Best Practices in Evaluating Worksite Health Promotion Programs

Jessica Grossmeier, MPH; Paul E. Terry, PhD; Aldo Cipriotti; Jeffrey E. Burtaine, MD

Setting the Stage

Although there is a consensus that program evaluation is a key component of a “best-practice” approach to designing and implementing employee health programs, it is less clear what constitutes best practice in program evaluation. A recent commentary on the cost benefit of worksite health promotion (WHP) programs addressed the skepticism among some that prevention programs may have merit but are unlikely to produce savings for society. Goetzel argues that worksites are uniquely suited for combining tenets of primary, secondary, and tertiary prevention to improve population health at a relatively low cost compared to the cost of treating conditions that could have been prevented. It is a salient commentary in an economic climate that has caused many organizations to reassess their WHP programs. Indeed, today’s financial challenges serve as a test of the extent to which U.S. employers view WHP programs as an investment in human capital.

Though evidence supporting the cost-effectiveness of WHP is substantial and growing, some purchasers of employee health benefits remain unconvinced that WHP programs can yield a positive return on investment (ROI). Similarly, expectations vary concerning how long it takes to achieve program outcomes and what level of ROI is achievable in the short term and in the long term. Not only do purchasers differ in their expectations about what constitutes program success, but researchers also note wide variation in the kinds of results achieved by WHP programs. What such program reviews share in common is the conclusion that WHP programs can be effective when they employ a best-practice approach. Though there are reports of significant increases in employer-sponsored employee health programs, only a small minority of these programs are comprehensive.

In recent years, several articles have been published to specify what is meant by best-practice programs. With this clarification of quality components in WHP comes the expectation that such well-designed and well-executed programs will deliver superior outcomes compared to common-practice approaches. What remain missing are commonly accepted standards and definitions that delineate what outcomes these best-practice programs will produce.

This edition of The Art of Health Promotion builds on the emerging literature that is defining best practices in WHP by focusing on the following issues:

- What are the measures of success for WHP programs?
- How can program performance metrics be organized into a comprehensive evaluation framework?
- What level of WHP programming is needed to achieve the best outcomes?
- What outcomes can be expected from a best-practice program?
- How does a best-practice evaluation framework support assessment of economic impacts?
- How can a best-practice evaluation framework be used to foster and maintain stakeholder support for WHP programs?
What are the Measures of Success for WHP Programs?

A best-practice evaluation strategy includes transparent and robust measurement of ROI, but it does not stop there. Just as designers of WHP programs aim to provide interventions along the full continuum of health (very poor to excellent health), evaluators must go upstream to articulate the logical, sequential cascade of program outputs and impacts that eventually yields a positive ROI. Such an approach provides not only foundational support for the process that was required to achieve cost savings, but also justification for a comprehensive evaluation strategy and the value of the measurement used from program launch through program outcomes assessment.

Although a best-practice evaluation approach is not new to the field of public health education or prevention practitioners, it has remained largely conceptual in WHP. Case examples are needed in WHP to more tangibly articulate what measures should be tracked at a given point in the program life cycle and how each measure provides foundational support for a comprehensive approach to best practices in evaluation and a credible determination of ROI.

Engagement Metrics

Program participation is a key driver for subsequent results (e.g., health risk reduction and health care cost decreases), so it is important to track multiple measures of program engagement over time, including initial registration, active program participation, and program completion. Although the WHP field has a solid track record assessing initial registration, there is considerable room for improvement around comprehensive program engagement. For example, annual health risk assessment (HRA) and biometric screening programs serve as evaluation mechanisms and gateways for follow-up behavior change and chronic condition management programs. HRA and screening participation, with specific intervals marking progress toward population-wide goals, provides worthwhile additional evaluative information and is relatively easy to track. Participation in follow-up programs requires more sophistication because programs begin with initial registration or acceptance of a triage call but only are successful if participants become actively engaged in the variety of modalities offered or remain engaged in multiple-contact programs. Sexner et al. suggest a series of metrics called a “participation cascade,” which documents a cohort’s movement from initial program enrollment to active participation to completion. Initial program enrollment rates represent the number of individuals targeted for intervention who agree to receive telephonic coaching calls or a series of print-based or Web-based behavior change materials. Active participation rates represent the number of enrolled individuals who complete at least one coaching call, receive at least one print-based mailing, or complete at least one Web-based health education unit.

