed light on employees' perceptions of workplace wellness programs, which may influence long-run effects. However, we add to a growing body of evidence from randomized trials that workplace wellness programs are unlikely to significantly improve employee health or reduce medical utilization in the short term.

Article Information

Author Contributions

Drs. Reif, Jones, and Molitor had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Reif, Jones, Payne, and Molitor

Acquisition, analysis, or interpretation of data: Reif, Chan, Jones, and Molitor

Drafting of the manuscript: All authors

Critical revision of the manuscript for important intellectual content: All authors

Statistical analysis: Reif, Jones, and Molitor

Obtained funding: Reif, Jones, Payne, and Molitor

Administrative, technical, or material support: All authors

Supervision: Reif, Jones, Payne, and Molitor

Conflict of Interest Disclosures

No conflicts of interest were reported.

Funding/Support

This research was supported by the National Institute on Aging of the National Institutes of Health under award number R01AG050701; the National Science Foundation under Grant No. 1730546; the Abdul Latif Jameel Poverty Action Lab (J-PAL) North America US Health Care Delivery Initiative; Evidence for Action (E4A), a program of the Robert Wood Johnson Foundation; and the W.E. Upjohn Institute for Employment Research. Illinois Human Resources provided in-kind logistical support for developing the program.

Role of the Funder/Sponsor

The funders had no role in the design or conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Additional Contributions

We thank Lauren Geary for outstanding project management. We thank Michele Guerra for her contributions and many hours volunteered to the project, without financial compensation. We thank Illinois Human Resources for invaluable institutional support, without financial compensation. We are also thankful for comments from Marian Huhman, Mark Stehr, and seminar participants at the American Society of Health Economists, without financial compensation.

References

1. Claxton G, Rae M, Damico A, Young G, McDermott D, Whitmore H. The Kaiser Family Foundation Employer Health Benefits 2019 Annual Survey. 2019.
2. Jones D, Molitor D, Reif J. What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study. *Quarterly Journal of Economics.* 2019;134(4):1747-1791.
3. Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G. Financial incentive–based approaches for weight loss: a randomized trial. *JAMA.* 2008;300(22):2631-2637.
4. Charness G, Gneezy U. Incentives to exercise. *Econometrica.* 2009;77(3):909-931.
5. Haisley E, Volpp KG, Pellathy T, Loewenstein G. The impact of alternative incentive schemes on completion of health risk assessments. *American Journal of Health Promotion.* 2012;26(3):184-188.
6. Royer H, Stehr M, Sydnor J. Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics.* 2015;7(3):51-84.
7. Meenan RT, Vogt TM, Williams AE, Stevens VJ, Albright CL, Nigg C. Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii. *Journal of Occupational and Environmental Medicine.* 2010;52(Suppl 1):S8.
8. Handel B, Kolstad J. Wearable technologies and health behaviors: New data and new methods to understand population health. *American Economic Review.* 2017;107(5):481-485.
9. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Affairs.* 2010;29(2):304-311.
10. 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. *American Journal of Health Promotion.* 2014;28(6):347-363.
11. Song Z, Baicker K. Effect of a workplace wellness program on employee health and economic outcomes: a randomized clinical trial. *JAMA.* 2019;321(15):1491-1501.
12. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the health belief model. *Health Education Quarterly.* 1988;15(2):175-183.
13. Bandura A. Health promotion by social cognitive means. *Health Education & Behavior.* 2004;31(2):143-164.
14. Claxton G, Rae M, Long M, Damico A, Sawyer B. The Kaiser Family Foundation Employer Health Benefits 2016 Annual Survey. 2016.
15. Manski CF. Measuring expectations. *Econometrica.* 2004;72(5):1329-1376.
16. Hendren N. Private information and insurance rejections. *Econometrica.* 2013;81(5):1713-1762.
17. Ioannidis JP, Stanley TD, Doucouliagos H. The power of bias in economics research. Oxford University Press Oxford, UK; 2017.

