HERO POSITION PAPER ON VALUE OF WELLNESS INITIATIVES HOW TO ASSESS THE EVIDENCE ON "WHAT WORKS" – A CASE STUDY EXAMPLE

Despite a large and growing evidence base demonstrating positive outcomes of best-practice health and well-being programs (HWB) and providing guidance on their design, implementation and evaluation,[1] media coverage of occasional studies with negative findings sometimes creates confusion among employers about what really works in wellness. This commentary is intended to help HERO members better assess such findings by providing tips on how to critically examine research on program effectiveness. To illustrate how these tips can be used in practice, we apply them here to the recently published Illinois Workplace Wellness Study (Illinois Wellness study).[2]

Be skeptical of claims from a single study that appears to refute a large body of previously conducted peer-reviewed research.

Media headlines often imply that a single study proves something is or is not true, so our first tip is to be very skeptical of claims from one study that appear to refute a large body of well-designed research. In the three decades since the workplace health promotion field emerged, hundreds of studies published in peer-reviewed scientific journals have identified organizational policies, interventions, and cultural supports that yield positive health and business outcomes. Systematic literature reviews have critically evaluated this research, finding strong evidence that well-designed initiatives are effective. In this context, any new study must be weighed against these hundreds of previous studies and subjected to scientific scrutiny. For example, if a single clinical trial on a promising new drug failed to support its effectiveness, researchers would be cautious about rushing to judgment. Perhaps the drug was not administered in the proper dose; perhaps patients were not compliant with taking it as directed. We must be similarly cautious about jumping to conclusions when a single wellness study fails to report positive results. A single study addresses the merits of one intervention, which may or may not have incorporated evidence-based, best practice approaches.

To critically assess the strengths and limitations of any study, it's important to look beyond the headline in drawing any conclusions. This may be an intimidating exercise for those not versed in the complexities of research, but there are several areas any health promotion professional can examine without advanced research training. In the case of the Illinois Workplace Wellness Study, the title of the working paper suggests their findings answer the question, "What do workplace wellness programs do?" Based on the working paper, it would more aptly have been titled, "What did the I**llinois Workplace Wellness** program do i**n its first year**?" Generalizing to other or more mature programs, employee populations or work settings is not appropriate. When a study includes use of financial incentives and discusses potentially discriminatory uses of data collected via a wellness program, readers need to be all the more cautious in drawing conclusions about whether the program design is consistent with their organizational values and alignment with legal and ethical standards. HERO convened experts from throughout the country to examine the role of financial incentives in wellness programs and concluded that more research (such as the Illinois Wellness Study) "is needed to demonstrate whether financial incentives have a differential impact on lower income employees and/or other variables such as generational, ethnic, or geographic differences." We are fundamentally opposed any use of incentives that could lead to privacy concerns or discriminatory practices.[3]

"What do workplace wellness programs do?"

"What did the Illinois Workplace Wellness program do in its first year?"

Determine whether the tested intervention represents a best practice approach.

Best-practice, evidence-based approaches to workplace HWB initiatives are characterized by a comprehensive set of strategies offering all employees—and ideally their family members—a menu of programs, services, and resources that engage them in their own HWB and increase their awareness of the influence of their behaviors on others. These strategies must rely on evidence-based behavior change principles. Best practice approaches are also supported by organizational policies, workplace environment, comprehensive communications, leadership practices, and social norms that make healthy choices easier. No single element or wellness program can effectively meet the variety of HWB needs of all employees in an organization. Multiple strategies are needed and, ideally, each is supported by an established body of evidence demonstrating its effectiveness.

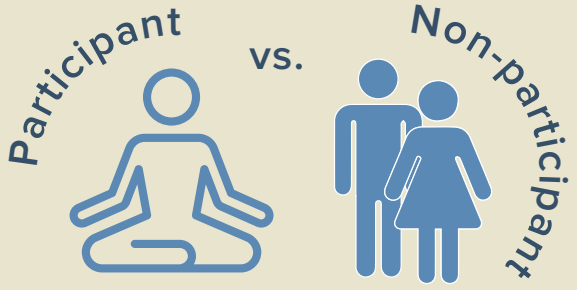
Best practice initiatives are also supported by strategic planning and ongoing evaluation to identify opportunities to improve upon the programs or their implementation. Some studies that fuel media claims that wellness does not work are based on programs that are poorly designed, poorly implemented, not evidence-based, or are incorporated into unsupportive workplace environments. Every peer-reviewed study includes a section describing the intervention design. Critically assess whether the intervention aligns with what you know about best-practice approaches. If you are not familiar with best practices, numerous scorecards[4-9] and well-conducted systematic literature reviews[10-12] identify elements of best-practice initiatives. These best practices elements cannot be assembled into a "one size fits all" approach to wellness program design. The basics of program planning start with an assessment of the problem one is solving specific to an organization and this should include the involvement of stakeholders the program is intended to serve, particularly those who are not likely to be drawn to wellness programs. One study showed, for example, significant variation in how best practices produced differential outcomes based on the demographic differences in the populations studied.[13]

