

Crafting a Foundation for Evaluating a Worksite Wellness Program

Teresa A. Hubley
University of Southern Maine

Kay Dutram
University of Southern Maine

Background: Businesses have been exposed to many positive accounts of the benefits of employee wellness to improve employee performance as well as reduce health and injury claims costs for the employer. However, many do not have the tools or experience to effectively demonstrate the benefits of a workplace wellness program for their own management and other stakeholders.

Purpose: This paper presents a) basic components for evaluating a workplace wellness program, b) observations gleaned from evaluating a wellness project, and c) a design for a simple evaluation system that provides information about the efficacy of the wellness program and establishes a foundation for more formal wellness program evaluation.

Setting: Maine

Intervention: Evaluating a wellness program and creating a system of measures for a sustainable evaluation system.

Research Design: Steps in wellness program evaluation are delineated (common definitions, evaluation design, i.e., logic model, data collection tools, data collection, comparison analyses, and

ROI) and carried out. Particulars for a sustainable wellness program evaluation are rendered.

Data Collection and Analysis: We describe how to apply the process and measurements depicted in the logic model. We illustrate a method for calculating return on investment (determine the ratio of known wellness program costs to decreases in injury claims costs and sick leave costs). We recommend components for a sustainable evaluation system based on our experience in actualizing the logic model.

Findings: Applying the analyses we found positive benefits of employee wellness in our Maine DOT case. Workers' compensation hours claims dropped from 875 hours in 2006, to 236 hours in 2007. Strains contributed roughly 17% of the overall injury costs reported for all three years, and almost one third of the costs for 2005. We computed a four-year ROI of \$2.90. To help establish regular and routine wellness program evaluation we describe and recommend additional data sources and measurement points.

Keywords: *employee wellness, wellness evaluation, Maine*

Employers are increasingly drawn to the development of employee wellness programs on the premise that employee job performance can be improved by promoting employee health, which in turn saves money the company

would have spent on health care (Baicker, Cutler, and Song, 2010). Wellness programs offer the promise that if they are effectively designed they will keep healthy employees healthy, support employees with health risks to improve their health

behaviors, and facilitate organizations to achieve workforce performance goals (Ozminkowski, et al. 2002). Published studies also show a savings-to-cost ratio for wellness programs that consistently ranges from \$1.40 to \$5.93 (Fogarty, 2007), an attractive outcome for budget-conscious managers.

Businesses have been exposed to many positive accounts of the benefits of wellness, which supports the concept of wellness programs improving employee performance as well as reducing health and injury claims costs for the employer (Baicker, Cutler, and Song, 2010). However, many do not have the tools or experience to effectively demonstrate the benefits of a workplace wellness program for their own management and other stakeholders. Outside evaluators may be engaged to produce reports that provide evidence for the utility of continuing to provide wellness programs. The ability of evaluators, whether internal or external, to provide useful information depends on the resources workplaces put towards identifying and capturing key information.

This paper presents basic considerations for evaluating workplace wellness programs, along with observations gleaned from a project in Maine, in which a public agency engaged outside evaluators to report on program efficacy and design an evaluation system that could continue to supply agency management with feedback about the efficacy of the program and highlight potential areas for improvement. The client agency expressed a desire to develop the capacity to carry out basic evaluation activities using its own internal staff. Thus, the contracted service included both an evaluation report based on data available to date and a design for simple evaluation processes that could

feed into another more formal evaluation in the future, while providing ongoing, day-to-day information to management. Below, we describe common features of wellness program evaluation and demonstrate how we applied these.

The Case for Wellness Program Evaluation

Evaluation of a wellness program is sometimes an activity new to the agency or business contracting for the service. A business case may need to be developed to ensure that the proposed project can proceed and that recommended ongoing activities will be supported. Fortunately, there is ample literature from both academic and institutional sources available to assist in building this case. Below are steps we found useful in facilitating our project.

Providing a Definition of "Evaluation"

For clients new to evaluation, evaluators may need to begin their work by establishing a clear, mutually understood definition of the term "evaluation" in order to meet the needs of both parties. Employee wellness resources are readily accessible. We used resources from WELCOA (Wellness Councils of America), a well-known resource for worksite wellness that can be a valuable touchstone in this discussion. The website provides evaluation resources that are written in Plain Language style and geared towards employers that desire to conduct their own in-house evaluations. For our Maine-based client, this was a trusted resource and therefore a key ally in creating mutual language.