Satisfaction Metrics

Once individuals are engaged in a program, it is important to ensure that they are satisfied with the program, the caliber of instruction or the quality of coaching, and/or the usefulness of program materials received in support of their behavior change efforts. This increases the likelihood that they will share their successes with others or that their experience can serve as a testimonial to drive further population-level engagement in programs. One positive experience also is more likely to build individual confidence and generate participation in other programs that might require commitment. Participant feedback ideally should be solicited from all individuals who participate in a program and across all programs offered. Specific strategies are needed to garner high response rates from these surveys because the utility of this feedback is contingent on learning from those who have failed as well as those who have succeeded in lifestyle improvement. Feedback may include overall satisfaction with a program, materials, or services, but also may ask participants to report if they learned new knowledge, acquired new skills, or met behavior-change goals as a result of their experience with the program. In accordance with 2009 recommendations by the National Committee for Quality Assurance, this information should be used to drive future program improvements and enhancements.

Health Behavior Change

Health behavior change can be assessed using print-based, Web-based, and telephonic survey instruments that are administered to participants in topic-specific programs. For lifestyle management programs, surveys can assess changes in physical activity, dietary intake, tobacco use, sleep hygiene, stress management techniques, and other topic areas. For chronic condition management, surveys may assess if participants are following clinical care team recommendations, prescribed medication regimens, and condition-specific self-care regimens. For consumerism programs, practitioners may assess changes in health care utilization patterns, preventive exam compliance, or information-seeking behaviors. Survey administration ideally follows a preprogram/postprogram assessment cycle but also may follow a postprogram-only design. The former is more robust but also requires more resources to administer and analyze.

Biometric Health and Clinical Impacts

Biometric health and clinical impacts represent the immediate outputs of disease prevention programs designed to elicit health behavior change or improve chronic condition management. Common biometrics include changes in blood pressure, cholesterol, triglycerides, weight, body mass index (BMI), and hemoglobin A1C levels. Clinical impacts are specific to chronic condition management programs and may include those stated previously plus adherence to condition-specific monitoring procedures, medication and treatment plan adherence, gaps in care, and condition-related emergency room visits, inpatient admissions, and length of hospital stay.

Population-Level Health Risk Reduction

In addition to changes in health behavior and clinical measures tied to specific intervention programs, it is important to track population-level indicators of health status. The measures should be available annually on a representative proportion of the overall population and also should be based on information that is representative of current health status as well as retrospective measures of health that can be captured by claims data review. For this reason, an HRA and biometric health screening programs represent a popular strategy for health assessment. Although HRAs vary in the types of health risks measured, they typically provide some composite measure of health practices (e.g., average number of health risks) that can be tracked over time at an aggregated level to yield population-level impacts in health improvements or declines. If repeat participation in the assessment is high, changes in the composite...
Some care costs, but they also require more savings to overcome their cost. The pursuit of ROI is a delicate balance requiring both enough investment in the right programs and the sustainability of the programs by providing them at a reasonable cost. It is likely that the lowest-cost programs will not yield the best results, but the highest-cost programs are not necessarily the most effective. The basic components of a comprehensive WHP initiative can vary in cost depending on how they are administered. For example, some organizations use cash incentives to foster participation in programs, whereas others integrate incentives into their health benefit plans to achieve the same result for less cost. Although the dashboard of measurement options supporting the effectiveness of WHP programs is growing in sophistication and rigor, ROI calculation still is more art than science and, therefore, is understandably suspect at credibly proving cost savings. What’s more, there is little evidence to show whether more rigorous ROI methodologies will satisfy those who question the value of a preventive approach to health care cost management. Therefore, to be credible, ROI must be supported by the aforementioned measures of program success.

How Can Program Performance Metrics Be Organized Into a Comprehensive Evaluation Framework?

WHP program performance can be evaluated using a variety of analytic tools. The model shown in Figure 1 identifies measures of importance and organizes them into a cohesive framework based on a logical cascade of events and a timeline to demonstrate when each measure is evaluated.

The foundational assumption of the model is that each measure, when shown to record a favorable outcome, also is a marker for the increasing likelihood that subsequent measures also will record outcomes changing in the desired direction. The proposed measures are aligned in an order projected to achieve the expected level of ROI. Moreover, if any of the measures fail to meet expected performance benchmarks, all of the measures that follow are in jeopardy. This allows practitioners to identify challenges early and work to overcome them in order to achieve an organization’s health goals and improvement targets of interest. For example, if participation rates are lagging, a variety of communication strategies can be used, including postcards mailed to the home, electronic messages posted on scrolling text message boards in certain facilities, outbound telephone calls to the home, or posters put up in facilities.