The Illinois Wellness study describes their wellness program as offering on-site biometric health screening, online health risk assessment, and wellness classes largely offered in a classroom setting to university employees at the Urbana-Champaign campus. While employees randomized to be offered these activities were encouraged to participate during paid work time, there is no other mention in the paper of other critical best practice components such as health-supporting policies, leadership support, and comprehensive communications strategies.

Ask whether enough time was allowed for the intervention to produce desired outcomes.

One of the most common mistakes we've observed in evaluations of worksite HWB initiatives is an insufficient evaluation period following the launch of interventions. Some studies include comprehensive interventions but do not allow sufficient time for study participants to complete programs and put newly learned behaviors into practice before measuring results. Every comprehensive initiative should assess leading and lagging indicators of program success.[14-17] For example, the Illinois Wellness study provides insights on who was attracted to participate in certain aspects of their first-year program, which represents a leading indicator of program performance. Best-practice process evaluation uses such information to improve the program to ensure lagging measures of performance are likely to be met as the program matures over time. Instead, the Illinois Wellness study researchers conducted process, impact and outcome evaluation simultaneously with less than a year of intervention exposure for most program participants. Most new programs suffer deficiencies that need to be identified and corrected before focusing on outcomes several years into the program. Furthermore, behavior change after exposure to skill building programs often takes time. Past evidence indicates that some of the outcome measures the researchers examined would not be expected to change in the short one-year evaluation period.[17]

Do not assume all research uses the most appropriate evaluation approach. Random assignment of study individuals into treatment and control groups has long been considered the ideal scientific approach for evaluating interventions when feasible because it provides the strongest evidence of cause and effect. Called a randomized controlled trial (RCT), this study design is commonly used to assess the effectiveness of pharmaceutical drugs and some other medical treatments. However, it is typically impractical to use in workplace health promotion evaluation because most employers are reluctant and may even find it unethical or possibly illegal to withhold valuable programs and information from specific groups of employees.[15] Further, this approach cannot be used to test the effectiveness of a comprehensive HWB initiative at a single site/location because both the treatment and control groups are exposed to critical social, environmental and cultural supports. The Illinois Wellness study was only able to use a RCT design because the narrow scope of the intervention excluded supportive policies, work environment, leadership practices and other cultural elements. Narrow programmatic approaches without these supports are not considered best practice because they are not likely to result in sustainable behavior change.



Because randomization of individual employees is rarely realistic for comprehensive workplace HWB initiatives, many published studies compare volunteer participants to non-participants and statistically control for group differences related to study outcomes. When possible, because of the remaining potential for selection bias in participant versus non-participant designs, the most rigorous studies aim to compare employees at a single work location where the comprehensive HWB initiative is implemented to those at a similar work location where it is not implemented. Because randomization of work sites or locations into intervention or non-intervention groups is seldom feasible, however, this design is less frequently observed in HWB research.[15]

Well-designed research requires carefully considering the outcomes to be studied and looking for published research or behavior change theory to identify solutions likely to influence study outcomes. Work begins with the end goal in mind, which leads program designers to shape the intervention to fit desired outcomes. When a program is focused on changing individual risk factors, this can yield individual level outcomes. With a more comprehensive approach, program goals include assessing and intervening on organizational and environmental policies that influence a population's health, and outcomes related to organizational leadership and organizational supports can also be examined.[18]

The Illinois Wellness study was based on interventions focused at the individual participant level but used only a health assessment, biometric screening, and a menu of classes to try to influence more than 40 different outcomes. The intervention effect is based on participation in any one of the array of courses offered. There is no analysis of the differential impact of one activity versus another, or how specific classes might have been linked to specific health behavior change or health outcomes. It is also unclear what evidence the researchers used to support the hypothesis that this limited wellness intervention alone could reasonably be expected to be successful in changing the observed outcomes within the timeframe of the study.

Be watchful for "confirmation bias."