The WELCOA website defines evaluation as a method for “determining the value of what you’ve done” [see www.welcoa.org]. This simple definition can be a starting point for a mutual work plan. Evaluators can use this definition to establish an understanding that in order to accomplish the goal of defining value that is useful to the client, one must first gather and analyze data about the processes and the desired outcomes of the wellness program and the evaluation project itself. In order to identify the relevant wellness activities and data, evaluators should confirm with program designers and administrators their desired outcomes, the activities and processes of the program, and how they relate to the desired outcomes. Program design information (i.e., what the program desires to achieve and how it plans to achieve it) can be used to map out measurement methods that will produce reliable quantitative information about program effectiveness. It may also be suggested that employee or organization-generated outcomes could be defined as part of the evaluation project or process. The client may be open to exploring many models of evaluation, depending on specific needs.

At this point, the degree of access to the wellness program, including its activities and data, and to employee feedback that will be allowed to evaluators should be established and formally agreed upon. A responsible evaluator would add the caveat that lack of data can compromise conclusions and that some outcomes detected will be less favorable in stakeholders’ judgment.

For our project, as we were dually tasked with evaluating the program to date and creating a system of measures to

be used into the future, we started by confirming program outcomes and design. We developed a logic model that summarized our understanding of how the project was meant to work and used crucial points raised by the client in describing the program. For example, the client expressed a strong preference for assuring that we showed the connection between injury prevention and wellness, two activities that were often separated bureaucratically by the agency but recognized by individuals responsible for these programs to be interdependent. We were also tasked with emphasizing the crucial nature of management support.

The logic model we developed for this assignment served to both confirm the mutual understanding of the resources, activities, and expected outcomes and to graphically depict the types of data and data gathering tools needed to measure the outcomes (Figure 1). For our immediate evaluation report, we opted to use the model and measures to gather as much data as possible.

Referencing the Universal Goal of Wellness

An evaluation client’s workplace wellness program may never have identified its own specific desired outcomes and how the activities pursued in the program relate to these outcomes. Fortunately for evaluation contractors, such programs have a general, common overall goal that can be a starting point. That is, workplace wellness programs, regardless of scope, generally seek to improve the health of employees in order to improve performance and reduce costs associated