18. Gerber A, Green D. *Field Experiments: Design, Analysis, and Interpretation.* W. W. Norton; 2012.
19. *Stata Statistical Software: Release 16* [computer program]. College Station, TX: StataCorp LLC; 2019.
20. Baicker K, Taubman SL, Allen HL, et al. The Oregon experiment—effects of Medicaid on clinical outcomes. *New England Journal of Medicine.* 2013;368(18):1713-1722.
21. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association.* 1996;91(434):444-455.
22. Finkelstein A, Taubman S, Wright B, et al. The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics.* 2012;127(3):1057-1106.
23. Westfall PH, Young SS. *Resampling-based multiple testing: Examples and methods for p-value adjustment.* Vol 279: John Wiley & Sons; 1993.
24. Abraham JM. Employer wellness programs—a work in progress. *JAMA.* 2019;321(15):1462-1463.

Tables and Figures

Table 1. Baseline Characteristics of the Study Population[a]

Variable	Treatment Group (n = 3300)	Control Group (n = 1534)
Age group, No. (%)		
<37 yr	1125 (34.1)	516 (33.6)
37-50 yr	1097 (33.2)	522 (34.0)
≥50 yr	1078 (32.7)	496 (32.3)
Age, mean (SD), y	43.8 (11.3)	44.0 (11.4)
Sex, No. (%)		
Male	1411 (42.8)	653 (42.6)
Female	1889 (57.2)	881 (57.4)
Race, No. (%)		
White	2758 (83.6)	1290 (84.1)
Non-white	542 (16.4)	244 (15.9)
Annual salary, $, No. (%)		
<40,000	798 (24.2)	374 (24.4)
40,000 to <50,000	660 (20.0)	327 (21.3)
50,000 to <75,000	1090 (33.0)	469 (30.6)
≥75,000	752 (22.8)	364 (23.7)
Employee class, No. (%)		
Faculty	662 (20.1)	301 (19.6)
Academic professional	1442 (43.7)	679 (44.3)
Civil service	1196 (36.2)	554 (36.1)
Health Alliance insurance, Oct 2015 to Jul 2016		
Any coverage, No. (%)	2184 (66.2)	1033 (67.3)
Months of coverage, mean (SD)	6.4 (4.7)	6.4 (4.7)
	Insurance Claims Subsample (n = 2184)	**Insurance Claims Subsample (n = 1033)**
Medical diagnoses, No. (%)		
Diabetes	106 (4.9)	66 (6.4)
Hypertension	289 (13.2)	151 (14.6)
Hyperlipidemia	337 (15.4)	171 (16.6)
Medical utilization, mean (SD), days		
Office/outpatient	2.4 (2.6)	2.7 (2.8)
Inpatient	0.1 (1.1)	0.1 (0.4)
ER	0.1 (0.5)	0.1 (0.4)

[a] Age, salary, and employee class are defined as of June 2016, two months prior to the start of the intervention. Medical diagnoses and medical utilization are measured over the period October 2015 to July 2016 and are derived from the insurance claims subsample, which includes all study participants enrolled in the Health Alliance plan.

Table 2. Mean Values and Effect of Wellness Program on Health Beliefs and Self-Reported Health Behaviors[a]

