What Do Workplace Wellness Programs Do?
Evidence from the Illinois Workplace Wellness Study

[Website Link]

Damon Jones
University of Chicago
and NBER

David Molitor
University of Illinois
and NBER

Julian Reif
University of Illinois
and NBER

NBER Summer Institute
July 24, 2019
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; Evidence for Action (E4A), a program of the Robert Wood Johnson Foundation; the Abdul Latif Jameel Poverty Action Lab (J-PAL) North America U.S. Health Care Delivery Initiative; and the W.E. Upjohn Institute for Employment Research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, any of our funders, or the University of Illinois.
Workplace wellness programs are growing rapidly

- Workplace wellness has become increasingly popular among US employers
  - Aims: reduce health care costs, improve employee health, increase productivity
  - 83% of large firms offer wellness programs, covering >50 million workers
Workplace wellness programs are growing rapidly

- Workplace wellness has become increasingly popular among US employers
  - Aims: reduce health care costs, improve employee health, increase productivity
  - 83% of large firms offer wellness programs, covering >50 million workers

- Workplace wellness is popular among policymakers
  - Affordable Care Act “Safeway Amendment” encourages these programs
  - Some advocate expanding to Medicare, Medicaid
Workplace wellness programs are growing rapidly

- Workplace wellness has become increasingly popular among US employers
  - Aims: reduce health care costs, improve employee health, increase productivity
  - 83% of large firms offer wellness programs, covering >50 million workers

- Workplace wellness is popular among policymakers
  - Affordable Care Act “Safeway Amendment” encourages these programs
  - Some advocate expanding to Medicare, Medicaid

- Workplace wellness programs are controversial
  - Do they actually improve health and productivity?
  - Do financial incentives shift costs onto certain groups of employees?
Prior evidence is mixed and limited

- Mixed results on effects of wellness programs
  - Meta analysis by Baicker, Cutler, and Song (2010) found
    - Medical cost reduction of $3.27 for every $1 spent on wellness
    - Absenteeism cost reduction of $2.73 for every $1 spent on wellness
  - Cost reduction estimates decreasing in study quality (Gowrisankaran et al. 2013)
  - Recent 18-month RCT finds no effect on medical spending or behaviors (Song and Baicker 2019)
Prior evidence is mixed and limited

- Mixed results on effects of wellness programs
  - Meta analysis by Baicker, Cutler, and Song (2010) found
    - Medical cost reduction of $3.27 for every $1 spent on wellness
    - Absenteeism cost reduction of $2.73 for every $1 spent on wellness
  - Cost reduction estimates decreasing in study quality (Gowrisankaran et al. 2013)
  - Recent 18-month RCT finds no effect on medical spending or behaviors (Song and Baicker 2019)

- Empirical challenges:
  - Selection bias from non-random participation
  - Potential for publication bias
  - Measurement and power
The Illinois Workplace Wellness Study

- Two-year randomized controlled trial at University of Illinois
  - 12,459 employees invited: faculty (26%), administrative and union service positions (74%)
  - Individual, random assignment to control or treatment groups
The Illinois Workplace Wellness Study

- Two-year randomized controlled trial at University of Illinois
  - 12,459 employees invited: faculty (26%), administrative and union service positions (74%)
  - *Individual, random assignment* to control or treatment groups

- Rich data linked at individual level allows for comprehensive evaluation
  - Administrative data on employment and health insurance claims
  - Administrative data on health behaviors (gym use, running events)
  - Survey data
  - Detailed biometric data

- Study was pre-registered in the AEA RCT Registry
Today: Two research questions

1. Who participates in workplace wellness?
   - Distributional consequences

2. What are the effects of workplace wellness after thirty months?
   - Medical spending
   - Employee productivity
   - Health behaviors
Summary of main findings

1. Significant advantageous selection on health
   - Workplace wellness programs might act as profitable screening mechanisms