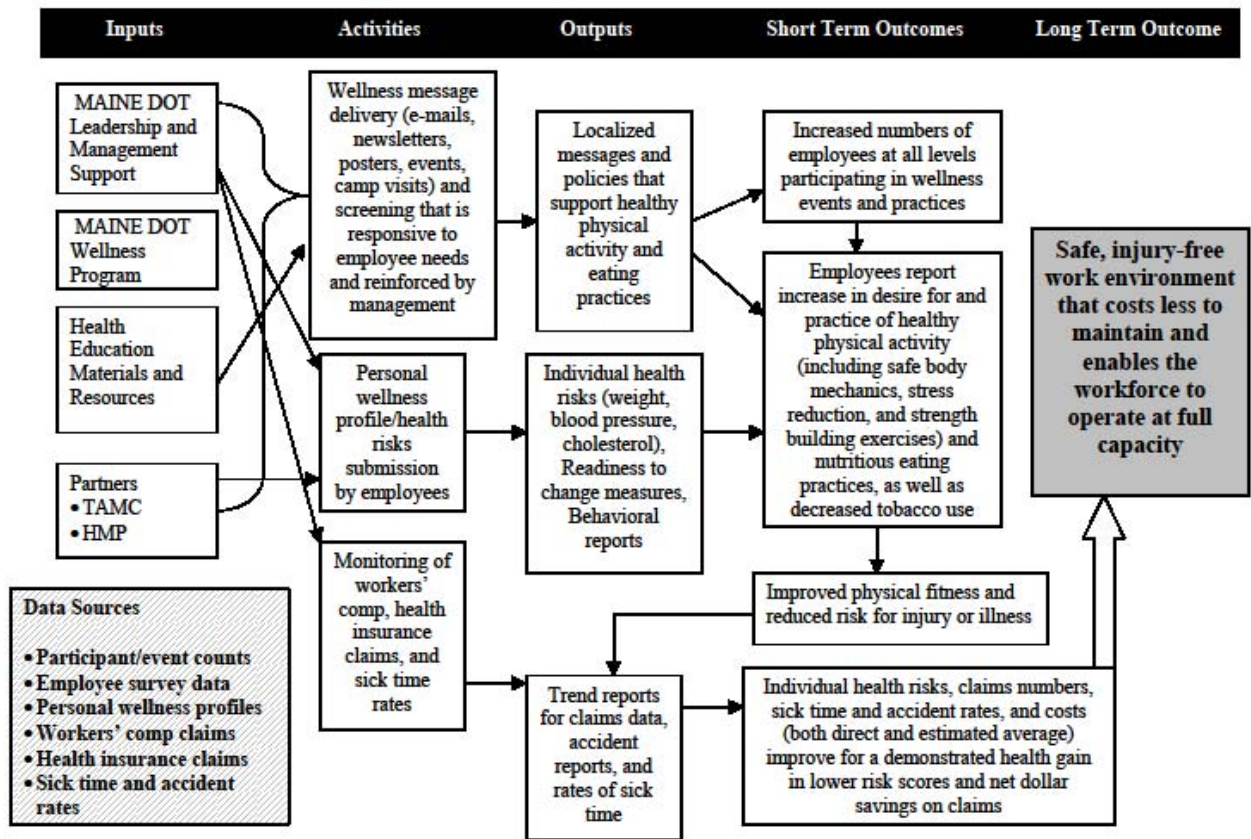


Figure 1. Department of Transportation Region 5 Wellness Program Logic Model

with health and safety. Activities typically address personal behaviors (such as smoking or improper lifting techniques) or company policies (such as allowance for physical activity breaks) that influence risks for costly health claims. For example, employees who are overweight or obese incur elevated lifetime health costs compared to normal-weight employees because of their greater tendency towards high blood pressure, joint damage, diabetes, and cardiovascular disease (Thompson, Edelsberg, Colditz, Bird, and Oster, 1999). Workplace wellness programs therefore often address behaviors that can influence body mass, such as physical activity and diet, in order to reduce the rates of chronic diseases and thereby lower health care claims costs and improve loss of

productivity due to sick time taken from the employer's perspective and improve quality of life for the employees.

Clear definition of expected outcomes, along with appropriate measures, helps lead to a self-sustainable method for systematically defining the impact of a wellness program for an agency or business. In the case of healthy body weight, for example, the long-term goal for this aspect of employee health is to achieve healthy weights for all employees and short-term goals may include decreasing the prevalence of employees who are overweight or obese in percentage increments. The appropriate measure for this set of goals is the trend among employee weights or weight status. If the program has not undertaken such a formal articulation of its intended impact

before, it may be necessary to establish a common understanding of terms such as “outcome,” “indicator,” and “measure” before proceeding to the design of an evaluation strategy. Again, a logic model can be an excellent tool for this part of the discussion. Once “expected outcomes” are used to develop a measurement system, allowances can be made for gathering data on unexpected outcomes.

Describing Key Elements of Wellness Evaluation

Because wellness programs all claim a common long-term outcome, there are standard considerations for a classic outcome evaluation approach in this area, as described in formal literature. The existence of these known elements helps reduce time needed to develop an evaluation plan, as a program can simply build on them. Below are general steps to designing an overall evaluation strategy for wellness programs.

Engage in Logic Model Development to Guide Creation of Measures. A logic model succinctly summarizes the assumptions and theories behind a program and identifies potential measurement methods and intervals (Polacsek, O’Brien, Lagasse, and Hammar, 2006). Logic models are compact enough to be easily presented to and discussed by diverse groups of people. They can be tailored to the audience, being either very formal (perhaps in table form) or casual and colorful (perhaps employing more visual elements, such as circles and splashes). Whatever form they take, they serve to communicate and ensure understanding of the aims, activities, and evaluation strategies of a wellness program to those who would

benefit from, to those who would pay for, and to those who would measure the program’s success. Establishing a definition of programming goals that matches the overall organizational mission and vision, as it relates to employee wellbeing, as well as the program activities that are intended to achieve this goal, is a helpful step towards ensuring the success of a wellness initiative (WELCOA, 2006).

Create or Acquire Appropriate Data Gathering Tools. Wellness program design should include development of a data gathering system capable of supporting reporting and program refinement needs as suggested by the program logic model, though such foresight is not always present at initial inception. The components of the system would ideally then allow the data to be shared and analyzed, with reasonable privacy safeguards in place. Although data would be analyzed and reported in the aggregate, it is important to be able to identify which individuals participated in the program so that overall program impact can be determined and reported in the context of sample size and to track health status improvements by groups or among wellness program participants over time. Some examples of important data sources to establish or confirm include (Goetzel, 2006; Hunnicutt, 2007a; Hunnicutt, 2007b; Polacsek, O’Brien, Lagasse, and Hammar, 2006; WELCOA, 2006).