Confirmation bias is the tendency of researchers to draw inferences from their study that align with their preexisting beliefs but that are not well supported by their data. One of the ways to identify confirmation bias is by looking for study findings within the research study to support each of the assertions or conclusions made by the authors. In the Illinois Wellness study the authors point to "advantageous selection" meaning employees who chose to participate in their offerings were those who, on average, had lower spending on health care. Though the authors state that their study was "not designed to examine recruitment efforts" they nevertheless conclude that a "primary benefit of these programs to employers may be their potential to attract and retain healthy workers with low medical spending." Conversely, the authors acknowledge a "more nuanced story" can be found in these results. While their data show that non-participants spent more "on average" than participants, results also indicate that employees with the highest health care costs were more likely to participate than were the very low cost employees. Given these findings, it could also be concluded that wellness programs attract the highest health care spenders and, therefore, represent a benefit that can address the needs of the costliest segment of the employee population.

The authors further suggest that "wellness programs may act as a screening device by encouraging employees who benefit most from these programs to join or remain at the firm." Related to this, results indicate very high wage earners did not participate in the program. This discussion of "screening" or "advantageous selection" seem to imply that wellness programs can be used to encourage less highly paid employees to stay at the firm while shifting the costs of the wellness program to very high wage earners. Somewhat counter to this inference, however, the study findings show that there were no post-intervention differences related to program participants attending a fun run and, further, the authors "do not find meaningful effects on the average number of days per month that an employee visits a campus recreation facility." One could interpret this finding to mean the programs had no additional benefit for employees who were already active in select, narrowly-defined wellness activities. In other words, rather than conclude that wellness could represent an opportunity to shift costs to less healthy employees, the study results are also consistent with an alternative and decidedly less pernicious conclusion that the program, as designed, simply failed to reach the right audience.

Identify unexpected findings to inform your future approach.

Sometimes research yields lessons or observations not represented in the original research questions. For example, a study's primary goal might be to examine the influence of a HWB initiative on health care costs but leads to discoveries about who is drawn to participate or which program elements most affect program impact. These unexpected discoveries are sometimes important enough to generate media headlines of their own but are overlooked because they are only mentioned briefly in discussing findings. Even if a study's primary research question is not clearly answered, such observations could make the study valuable for future program design or implementation strategies. For example, the authors often referred to "selection bias" and its relationship to recruitment of employees even though their study was not designed to assess recruitment and retention issues. Though the study showed a $4 average monthly health care spending difference between the treatment and control group, the authors focused more on the $1,574 health spending differential between the 39% who volunteered for the study compared to the majority of the population who elected not to be involved with the program. Since this is planned as a multi-year study, one simple way the researchers could address their interest in retention issues is to compare the employee retention trends in coming years between the treatment group, the control group and the 61% of the worksite who were non-participants. Since there is an existing University of Illinois wellness program

available to all employees that runs independent of this research intervention,[19] researchers would need to be intentional about trying to control for contamination effects between the two programs.

Other examples of unexpected findings from the Illinois Wellness study include experimental testing of several incentive strategies within the participant population. Researchers found that financial incentives increased participation but with diminishing returns. The researchers also found that specific incentive strategies yielded higher participation rates, and participants exposed to programs had more positive perceptions about the organization. Any of these findings might have been worthy of headlines but were overlooked by the media. The full value of this research can only be gained by reading the study in detail to assess whether there are ways to apply what researchers learned to future programs. .

Conclusion

Research on the effectiveness of workplace HWB initiatives continues to evolve and every new study – whether its results are supportive of HWB or not – must be subjected to critical scrutiny to determine whether its design and findings support media claims.

The Illinois Wellness study represents very early findings from a limited intervention and study findings must be interpreted with that caveat in mind. The study found that health assessment, biometric screening, and an array of classroom-based educational activities was insufficient to yield health or financial outcomes in the first year of program launch. Appropriately, the study is ongoing, and we will be following it closely to better understand how their program may be linked to the more than 40 outcomes being tracked over time.

In the meantime, here are some questions HERO members can use to critically assess published research studies:

- ☐ What are the characteristics of the studied population? To whom else may these findings be applicable?
- ☐ What are the characteristics of the tested intervention? Does it conform with widely accepted best practices and other published evidence about what is effective?
- ☐ Did the researchers allow enough time between initial intervention exposure and follow-up outcomes measurement? Were there leading indicators of program performance that would detect early that the intervention may not yield expected outcomes?
- ☐ Does the study design take into account the type of program being evaluated? Are comparison groups used and if so, were there any pre-intervention differences between the groups that were controlled for in data analysis?
- ☐ Is there strong scientific grounding that suggests the studied programs should yield the expected study outcomes?
- ☐ Does the discussion of findings suggest “confirmation bias?” Were there examples where researchers draw inferences from their study that seem to align with their preexisting beliefs but that are not well supported by their data?
- ☐ Were there any unexpected findings or lessons learned that have implications for future programs or initiatives?