2. No significant causal effects on medical spending or employee productivity
   - We rule out the widely cited return on investment for wellness programs (Baicker, Cutler, and Song 2010)
Study Design:
Illinois Workplace Wellness Study
Background on workplace wellness programs

- Three main components:
  1. Biometric health screening
  2. Health risk assessment (HRA)
  3. Wellness activities

- We designed a “gold-standard” wellness program (iThrive)
  - Includes all three components above
  - Includes financial incentives tied to participation
  - Allows employees to take paid time off to participate
Study enrollment (July 2016)

- Enrollment predicated on completing a 15-minute online survey

- Survey invitations sent to 12,459 employees
  - Postcard notification sent to employee home address (July 6, 2016)
  - Email invitation to employees, with personalized link to online survey (July 11)

- Employees were offered $30 Amazon.com gift card, plus a chance “to participate in a second part of the research study"
You have been selected to take an online survey as part of the Illinois Workplace Wellness Study.

The purpose of this survey is to better understand health behaviors and wellness on campus.

Check your University of Illinois email on July 11th for instructions and a link to the survey.

All respondents will receive a $30 Amazon.com Gift Card for completing the survey.

For more information: WellnessStudy@illinois.edu
9. Are you currently trying to increase your physical activity or exercise?
   - [ ] Yes
   - [ ] No
Online survey interface: nudges to answer questions

Illinois Workplace Wellness Study

You left an answer blank. If this was intentional, please click “Next” and proceed with the rest of the survey. If not, please answer the question.

9. Are you currently trying to increase your physical activity or exercise?
   ○ Yes
   ○ No

Back  Next
Survey responses were very complete

- **Response:** 4,834 employees (38.8%) successfully completed baseline survey
  - Survey open for three weeks, with periodic email reminders

- **Completeness:** Fewer than 1% of respondents skipped *any* question

- **Validity:** Age, sex responses match University administrative records closely
  - Last questions on the survey
  - 99.4% within 1 year of age reported in University records
Experimental design: Illinois Workplace Wellness Study

Study Sample (completed baseline survey, $N = 4,834$)

- **Control Group** ($N = 1,534$)
- **Treatment Group** ($N = 3,300$)

Randomization

- **Year 1 Follow-up**
- **iThrive Wellness Program**
- **Year 1 Follow-up**
- **Year 2 Follow-up**
- **iThrive Wellness Program**
- **Year 2 Follow-up**
Participating in iThrive involved two steps

Step 1: Screening

*Biometric Screening*
- 8 different on- and off-campus locations
- Fingerstick + blood pressure, height, weight, waist circumference

*Health Risk Assessment (HRA)*
- Online questionnaire designed to assess lifestyle habits
Participating in iThrive involved two steps

Step 1: Screening

*Biometric Screening*
- 8 different on- and off-campus locations
- Fingerstick + blood pressure, height, weight, waist circumference

*Health Risk Assessment (HRA)*
- Online questionnaire designed to assess lifestyle habits

Step 2: Wellness Activities

- Many options, e.g., Weight Watchers, smoking cessation, stress management
- Classes ranged from 6-12 weeks
- Completion defined as attending at least 3/4 of sessions
Participating in iThrive involved two steps

**Step 1: Screening — up to $200 reward**

*Biometric Screening*
- 8 different on- and off-campus locations
- Fingerstick + blood pressure, height, weight, waist circumference

*Health Risk Assessment (HRA)*
- Online questionnaire designed to assess lifestyle habits

**Step 2: Wellness Activities — up to $75/semester**

- Many options, e.g., Weight Watchers, smoking cessation, stress management
- Classes ranged from 6-12 weeks
- Completion defined as attending at least 3/4 of sessions
My Portal

My Portal gives you information about your progress in iThrive, a program to promote health and wellness among campus faculty and staff. iThrive offers you the opportunity to participate in valuable health screening and wellness activities at no cost to you. In addition, you can receive financial rewards for completing certain elements of iThrive.