	Group Mean (SD)		Effect of Wellness Program Eligibility		
Outcome	Treatment	Control	Effect (95% CI)	*P* Value	Adjusted *P* Value[b]
Health beliefs, 2017					
Height, cm	170.5 (9.8)	170.9 (10.2)	-0.22 (-0.92 to 0.49)	.55	.96
Weight, kg	83.9 (21.7)	84.1 (21.7)	0.03 (-1.91 to 1.97)	.98	>0.99
Chance of BMI > 30, %	46.2 (40.1)	46.8 (39.5)	-0.87 (-4.53 to 2.80)	.64	.96
Chance of high cholesterol, %	37.1 (28.2)	40.2 (27.7)	-3.01 (-5.70 to -0.31)	.03	.24
Chance of high blood pressure, %	29.0 (26.0)	31.6 (27.1)	-2.41 (-4.96 to 0.14)	.06	.41
Chance of impaired glucose, %	28.3 (24.3)	31.0 (24.0)	-2.68 (-5.01 to -0.36)	.02	.22
Health beliefs, 2018					
Height, cm	170.4 (10.1)	170.8 (10.9)	-0.33 (-1.11 to 0.45)	.40	.95
Weight, kg	83.5 (21.4)	83.9 (22.1)	-0.61 (-2.70 to 1.47)	.56	.96
Chance of BMI > 30, %	46.2 (40.1)	46.1 (39.4)	-0.10 (-3.97 to 3.78)	.96	>0.99
Chance of high cholesterol, %	37.3 (27.6)	39.2 (28.5)	-1.74 (-4.57 to 1.09)	.23	.83
Chance of high blood pressure, %	29.4 (25.5)	32.4 (26.5)	-2.93 (-5.53 to -0.33)	.03	.24
Chance of impaired glucose, %	28.4 (25.3)	29.5 (25.2)	-1.00 (-3.59 to 1.58)	.45	.95
Self-reported health behaviors, 2017, %					
Has primary physician	89.4 (30.8)	85.9 (34.8)	3.20 (0.09 to 6.30)	.04	.42
No tobacco use	94.8 (22.3)	94.4 (23.0)	0.46 (-1.76 to 2.67)	.69	>0.99
Exercise 1+ times/week	92.6 (26.2)	93.4 (24.9)	-0.84 (-3.26 to 1.58)	.50	>0.99
Exercise 3+ times/week	57.6 (49.4)	53.1 (49.9)	4.73 (0.03 to 9.44)	.05	.43
Exercise for 20 minutes	93.4 (24.9)	94.2 (23.5)	-0.83 (-3.14 to 1.48)	.48	>0.99
Exercise for 40 minutes	49.2 (50.0)	50.1 (50.0)	-0.86 (-5.66 to 3.95)	.73	>0.99
Never anxious/depressed	32.1 (46.7)	31.6 (46.5)	0.18 (-4.28 to 4.65)	.94	>0.99
Never or sometimes anxious/depressed	86.9 (33.8)	87.1 (33.6)	-0.50 (-3.69 to 2.69)	.76	>0.99
Self-reported health behaviors, 2018, %					
Has primary physician	92.2 (26.9)	86.1 (34.6)	6.13 (3.04 to 9.22)	<.001	<0.01
No tobacco use	95.2 (21.5)	93.0 (25.6)	2.60 (0.16 to 5.04)	.04	.38
Exercise 1+ times/week	91.0 (28.7)	89.9 (30.1)	1.26 (-1.79 to 4.31)	.42	>0.99
Exercise 3+ times/week	52.3 (50.0)	47.8 (50.0)	4.42 (-0.55 to 9.39)	.08	.58
Exercise for 20 minutes	92.1 (27.0)	91.2 (28.3)	0.89 (-2.00 to 3.78)	.54	>0.99
Exercise for 40 minutes	46.8 (49.9)	46.5 (49.9)	0.50 (-4.53 to 5.54)	.84	>0.99
Never anxious/depressed	32.6 (46.9)	31.5 (46.5)	0.68 (-4.05 to 5.42)	.78	>0.99
Never or sometimes anxious/depressed	85.8 (35.0)	84.6 (36.1)	1.31 (-2.33 to 4.95)	.48	>0.99
Standardized treatment effect[c]					
Health beliefs			-0.07 (-0.12 to -0.01)	.02	
Self-reported health behaviors			0.04 (-0.00 to 0.08)	.05	

[a] This table reports effects of the wellness program. All regressions included stratification variables as controls. All outcome variables were obtained during the on-site screening in either 2017 (12-month follow-up) or 2018 (24-month follow-up). The sample size of the regressions ranged from 1,739 to 1,999 because fewer subjects participated in the 2018 screening than in the 2017 screening and because some outcomes were occasionally missing or illegible.

[b] Adjusted *P* values account for the number of hypotheses tested in each domain. We tested 12 hypotheses in the health beliefs domain and 16 hypotheses in the self-reported health behaviors domain.

[c] The standardized treatment effect gives equal weight to each outcome within a domain and includes both the 2017 and 2018 outcomes. The standardized treatment effect for health beliefs excludes height and weight.