- Program registration sheets (provide data on how many attended and how many completed in order to show participation rates);
- Participant satisfaction surveys;

- Self-reported behavior surveys (provide current behavioral data for baselines and follow-ups and measure stages of change);
- Health Risk Assessments (provide objective biometric data such as blood pressures, as well as subjective health or lifestyle behavioral data);
- Productivity questionnaires or employer data (provide data on sick leave absenteeism and other productivity measures to show program impact),
- Claims systems (provide medical and workers' compensation use data)

Evaluators may find that some or all of the tools necessary to connect outcomes and activities are missing at the start of the evaluation cycle and therefore may need to deliver their results with appropriate qualifications and recommendations.

Collect Data at Baseline and Follow-up to Show Effects Over Time. Before a program begins, designers should gather baseline data from the population on indicators identified through the logic modeling process and then measure the same indicators again at future intervals in order to track progress at the aggregate/population level (Harden, Peersman, Oliver, Mauthner, and Oakley, 1999; Jeffrey et al., 1993). Evaluators called in "after the fact" often not find complete baseline data in place.

Use Comparison Groups to Show Impact Across the Population. When studying the impact of a wellness program, the ideal (though not always possible) research model includes a comparison group, such as another worksite, that does not receive

the program but does provide data on all the measures in order to show how the program influences variables (Harden, Peersman, Oliver, Mauthner, and Oakley, 1999; Jeffrey et al., 1993; Chapman, 2005; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Baiker, Cutler, and Song, 2010). In the absence of a true comparison or control group, one option that is still considered sound research design is to use participants as their own controls, i.e. measure continuously before, during, and after interventions (Chapman, 2005) or apply interventions some times and not others and compare outcomes over time (Reichardt 2011).

Conduct Return on Investment (ROI) Analyses. Many of the studies we reviewed recognized the importance of showing a positive return on investment for wellness programs to management, even in cases where this was not the main subject of the article (Jeffrey, et al. 1993; Ozminkowski, et al., 2002). As discussed below in greater detail, ROI is the amount of money saved through lowering health-associated costs divided by the amount spent in conducting a wellness program. The resulting figure is often reported as savings per every dollar spent (Fogarty, 2007).

Common Indicators

Several of the items we reviewed in preparing to help our client create an evaluation strategy reflected recurrent themes. In particular, whether they employed more scientific approaches or less rigorous designs, the studies referenced in these articles often relied on the same groups of measures (i.e. discrete units that can be observed and counted), and employing similar indicators (i.e. trends among the units). For example,

most studies suggest weight or Body Mass Index as measures of program elements aimed at affecting eating and physical activity behaviors. A positive effect in this measure would be shown by a downward trend in the rate across the population.

Table 1 presents indicators (defined here as “measures that include trends showing a positive effect at the population level”) that occurred across the sources we reviewed while developing our case.

Key Elements of ROI

While the specific means of calculating the ROI vary slightly from study to study, all use some variation of the ratio of program delivery costs to claim costs borne by the employer that are thought to be affected by the wellness program. Steps to conducting an ROI study are summarized below.

Determine the Cost of Delivering the Program. Identify the salary, material, and logistical costs for a program. Jeffrey et al. (1993) found that the program they studied cost \$1500 per intervention site to deliver. Their sample consisted of a total of 32 intervention sites in various industries with varying staff sizes between 400 and 900 employees. The program cost reported consisted of instructor time and cost of materials. As noted in our case study, travel time and expense can be significant and should be considered in the model as appropriate.

Establish the Cost of Employee Health Risks Over Time. Attach a cost figure to

health risks that may affect a workplace. A simple method is to identify how much money is spent by the employer to ameliorate key conditions created by health risks (such as respiratory illness due to smoking, etc.). Determine the number of valid unduplicated insurance or injury claims related to wellness concerns and translate this number into dollars using payments made on claims. Combined trends in claims costs from recent years can then be projected into the future using multiple regression to show potential savings or loss. Using this method, Ozminkowski, et al. (2002) found a combined savings of \$224.66 per employee for participants in a wellness program, despite an uptick in emergency room visits (i.e. because all other claim numbers related to wellness issues addressed by the program dropped).