To earn rewards and to participate in Wellness Activities, you must complete your Health Screening by Friday, September 16th and the Health Assessment by Friday, September 30.

Your participation reward: $200.00 of $350.00 earned so far

**Step 1: Health Screening & Assessment**

The first step in iThrive is to complete your Health Screening and Health Assessment. After you complete your Health Screening, you will be able to access your online Health Assessment. [Learn more about Health Screening & Assessment »](#)

Congratulations! You have completed your Health Screening and Health Assessment.

**Reward for completing both the Health Screening and Health Assessment:** $200.00

- ✔ Health Screening completed
- ✔ Health Assessment completed

**Step 2: Wellness Activities**

After you have completed Step 1, you may register to participate in a wellness activity. You may use the information provided to you in your Health Assessment to select a program that best addresses an area of your health that you would like to improve. [Learn more about Wellness Activities »](#)

Registration for Fall Activities is now closed. More information about Spring Activity registration will be made available soon.

**Reward for completing Fall activity:** $75.00

**Reward for completing Spring activity:** $75.00

- ✗ Fall activity not completed. Registered for HealthTrails
- ✗ Spring activity not completed
Data and Results:
Illinois Workplace Wellness Study
We construct 42 outcomes from our datasets

- Medical spending and utilization (8 outcomes)
- Health status and behaviors (17 outcomes)
- Employment and productivity (17 outcomes)
17 measures of productivity from survey and admin data

- **Administrative measures (5 outcomes)**
  - Job title change, annual salary \(\rightarrow\) “job promotion”
  - Job retention
  - Sick leave taken

- **Online survey measures (11 outcomes)**
  - Worked 50+ hours/week
  - Job satisfaction, job search
  - Feel productive at work, feel happy at work, received promotion

- **First principle component of all available measures (1 outcome)**
  - “Productivity index”
### Baseline summary statistics: administrative data

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treat</th>
<th>( p )-value</th>
<th>( N )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.426</td>
<td>0.428</td>
<td>0.90</td>
<td>4,834</td>
</tr>
<tr>
<td>Age 50+</td>
<td>0.323</td>
<td>0.327</td>
<td>0.82</td>
<td>4,834</td>
</tr>
<tr>
<td>White</td>
<td>0.841</td>
<td>0.836</td>
<td>0.65</td>
<td>4,834</td>
</tr>
<tr>
<td>Faculty</td>
<td>0.196</td>
<td>0.201</td>
<td>0.72</td>
<td>4,834</td>
</tr>
<tr>
<td><strong>Health claims</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total spending (dollars/month)</td>
<td>506</td>
<td>465</td>
<td>0.32</td>
<td>3,222</td>
</tr>
<tr>
<td><strong>Health behaviors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Running event participant</td>
<td>0.107</td>
<td>0.118</td>
<td>0.13</td>
<td>4,834</td>
</tr>
<tr>
<td>Gym visits (days/year)</td>
<td>7.36</td>
<td>6.78</td>
<td>0.48</td>
<td>4,834</td>
</tr>
<tr>
<td><strong>Productivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sick leave (days/year)</td>
<td>6.05</td>
<td>6.13</td>
<td>0.71</td>
<td>4,834</td>
</tr>
<tr>
<td>Annual salary (dollars)</td>
<td>61,528</td>
<td>61,736</td>
<td>0.84</td>
<td>4,770</td>
</tr>
<tr>
<td><strong>Sample size</strong></td>
<td>1,534</td>
<td>3,300</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Treatment group participation in our workplace wellness program

Year 1:
- Screening: 58%
- HRA: 56%
- Fall Activity: 27%
- Spring Activity: 22%

Year 2:
- Screening: 43%
- HRA: 39%
- Fall Activity: 13%
- Spring Activity: 10%
We present two sets of results

1. Selection into workplace wellness

2. Causal effects of workplace wellness
   - Short-run (12 months)
   - Longer-run (24-30 months) [new!]
Results 1: Who participates in workplace wellness programs?