Table 3. Mean Values and Effect of Wellness Program on Biometrics[a]

	Group Mean (SD)		Effect of Wellness Program Eligibility		
Outcome	Treatment	Control	Effect (95% CI)	*P* Value	Adjusted *P* Value[b]
Biometric outcomes, 2017					
Continuous measures					
Height, cm	170.3 (9.4)	170.3 (9.4)	0.24 (-0.41 to 0.89)	.48	>0.99
Weight, kg	84.1 (21.9)	84.3 (22.0)	-0.05 (-2.01 to 1.91)	.96	>0.99
Waist, cm	95.2 (16.8)	95.6 (17.1)	-0.37 (-1.89 to 1.16)	.64	>0.99
BMI	28.9 (7.0)	29.0 (7.1)	-0.12 (-0.76 to 0.53)	.72	>0.99
Blood pressure, mm Hg					
Systolic	123.8 (13.7)	124.9 (14.9)	-1.07 (-2.37 to 0.24)	.11	.89
Diastolic	75.5 (9.1)	75.8 (8.9)	-0.34 (-1.16 to 0.47)	.41	>0.99
Lipid panel					
Total cholesterol, mg/dL	187.4 (41.2)	185.9 (38.8)	1.69 (-2.04 to 5.41)	.37	>0.99
HDL cholesterol, mg/dL	54.4 (17.3)	54.8 (17.9)	-0.38 (-1.93 to 1.17)	.63	>0.99
Total cholesterol / HDL cholesterol	3.7 (1.2)	3.6 (1.1)	0.06 (-0.05 to 0.17)	.26	>0.99
LDL cholesterol, mg/dL	107.4 (34.6)	106.6 (33.1)	1.07 (-2.21 to 4.35)	.52	>0.99
Triglycerides, mg/dL	129.1 (70.1)	124.5 (61.3)	4.02 (-2.12 to 10.17)	.20	.98
Glucose, mg/dL	94.1 (20.5)	93.4 (21.0)	0.43 (-1.56 to 2.41)	.67	>0.99
Binary measures, %					
Obesity (BMI ≥30)	35.5 (47.9)	33.9 (47.4)	1.36 (-3.09 to 5.81)	.55	>0.99
Hypertension (systolic ≥130 or diastolic ≥80)	49.8 (50.0)	50.4 (50.0)	-0.63 (-5.23 to 3.98)	.79	>0.99
High LDL cholesterol (≥100 mg/dL)	56.4 (49.6)	57.7 (49.4)	-1.06 (-5.86 to 3.74)	.66	>0.99
High glucose (≥100 mg/dL)	25.1 (43.4)	22.5 (41.8)	2.54 (-1.44 to 6.52)	.21	.98
Biometric outcomes, 2018					
Continuous measures					
Height, cm	170.2 (9.5)	170.2 (9.6)	0.02 (-0.68 to 0.72)	.95	>0.99
Weight, kg	84.1 (21.7)	84.8 (22.2)	-0.86 (-2.97 to 1.24)	.42	>0.99
Waist, cm	95.2 (17.3)	95.3 (17.4)	-0.11 (-1.80 to 1.59)	.90	>0.99
BMI	29.0 (6.9)	29.3 (7.5)	-0.40 (-1.11 to 0.31)	.27	>0.99
Blood pressure, mm Hg					
Systolic	122.3 (13.8)	122.4 (14.0)	-0.15 (-1.48 to 1.18)	.83	>0.99
Diastolic	76.3 (9.8)	76.2 (9.8)	0.04 (-0.91 to 1.00)	.93	>0.99
Lipid panel					
Total cholesterol, mg/dL	178.8 (40.6)	178.3 (37.8)	0.70 (-3.19 to 4.59)	.72	>0.99
HDL cholesterol, mg/dL	54.1 (16.7)	54.0 (17.1)	-0.09 (-1.64 to 1.47)	.91	>0.99
Total cholesterol / HDL cholesterol	3.5 (1.1)	3.5 (1.2)	0.01 (-0.10 to 0.13)	.80	>0.99
LDL cholesterol, mg/dL	101.7 (35.1)	101.2 (33.7)	0.89 (-2.71 to 4.48)	.63	>0.99
Triglycerides, mg/dL	120.2 (65.0)	119.4 (62.6)	1.13 (-5.48 to 7.74)	.74	>0.99
Glucose, mg/dL	103.4 (18.7)	103.8 (17.6)	-0.52 (-2.32 to 1.27)	.57	>0.99
Binary measures, %					
Obesity (BMI ≥30)	36.5 (48.2)	36.3 (48.1)	-0.15 (-4.88 to 4.59)	.95	>0.99
Hypertension (systolic ≥130 or diastolic ≥80)	49.3 (50.0)	47.8 (50.0)	1.01 (-3.91 to 5.92)	.69	>0.99
High LDL cholesterol (≥100 mg/dL)	47.5 (50.0)	48.1 (50.0)	0.31 (-4.94 to 5.57)	.91	>0.99
High glucose (≥100 mg/dL)	55.3 (49.7)	52.8 (50.0)	2.83 (-2.01 to 7.68)	.25	>0.99
Standardized treatment effect[c]					
Biometric outcomes			-0.00 (-0.05 to 0.04)	.83	