Other figures besides payment on claims can be added into a cost model for health risks in order to make the model more robust. Goetzel et al. (2005) included dependent expenses (related outlays to family members through insurance) and productivity costs (measured in salary associated with lost work days due to absences and assorted costs related to replacing employees at turnover) in their cost model. Roslender, Stevenson, and Kahn (2006) also specifically tracked costs related to absenteeism due to accident and included this in their model. A study by Thompson et al. (2001) focused on the costs of medications used to combat conditions created by excessive weight and blended

Table 1
Common Indicators

Indicators	Sources
Increase in number and percent of employees participating in the wellness program	Goetzel et al., 2005; Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Hunnicutt, 2007a; Jeffrey et al., 1993; Ozminkowski, et al., 2002; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998; Roslender, Stevenson, and Kahn, 2006; WELCOA, 2006
Increase in number and percent indicating higher stage of change readiness and engagement	Hunnicutt, 2007a; Polacsek, O'Brien, Lagasse, and Hammar, 2006; WELCOA, 2006
Decrease in blood pressure	Goetzel et al., 2005; Goetzel, 2006; Hunnicutt, 2007a; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998; WELCOA, 2006
Decrease in total cholesterol	Goetzel et al., 2005; Goetzel, 2006; Hunnicutt, 2007a; Polacsek, O'Brien, Lagasse, and Hammar, 2006
Decrease in blood glucose	Goetzel et al., 2005; Goetzel, 2006; Hunnicutt, 2007a
Decrease in reported dietary intake of fats and sugars	Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Goetzel et al., 2005; Goetzel, 2006; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998
Increase in reported dietary intake of fruits and vegetables	Harden, Peersman, Oliver, Mauthner, and Oakley, 1999; Goetzel et al., 2005; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998
Decrease in weight/body mass index/waist circumference	Goetzel et al., 2005; Goetzel, 2006; Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Jeffrey et al., 1993; Hunnicutt, 2007a; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998; Thompson et al., 2001; WELCOA, 2006
Decrease in smoking rate	Goetzel et al., 2005; Goetzel, 2006; Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Jeffrey et al., 1993; WELCOA, 2006
Decrease in reported stress/depression	Goetzel et al., 2005; Goetzel, 2006; Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Roslender, Stevenson, and Kahn, 2006; WELCOA, 2006
Increase in reported physical activity	Goetzel et al., 2005; Goetzel, 2006; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Resnicow et al., 1998; WELCOA, 2006
Decrease in absenteeism and workers' comp claims	Brooks, 2006; Chapman, 2005; Goetzel et al., 2005; Polacsek, O'Brien, Lagasse, and Hammar, 2006; Roslender, Stevenson, and Kahn, 2006; WELCOA, 2006
Decrease in health care claims for emergency room visits, outpatient visits, doctor visits, mental health visits, and inpatient days	Brooks, 2006; Chapman, 2005; Goetzel et al., 2005; Hunnicutt, 2007a; Ozminkowski et al., 2002; Thompson et al., 2001; WELCOA, 2006
Increase in communications from management in support of the wellness program	Harden, Peersman, Oliver, Mauthner and Oakley, 1999; Ozminkowski, et al., 2002; WELCOA, 2006
Increase in job satisfaction/program satisfaction	Hunnicutt, 2007a; Resnicow et al., 1998

these with general claims costs, finding a total cost of \$15,500 over nine years for those with healthy weight versus \$18,700 for those who were overweight and \$21,700 for obese employees.

General costs per person per year at the level of larger population groups (such as nations and state) for specific conditions are available in order to make a related argument that, in general, there are potential savings to be had from addressing health risks. For example, Thompson et al. 2001 reports excess costs for hypertension (\$542.00) and high cholesterol (\$155.00). The Centers for Disease Control and Prevention's *Morbidity and Mortality Weekly* (CDC, 2002) reported excess cost of \$1,623.00 per year per person for smokers and Wang et al. (2001) reported excess cost of \$1,250.00 for sedentary persons. Thompson et al. (1999) found that the average "extra lifetime costs" attributable to excess BMI can rise as high as \$15,000. The WELCOA 2006 case study of International Trucking's "Vital Lives" corporate wellness program determined that employees with diabetes cost the company almost \$8,000 more per employee on an annual basis in health care alone. These are very general figures, however, best used to frame an argument for engaging in wellness activities but too broad to provide specific support for a program.

Establish a Cost-Benefit Ratio. Compare the costs of delivery to the costs of health risks as a straight ratio (costs caused by health risks per person/program delivery costs) and add in time (so that the equation becomes *change* in wellness-related costs per person/program delivery costs). Chapman (2005) reported that across 56 peer-reviewed articles, this method yielded an average return on

every dollar spent on wellness programs of \$5.81. Baiker, Cutler, and Song (2010) commented that Chapman's more lenient inclusion criteria for this study returned a relatively generous figure but they themselves reported an ROI of \$3.27 per dollar spent per employee in medical costs and \$2.73 in costs related to absenteeism. A study by Goetzel (2006) showed the cost of employee healthcare expenditures in general rising 6% in the year 2006 and approaching \$9,000 per employee. In contrast, this author found that across several studies wellness programs yielded an average return of \$1.40 to \$4.90 per dollar spent on wellness programs and another such review showed average reductions in health costs of 26% and 27% as measured by absenteeism (i.e., a lower rate of absenteeism implying fewer sick days and hours paid for by the company and higher productivity).