Cited References

1. Goetzel RZ, Henke RM, Tabrizi M, et al. Do workplace health promotion (wellness) programs work? Journal of Occupational and Environmental Medicine. 2014; 56(9): 927-934.
2. Jones D, Molitor D, Reif J. What do workplace wellness programs do? Evidence from the Illinois Workplace Wellness Study. National Bureau of Economic Research. Available at: www.nber.org/papers/w24229
3. ACOEM, ACE, AHA, et al. A response to proposed Equal Employment Opportunity Commission regulations on employer-sponsored health, safety, and well-being initiatives. Journal of Occupational and Environmental Medicine. 2016;58(3): e103-e110.
4. Fanarow GC, Calitz C, Arena R, et al. Workplace wellness recognition for optimizing workplace health: A presidential advisory from the American Heart Association. Circulation. 2015;131. DOI:10.1161/CIR.0000000000000206
5. Safeer R, Bowen W, Maung Z, Lucik M. Using the CDC Worksite Health Scorecard to assess employer health promotion efforts: A case study at Johns Hopkins Medicine. Journal of Occupational and Environmental Medicine. 2018;60(2):e98-e015.
6. Weaver GM, Mendenhall BN, Hunnicutt D, Picarella R, Leffelman B, Perko M, Bibeau DL. Performance against WELCOA's worksite health promotion benchmarks across years among selected US organizations. American Journal of Health Promotion. 2016. DOI:10.1177/0890117116679305
7. Goetzel RZ, Henke RM, Benevent R, et al. The predictive validity of the HERO Scorecard in determining future health care costs and risk trends. Journal of Occupational and Environmental Medicine. 2014;56(2):136-144.
8. Grossmeier J, Fabius R, Flynn JP, Noeldner SP, Fabius D, Goetzel RZ, Anderson DR. Linking workplace health promotion best practices and organizational financial performance: Tracking market performance of companies with highest scores on the HERO Scorecard. Journal of Occupational and Environmental Medicine. 2016;58(1):16-23.
9. Sorensen G, Sparer E, Williams JAR, et al. Measuring best practices for workplace safety, health and wellbeing: The Workplace Integrated Safety and Health Assessment. Journal of Occupational and Environmental Medicine. 2018. DOI:10.1097/JOM0000000000001286
10. Goetzel RZ and Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. Annual Review of Public Health. 2008;29: 303-323.
11. Goetzel RZ, Roemer EC, Liss-Levinson RC, and Samoly DK. Workplace health promotion: policy recommendations that encourage employers to support health improvement programs for their workers. Partnership for Prevention, December 2008.
12. The Community Guide. Evidence-based findings on worksite health promotion interventions. Available at: www.thecommunityguide.org/content/evidence-based-findings-worksite-health-promotion-interventions
13. Terry P, Grossmeier J, Mangen D, Gingerich S. Analyzing best practices in employee health management: how age, sex, and program components relate to employee engagement and health outcomes. Journal of Occupational and Environental Medicine. 2013; 55:378–392.
14. Grossmeier J, Terry PE, Cipriotti A, Burtaine JE. Best practices in evaluating worksite health promotion programs. American Journal of Health Promotion. 2010;24(3), TAHP 1-9,iii.
15. Goetzel RZ, Roemer EC, Kent K, et al. Health promotion in the workplace program evaluation. In: O'Donnell M (ed.). Health Promotion in the Workplace, 4th Ed. 2014. Troy, MI: American Journal of Health Promotion: 199-263.
16. HERO, PHA. Program Measurement and Evaluation Guide. 2015. Available at: https://hero-health.org/resources/committee-publications/
17. Grossmeier J, Terry PE, Anderson DR, Wright S. Financial impact of population health management programs: Reevaluating the literature. Population Health Management. 2012; 15(3):129-134.
18. Terry PE, Seaverson EL, Grossmeier J, Anderson DR. Association between nine quality components and superior worksite health management program results. Journal of Occupational and Environmental Med. 2008;50:633–641.
19. University of Illinois. Campus Wellbeing Services. Available at: http://humanresources.illinois.edu/campus-wellbeing-services/about.html

HERO