[a] This table reports effects of the wellness program. All regressions included stratification variables as controls. All outcome variables were obtained during the on-site screening in either 2017 (12-month follow-up) or 2018 (24-month follow-up). The sample size of the regressions ranged from 1,662 to 2,004 because fewer subjects participated in the 2018 screening than in the 2017 screening and because some outcomes were occasionally missing or illegible.

[b] Adjusted *P* values account for the 32 hypotheses tested in this domain.

[c] The standardized treatment effect gives equal weight to each outcome within a domain and includes both the 2017 and 2018 outcomes. It excludes height, BMI, total cholesterol / HDL cholesterol, LDL cholesterol, and the four binary measures.

Table 4. Mean Values and Effect of Wellness Program on Medical Diagnoses and Utilization[a]

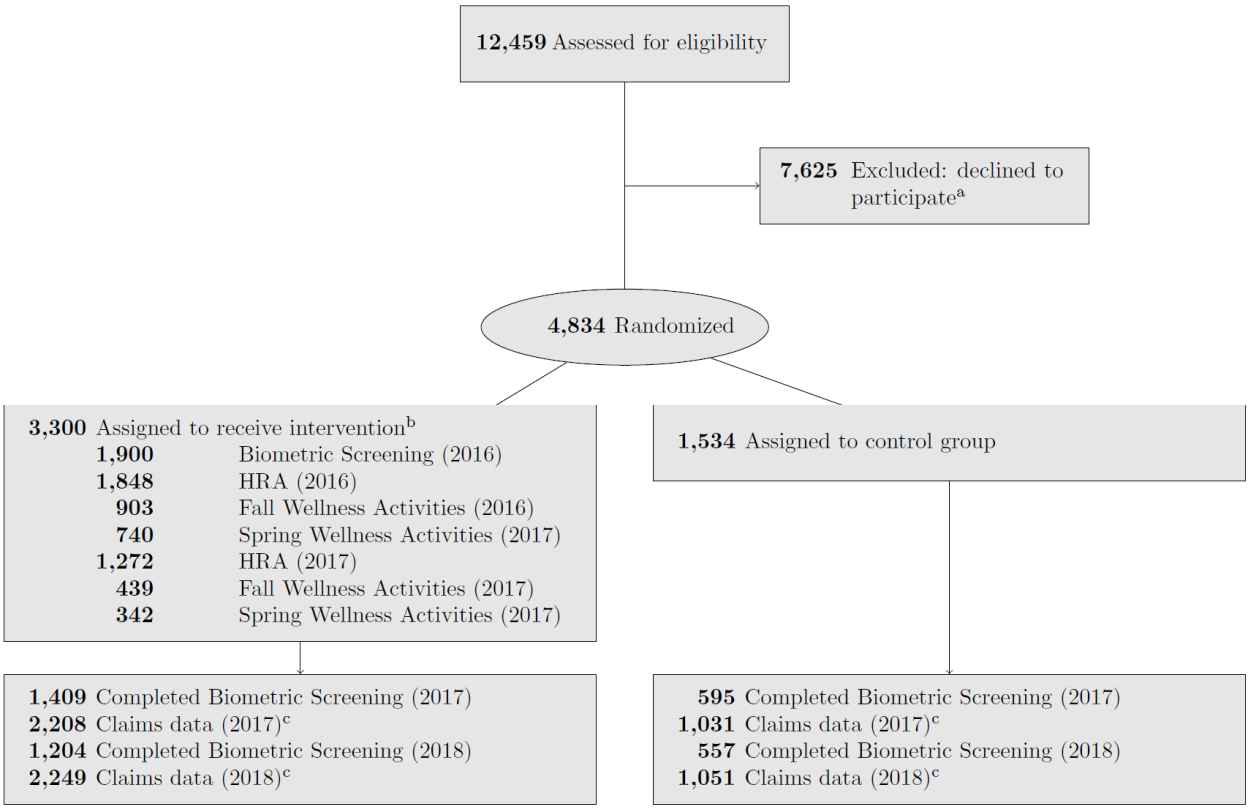
	Group Mean (SD)		Effect of Wellness Program Eligibility		
Outcome	Treatment	Control	Effect (95% CI)	*P* Value	Adjusted *P* Value[b]
Medical diagnoses, 2017, %					
Diabetes	5.6 (23.0)	6.8 (25.2)	0.26 (-0.61 to 1.12)	.56	.93
Hypertension	15.3 (36.0)	18.1 (38.5)	-1.57 (-3.70 to 0.56)	.15	.50
Hyperlipidemia	18.7 (39.0)	19.5 (39.6)	0.40 (-2.13 to 2.93)	.76	.98
Medical diagnoses, 2018, %					
Diabetes	6.3 (24.3)	7.8 (26.9)	-0.09 (-1.14 to 0.96)	.86	.98
Hypertension	19.6 (39.7)	22.5 (41.8)	-1.55 (-3.87 to 0.77)	.19	.57
Hyperlipidemia	25.5 (43.6)	26.5 (44.2)	0.30 (-2.47 to 3.07)	.83	.98
Medical utilization, 2017					
Office/outpatient (# days with at least 1 claim)	3.20 (3.28)	3.31 (3.44)	0.05 (-0.16 to 0.26)	.64	.96
Inpatient (# days with at least 1 claim)	0.09 (0.68)	0.08 (0.59)	0.02 (-0.03 to 0.06)	.52	.96
ER (# days with at least 1 claim)	0.13 (0.47)	0.15 (0.53)	-0.02 (-0.06 to 0.02)	.34	.87