(Optional) Create Impact Scenarios. Goetzel et al. (2005) used program cost data and claim cost data to project potential savings over time given large, modest, and break-even gains. Using this method, the authors found that even modest gains in indicators (0.1% change per year) generated large savings. The WELCOA 2006 "Vital Lives" case study projected annual savings realized from corporate wellness activities if they had 100% participation (rather than the current variable participation). The program identified a \$5.2 million savings at the current level for seven wellness programs (with costs including the conduct of an annual Health Risk Appraisal) and projected a \$21.1 million dollar savings at the 100% level.

The Case Applied

The Maine Department of Transportation (DOT) Case

In 2007, the Maine DOT engaged our team from the Edmund S. Muskie School of Public Service at the University of Southern Maine to conduct an evaluation of a wellness program in one of the service regions that had been operating since 2004, and appeared to have achieved success in improving the health and absenteeism of workers. The request included design of an evaluation plan that would guide data gathering and analysis into the future and a plan to assist in replicating the program's successes across other regions.

Work began with the development of a logic model as described previously. Subsequently, the logic model was used to create an evaluation plan that could be partially pursued in-house once data needs were identified, as shorter term measures were designed to be easily extracted and reported. Because program evaluation was not included in the original program design, many data elements were missing and those that existed were spread across different systems and sources. Most significantly, the DOT had not been tracking measures at the individual level making it difficult to ascribe any specific impact to the program, since no reliable conclusion could be drawn about how the population treated by the wellness program improved on any indicator.

In spite of the dearth of data, some positive effects could be discerned. For example, there appeared to be a positive effect related to intensive injury prevention efforts focused on strains in particular (as this program was universal).

This was the one area where, with available data, an ROI could be partially computed (Figure 2).



Figure 2. In Sum, Workers' Compensation Hours

Claims dropped from 875 hours in 2006, to 236 hours in 2007. Strains contributed roughly 17% of the overall injury costs reported for all three years, and almost one third of the costs for 2005. While the number of injuries reported rose slightly each year from 2005 to 2007, the cost of injuries incurred dropped each year, from \$100,236 in 2005, to \$80,357 in 2006, to \$31,105 in 2007. With somewhat imprecise program cost data at hand we computed an average employee wellness program cost of \$98.06 per participant per year and a four-year ROI of \$2.90, using changes in costs associated with injury claims data alone. The ROI would have provided a much stronger argument with the addition of data regarding health claims costs related to other conditions, such as lower blood pressure, and more precise cost data for program delivery. Of course, it is possible that with more data entered into the equation the ROI could prove to be negative. This is a potential consequence of which efforts to enhance an ROI figure in any case.

Table 2 shows indicators that emerged from the logic model in the context of available data and suggested changes in data gathering procedures. Final report recommendations included computing more precise cost figures and developing a system that could track related health claims and their costs, as well as participation in the wellness program at the individual level, while reporting them back at the aggregate level. While there were reports generated from Health Risk Appraisals (HRA) on health claims, the HRA system reports used by the region did not address specific conditions that were the focus of the wellness program. For example, for smoking cessation, there was no one diagnostic group that matched this goal and this made it difficult to compute costs or savings associated to smoking cessation wellness program activities. From the HRAs conducted yearly it appears the smoking rate decreased significantly over time for this region so the program may have produced this effect but this relationship could not be verified. We also recommended development of a formal method for eliciting employee feedback, which, we had been informed, had been of some benefit to customizing the program but was not done consistently. Employee feedback can also be used to measure the degree to which the program generally improves the workplace experience for workers, with the caution that there may be little to no relationship between satisfaction with the worksite and program and actual behavior change.

Conclusion

As a result of this work, which was completed and reported out in October 2008, The Maine DOT has decided to

move ahead with development of an integrated data tracking system. Meanwhile, some aspects of the regional program have been replicated elsewhere and have generated some new data, which can be gathered and reported as “mid-point checks.” In the coming years, the agency, having been convinced that wellness saves money and boosts performance, has made a commitment to maximizing their ability to measure the impact of the wellness program at “major checkpoints,” even if this requires contracting external consultants again. Our work contributed to the client’s effort to maintain a responsive and effective wellness program by developing an evaluation plan that identifies the steps they can easily take to carry on with data gathering and analysis activities to support program measurement.