\[ X_i = \alpha + \theta x P_i + \varepsilon_i \]

- Observations: employees assigned to the treatment group
- \( X_i \) selection variable, “pre-determined” prior to intervention
- \( P_i \) indicator for completing both health screening and HRA in the first year

\textit{WYOUNG:} Stata module for performing multiple hypothesis testing

\texttt{ssc install wyoung}
Results 1: Who participates in workplace wellness programs?

\[ X_i = \alpha + \theta x P_i + \varepsilon_i \]

- Observations: employees assigned to the treatment group
- \( X_i \) selection variable, “pre-determined” prior to intervention
- \( P_i \) indicator for completing both health screening and HRA in the first year
- Family-wise \( p \)-values adjust for multiple outcomes (Westfall and Young 1993)
Results 1: Who participates in workplace wellness programs?

\[ X_i = \alpha + \theta x P_i + \varepsilon_i \]

- Observations: employees assigned to the treatment group
- \( X_i \) selection variable, “pre-determined” prior to intervention
- \( P_i \) indicator for completing both health screening and HRA in the first year
- Family-wise \( p \)-values adjust for multiple outcomes (Westfall and Young 1993)

**WYOUNG**: Stata module for performing multiple hypothesis testing

Install by typing “ssc install wyoung” at Stata prompt
Participants already had lower health spending

Figure: Pre-intervention, average monthly medical spending

Vertical bars display 95% confidence intervals on the difference in means.
Participants already had healthier behaviors

**Figure: Average annual gym visits**

- Did Not Complete Screening (N=1,452): 5.6 visits/year
- Completed Screening (N=1,848): 7.7 visits/year

**Figure: Running event participation**

- Did Not Complete Screening (N=1,452): 6.8%
- Completed Screening (N=1,848): 15.7%

Vertical bars display 95% confidence intervals on the difference in means.
Participants were spending less time at work

**Figure: Worked 50+ hours/week (survey)**

<table>
<thead>
<tr>
<th>Category</th>
<th>Did Not Complete Screening (N=1,452)</th>
<th>Completed Screening (N=1,845)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worked 50+ hours/week</td>
<td>20.5%</td>
<td>14.7%</td>
</tr>
</tbody>
</table>

**Figure: Sick leave taken**

<table>
<thead>
<tr>
<th>Category</th>
<th>Did Not Complete Screening (N=1,452)</th>
<th>Completed Screening (N=1,848)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sick leave taken</td>
<td>5.8 days/year</td>
<td>6.4 days/year</td>
</tr>
</tbody>
</table>

Vertical bars display 95% confidence intervals on the difference in means.
Results 2: Causal effects of workplace wellness

\[ Y_i = \alpha + \beta T_i + \gamma X_i + \varepsilon_i \]

- Observations: employees assigned to treatment or control group
- \( Y_i \) outcome variable
  - We consider 42 different survey and administrative data outcomes
- \( T_i \) indicator for treatment assignment
- \( X_i \) controls (none or post-Lasso)
Intervention reduced number of employees with no health screenings

**Figure:** Never had a screening (12-month survey)

<table>
<thead>
<tr>
<th></th>
<th>Control (N=1,157)</th>
<th>Treatment (N=2,410)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>8.5%</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

**Figure:** Never had a screening (24-month survey)

<table>
<thead>
<tr>
<th></th>
<th>Control (N=991)</th>
<th>Treatment (N=2,029)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Never</td>
<td>5.8%</td>
<td>2.9%</td>
</tr>
</tbody>
</table>

Vertical bars display 95% confidence intervals on the difference in means.
No causal effect on medical spending after 30 months

Pearson's chi-squared test for equality: p-value = 0.176
Kolmogorov-Smirnov test for equality: p-value = 0.373
No causal effect on health behaviors after 30 months

**Figure: Average annual gym visits**
- Control (N=1,534): 5.3 visits/year
- Treatment (N=3,300): 4.9 visits/year

**Figure: Running event participation**
- Control (N=1,534): 4.8%
- Treatment (N=3,300): 5.3%

Vertical bars display 95% confidence intervals on the difference in means.
No causal effect on employee productivity after 30 months

**Figure: Received job promotion**

- Control (N=1,158): 36%
- Treatment (N=2,477): 36%

**Figure: Sick leave**

- Control (N=1,534): 5.8 days/year
- Treatment (N=3,300): 5.9 days/year

Vertical bars display 95% confidence intervals on the difference in means.
We rule out 84% of prior studies on spending, absenteeism.