Medical utilization, 2018					
Office/outpatient (# days with at least 1 claim)	6.46 (6.16)	6.67 (6.54)	0.08 (-0.30 to 0.46)	.68	.95
Inpatient (# days with at least 1 claim)	0.20 (1.41)	0.23 (2.59)	-0.03 (-0.19 to 0.14)	.77	.96
ER (# days with at least 1 claim)	0.26 (0.79)	0.28 (1.13)	-0.02 (-0.10 to 0.05)	.56	.96
Standardized treatment effect[c]					
Medical diagnoses			-0.01 (-0.04 to 0.02)	.59	
Medical utilization			-0.00 (-0.05 to 0.04)	.92	

[a] This table reports effects of the wellness program. All regressions included stratification variables, baseline medical diagnoses, and baseline medical utilization as controls. All regressions were weighted by the employee's number of months of insurance coverage in the post-intervention period. The 2017 period is defined as August 2016 to July 2017, and the 2018 period is defined as August 2016 to July 2018. The sample size of the regressions ranged from 3,164 to 3,167.

[b] Adjusted *P* values account for the number of hypotheses tested in each domain. We tested 6 hypotheses in the medical diagnoses domain and 6 hypotheses in the medical utilization domain.

[c] The standardized treatment effect gives equal weight to each outcome within a domain and includes both the 2017 and 2018 outcomes.

Figure. Flow of Participants in the Illinois Workplace Wellness Study



[a] All eligible employees were invited to enter the study by taking a baseline survey. Those who did not complete the survey were not included in the study.

[b] Participants who received the intervention were invited to participate in wellness program components over the course of two years. Participation varied across the various components.

[c] Claims data were collected for those participants in the treatment and control group who were enrolled in the Health Alliance insurance plan.