For our team, the greatest single challenge faced in this project was the lack of useful outcome data for a program that had been running for a few years and had little more than tantalizing clues to its effectiveness. Before reaching this particular stumbling block, however, it was necessary to engage the client in developing an understanding of what mechanisms would need to be in place to gather useful data. For their part, the client faced a serious barrier to assessing the program’s impact in that all the necessary data were not available through one source and in one format and no effort was made in the early design phases of the program to produce data that might be used to monitor and improve the program into the future.

An ongoing challenge that arises from this case is understanding and crafting a realistic cost-benefit model for this or any wellness program. While efforts from the literature have involved material costs and

Table 2
Recommended Evaluation Plan for MeDOT Wellness

Measure	Status of Data Source
“Beginning” measures for initial start-up and maintenance	
Development of a program that is tailored to local needs and issues	The region studied achieved this goal by designing and maintaining the program through a local wellness team
# Program modules delivered and Health Risk Assessments completed at # of sites	These numbers were available through HRA summary reports and program calendars
# Participants as % of site employees, all positions	These figures were available only as estimates for modules delivered, as the program did not have a formal sign-up method. However, modules at the 19 worksites were considered mandatory trainings and reportedly highly attended. For HRAs, the completion rate rose over the years studied from around 40% to around 60%.
# Materials distributed by type	No records were available
“Mid-Point Check” measures for tracking effects that are evident only after the program has been operating for some time	
# Follow-up contacts made (i.e. number of times wellness staff interacted with individuals as a result of programming)	No records were available
# Supportive policies made and enforced	On a wider scale, the consideration of wellness itself was a shift in the region’s philosophy. For example, all menus for group gatherings were required by policy to be approved by nutritionists and “flex-breaks” were created to allow group walks and other physical activities. No information on enforcement was consistently available.
% employees showing satisfaction with the program and plans to change behavior in response to program	There was no system to specifically track satisfaction and the behavior changes reflected in the HRA could not be traced back to program attendance or satisfaction because of the lack of tracking on the individual level. Anecdotal evidence, such as a noticed positive shift in the willingness of employees to ask questions at trainings, was offered by those conducting the program at time of our evaluation.
# and % of respondents showing significant changes in major risk areas, such as smoking, weight, fruit and vegetable consumption, blood pressure, and cholesterol, and in the stage of change readiness or adoption	This data was gleaned from HRA reports but, as noted above, the impact of the wellness programming itself could not be traced. The reports showed that the work region as a whole continued to be highly at risk for high BMI and inactivity but had lowered smoking and diabetes rates, while cholesterol and hypertension were roughly the same. Readiness to change (also measured on the HRA) was strongest in tobacco use and lower in healthy eating habits, healthy weight maintenance, and physical activity.
“Major Checkpoint” measures to be statistically tested every few years	
Cost of program delivery (including partnership maintenance)	The only figures available were the cost of providing a stipend to the partner agency that co-produced the program, the cost of administering the HRA, and some staff salaries as related to training and screening. Missing were travel costs (significant in a large region with several worksites), non-HRA material costs, and other relevant costs (if any) incurred by the partner agencies.
# and cost of workers’ comp claims	These were available through the workers’ comp claims system and local injury reports
# and cost of health claims by type	As noted, the types offered in current reports did not match up with program goals and individuals could not be tracked.
Absenteeism rates	These were available through regional reports.
# and % of respondents showing positive change in vital signs	These were available in HRA reports. As noted above, BMI, hypertension, and cholesterol remained constant. However, no connection could be made between the individuals taking the HRA and those involved in the wellness program, again, because of lack of individual tracking mechanisms.
Highlights of participant feedback	No system existed to elicit this data, although some comments were available from staff who administered the program. We suggested use of surveys and the occasional focus group.
Evidence that participant feedback is used to continuously improve the program	The wellness team in this region did report applying participant suggestions to their programming. For example, a lunch box packing demo was developed after hearing comments that workers were not sure exactly how the healthy eating principles discussed could be applied to their own behaviors. However, such direct use of feedback to modify the program is not documented systematically.
Variations and comparisons across regions in measures, program design	At the time of the original project, only one region had administered the program.

salaries on the cost side, our case especially in a large rural district, should suggested travel time and expenses, be included and there are bound to be

other costs that can only be identified by the stakeholders. Likewise, the savings side of the model (needed to compute an ROI) may prove to have much wider implications than “fewer sick days” and “fewer claims.” The participants and managers involved in the program could likely find useful figures, such as lower costs for medication. However, considering the balance between costs and benefits raises the possibility of ambiguities in the formula. For example, employees who take out time from work to exercise may be healthier but they are, in fact, working fewer hours, just as those who eat more fruits and vegetables may be bringing more lunches from home and skipping the workplace cafeteria, vending machine, and events that serve cake.