Orange shaded star is our RCT estimate. Red brackets depict 95% confidence intervals.
Why do our results differ from prior studies?

1. Treatment effects in our program/setting may differ (effect heterogeneity)

2. Prior studies may suffer from selection bias (most are non-RCT)
Why do our results differ from prior studies?

1. Treatment effects in our program/setting may differ (effect heterogeneity).

2. Prior studies may suffer from selection bias (most are non-RCT).

What happens if we estimate an observational model using our data?
Observational model estimates a reduction in medical spending

$-104/month$

$52/month$

$-150$

$-100$

$-50$

$0$

$50$

$100$

$150$

OLS

IV (RCT)

OLS and IV controls selected using post-Lasso. Vertical bars display 95% confidence interval for IV estimate.
Observational model estimates an improvement in health behaviors

**Figure: Average annual gym visits**

- **OLS**: 2.16 visits/year
- **IV (RCT)**: 0.76 visits/year

**Figure: Running event participation**

- **OLS**: 2.4%
- **IV (RCT)**: -1.1%

OLS and IV controls selected using post-Lasso. Vertical bars display 95% confidence interval for IV estimate.
Observational model estimates a reduction in job exit after 12 months

OLS: -6.3%
IV (RCT): -2.3%

OLS and IV controls selected using post-Lasso. Vertical bars display 95% confidence interval for IV estimate.
What have we learned from the Illinois Workplace Wellness Study?

1. Participants were already healthier and had lower medical spending
   - Wellness programs shift costs onto less healthy employees
   - Wellness programs may be effective way to attract healthy workers

2. No effects on medical spending, health behaviors, or productivity after 30 months
   - We rule out majority of estimates from prior studies

3. Observational health studies likely to suffer selection bias, even with rich controls (Oster 2019)
Illinois Workplace Wellness Study

www.nber.org/workplacewellness
Appendix Slides:
Illinois Workplace Wellness Study
## Associations between 2017 survey and admin measures of productivity

<table>
<thead>
<tr>
<th></th>
<th>(1) Salary (% change)</th>
<th>(2) Job promotion</th>
<th>(3) Job title change</th>
<th>(4) Job terminated</th>
<th>(5) Sick leave (days/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any sick days in past year [survey]</td>
<td>-0.009* (0.005)</td>
<td>0.005 (0.014)</td>
<td>0.009 (0.014)</td>
<td>-0.026*** (0.007)</td>
<td>3.242*** (0.240)</td>
</tr>
<tr>
<td>Worked 50+ hours/week [survey]</td>
<td>0.006 (0.007)</td>
<td>-0.032* (0.018)</td>
<td>-0.036** (0.018)</td>
<td>0.011 (0.010)</td>
<td>-3.278*** (0.295)</td>
</tr>
<tr>
<td>Very/somewhat satisfied with job [survey]</td>
<td>0.026*** (0.005)</td>
<td>0.050*** (0.017)</td>
<td>0.043** (0.017)</td>
<td>-0.026** (0.011)</td>
<td>-1.440*** (0.334)</td>
</tr>
<tr>
<td>Received promotion [survey]</td>
<td>0.050*** (0.005)</td>
<td>0.229*** (0.013)</td>
<td>0.225*** (0.013)</td>
<td>-0.013* (0.007)</td>
<td>0.007 (0.246)</td>
</tr>
<tr>
<td>Job search very likely [survey]</td>
<td>0.003 (0.007)</td>
<td>-0.049** (0.019)</td>
<td>-0.046** (0.020)</td>
<td>0.166*** (0.018)</td>
<td>-1.522*** (0.347)</td>
</tr>
<tr>
<td><strong>Outcome mean</strong></td>
<td>0.061</td>
<td>0.184</td>
<td>0.192</td>
<td>0.045</td>
<td>6.473</td>
</tr>
</tbody>
</table>