Despite the complexity of an evaluation study, this case leaves us confident that client agencies can conduct simple self-studies while gathering the data that will help determine longer term impact in a more rigorous manner. These self-studies can bolster the case for a wellness program, as well as make it more responsive and accountable.

References

- Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health Affairs, 29*, 304-311.
- Brooks, B. (2006). Shifting the focus of strategic occupational injury prevention: Mining free-text, workers compensation claims data. *Safety Science, 46*, 1-21.
- Centers for Disease Control and Prevention. (2002). Annual smoking-attributable morbidity, years of potential life lost, and economic costs—United States, 1995-1999. *Morbidity and Mortality Weekly Report, 51*, 300-303.
- Chapman, L. S. (2005). Meta-evaluation of worksite health promotion economic return studies: 2005 update. *The Art of Health Promotion, July/August*, 1-10.
- Fogarty, S. (2007). Evaluating ROI for wellness programs. *Benefits and Compensation Digest, August 2007*, 22-25.
- Goetzel, R. Z., Ozminkowski, R. J., Baase, C. M., & Billotti, G. M. (2005). Estimating the return-on-investment from changes in employee health risks on the Dow Chemical Company's health care costs. *The Journal of Occupational and Environmental Medicine, 47*, 759-768.
- Goetzel, R. Z. (2006). The business case for health and productivity management: What's the return on investment? Presentation for the National Health Communication Conference, Washington DC.
- Harden, A., Peersman, G., Oliver, S., Mauthner, M., & Oakley, A. (1999). A systematic review of the effectiveness of health promotion interventions in the workplace. *Occupational Medicine, 49*, 540-548.
- Hunnicut, D. (2007a). The fundamentals of evaluation: Why, what, how and when you should evaluate your wellness program. *Absolute Advantage Magazine, 6*(10), 13-19.
- Hunnicut, D. (2007b). More than just numbers: nation's leading expert speaks out on the art and science of evaluating worksite wellness programs. *Absolute Advantage Magazine, 6*(10), 24-29.
- Hunnicut, D. & Sabbag, A. (2008). You can't change what you can't measure: the importance of providing an opportunity for health screening”.

- Absolute Advantage Magazine*, 7(2), 30-37.
- Jeffery, R. W., Forster, J. L., French, S. A., Kelder, S. H., Lando, H. A., McGovern, P. G., Jacobs Jr., D. R., & Baxter, J. E. (1993). The healthy worker project: a work-site intervention for weight control and smoking cessation. *American Journal of Public Health*, 83, 395-401.
- Ozminkowski, R. J., Ling, D., Goetzel, R. Z., Bruno, J. A., Rutter, K. R., Isaac, F., & Wang, S. (2002). Long term impact of Johnson and Johnson's health and wellness program on health care utilization and expenditures. *The Journal of Occupational and Environmental Medicine*, 44, 21-29.
- Polacsek, M., O'Brien, L.M., Lagasse, W., & Hammar, N. (2006). Move & improve: a worksite wellness program in Maine. *Preventing Chronic Disease*, 3, 1-8.
- Reichardt, C.S. (2011). Evaluating Methods for Estimating Program Effects. *American Journal of Evaluation*, 32, 246-272.
- Resnicow, K., Davis, M., Smith, M., Baranowski, T., Lin, L. S., Baranowski, J., Doyle, C., & Wang, D. T. (1998). Results of the teachwell worksite wellness program. *American Journal of Public Health*, 88, 250-257.
- Roslender, R., Stevenson, J., & Kahn, H. (2006). Employee wellness as intellectual capital: an accounting perspective. *Journal of Human Resource Costing and Accounting*, 10, 48-64.
- Thompson, D., Edelsberg, J., Colditz, G. A., Bird, A. P., & Oster, G. (1999). Lifetime health and economic consequences of obesity. *Archives of Internal Medicine*, 159, 2177-2183.
- Thompson, D., Brown, J. B., Nichols, G. A., Elmer, P. J., & Oster, G. (2001). Body mass index and future healthcare costs: a retrospective cohort study. *Obesity Research*, 9, 210-218.
- Wang, G., Helmick, C., Macera, C., Zhang, P., & Pratt, M. (2001). Inactivity-associated medical costs among us adults with arthritis. *Arthritis Care and Research*, 45, 439-445.
- WELCOA. (2006). *International: A Platinum Well Workplace Case Study*. Retrieved March 28, 2008 from http://www.welcoa.org/freeresources/pdf/case_study_international.pdf