Notes: Each row and column reports estimates from a separate regression, with dependent variable given by column header.
Baseline summary statistics: online survey data

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treat</th>
<th>(p)-value</th>
<th>(N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ever screened</td>
<td>0.885</td>
<td>0.892</td>
<td>0.503</td>
<td>4,834</td>
</tr>
<tr>
<td>Physically active</td>
<td>0.359</td>
<td>0.382</td>
<td>0.134</td>
<td>4,834</td>
</tr>
<tr>
<td>Trying to be active</td>
<td>0.822</td>
<td>0.809</td>
<td>0.278</td>
<td>4,834</td>
</tr>
<tr>
<td>Current smoker</td>
<td>0.072</td>
<td>0.065</td>
<td>0.428</td>
<td>4,833</td>
</tr>
<tr>
<td>Current smoker (other)</td>
<td>0.085</td>
<td>0.085</td>
<td>0.960</td>
<td>4,833</td>
</tr>
<tr>
<td>Former smoker</td>
<td>0.198</td>
<td>0.196</td>
<td>0.870</td>
<td>4,833</td>
</tr>
<tr>
<td>Drinker</td>
<td>0.657</td>
<td>0.645</td>
<td>0.423</td>
<td>4,830</td>
</tr>
<tr>
<td>Heavy drinker</td>
<td>0.050</td>
<td>0.049</td>
<td>0.955</td>
<td>4,829</td>
</tr>
<tr>
<td>Chronic condition</td>
<td>0.729</td>
<td>0.726</td>
<td>0.824</td>
<td>4,834</td>
</tr>
<tr>
<td>Excellent/v. good health</td>
<td>0.586</td>
<td>0.602</td>
<td>0.281</td>
<td>4,834</td>
</tr>
<tr>
<td>Not poor health</td>
<td>0.989</td>
<td>0.989</td>
<td>0.882</td>
<td>4,834</td>
</tr>
<tr>
<td>Physical problems</td>
<td>0.392</td>
<td>0.388</td>
<td>0.793</td>
<td>4,834</td>
</tr>
<tr>
<td>Lots of energy</td>
<td>0.310</td>
<td>0.330</td>
<td>0.171</td>
<td>4,834</td>
</tr>
<tr>
<td>Bad emotional health</td>
<td>0.308</td>
<td>0.288</td>
<td>0.162</td>
<td>4,834</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.545</td>
<td>0.533</td>
<td>0.438</td>
<td>4,834</td>
</tr>
<tr>
<td>High BP/cholesterol/glucose</td>
<td>0.308</td>
<td>0.295</td>
<td>0.354</td>
<td>4,834</td>
</tr>
<tr>
<td>Sedentary</td>
<td>0.545</td>
<td>0.542</td>
<td>0.846</td>
<td>4,833</td>
</tr>
<tr>
<td>Pharmaceutical drug utilization</td>
<td>0.723</td>
<td>0.706</td>
<td>0.205</td>
<td>4,830</td>
</tr>
<tr>
<td>Physician/ER utilization</td>
<td>0.772</td>
<td>0.748</td>
<td>0.077</td>
<td>4,833</td>
</tr>
<tr>
<td>Hospital utilization</td>
<td>0.038</td>
<td>0.027</td>
<td>0.054</td>
<td>4,833</td>
</tr>
</tbody>
</table>

Sample size

<table>
<thead>
<tr>
<th>Control</th>
<th>Treat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,534</td>
<td>3,300</td>
</tr>
</tbody>
</table>