In addition to organizing performance measures into a logical and sequential roadmap for setting goals and expectations, the model shown in Figure 1 also provides a framework for when the measures can be assessed. As a program is launched, the focus should be on program implementation and process evaluation metrics to ensure that the program is gaining early acceptance and traction within the eligible population. Impacts can be assessed as soon as repeat measures are available for comparison. However, enough time must be permitted to pass before significant impacts can be expected. Some health behaviors (e.g., stress management or tobacco use) are more difficult to change than others (e.g., preventive health exams) and may take multiple attempts to result in associated health outcomes. For this reason, changes in composite health practices are best assessed after 6 to 12 months of programming. Outcomes that require cultural support and significant intervention (e.g., weight management) may take up to 2 years to see in terms of population-level shifts in health trends.

If financial impacts are desired during the first year of programming, claims-based savings will most likely need to be targeted at common self-limiting conditions, and programming will need to be focused on the benefits of medical self-care. Changes in utilization related to major
lifestyle risk factors, on the other hand, are difficult to produce at a population level without a significant and comprehensive population health management approach. In fact, recent studies have demonstrated that even the best programs can fail to yield a break-even ROI after only 1 year of programming. 6, 27, 28

What Level of WHP Programming Is Needed to Achieve the Best Outcomes?

A recently published article represents the first to use a deductive approach to identify commonly cited elements of best-practice programs and test the hypothesis that programs with best-practice elements drive outcomes. 11 The authors concluded that a comprehensive program strategy and design is vital to improving health outcomes and fostering positive behavior change. The following case study of the program implemented by the Volvo Group companies in North America, including Mack Trucks, Inc. (Volvo), exemplifies the application of the study’s nine best-practice elements. The case study demonstrates the level of programming required to yield positive outcomes and a level of financial return needed to justify continued investment in the program.

Program Overview

Volvo introduced its Health for Life program in 2004 as a comprehensive employee health management program designed for all employees in North America. The Health for Life program’s purpose is to build a culture of health that supports and encourages healthy lifestyles by providing all employees with ongoing opportunities to maintain or improve their health status across a variety of health management interventions.

Senior Management Support

The most senior program champion is the global chief executive officer (CEO), who recognizes and supports health and well-being throughout the organization. Support also has cascaded down to all of the country-level CEOs and site leaders in North America. The country medical director provides overall program direction and oversight. Although support initially came from the global CEO, it is worth noting that the program did not gain acceptance and financing in the traditional business manner. A business plan was not developed with projected ROI and brought to the “C” suite (e.g., the CEO and the chief financial officer) for approval and funding; instead, the idea was brought to the health benefits department. The department understood the concept and the plan, and costs were put into the overall health care budget. Compared to the total health care budget, the projected WHP budget was so small that it essentially represented a rounding margin in the overall budget. Because the WHP budget was such a small relative portion and was encompassed within the overall health care budget, funding was obtained without immediate need to promise a short-term ROI.

During the inception of the Health for Life program in 2004, senior executives spearheaded various health initiatives, including the construction of state-of-the-art fitness centers at major locations in North America. Other initiatives supported by the CEO include the Volvo Group Health and Well-Being Award. Instituted in 2006, this internal award recognizes best practices and rewards health promotion efforts within the organization. This award has become a permanent program to increase awareness of the importance of well-being within the company. As part of this reward program, the Health for Life program received a special honorable prize in 2006 and was recognized globally during an internal awards ceremony. In addition, in 2007 and 2008, the Health for Life program was awarded the Gold award for Best Employers for Healthy Lifestyles by the National Business Group on Health (NBGH). The program recently achieved its highest level of recognition by NBGH and was awarded the platinum-level award in 2009.

Comprehensive Program Design

The Health for Life program relies on a comprehensive population health management approach. Although there are many definitions of
what constitutes a comprehensive program. Most subject matter experts agree that it includes population-level assessment, targeted follow-up programs for those at highest risk, and population-level health awareness and education activities for all employees. The comprehensive program design includes an on-site program management team consisting of managers, coordinators, and specialists. In addition, the program delivers HRAs, biometric health screenings, preventive health screenings, incentive models, focused behavior change programs, lifestyle management and disease management programs, national wellness campaigns, personal online health Web sites, and on-site fitness centers. It provides a nationwide model of consistent offerings to employees at all worksites while also allowing for local-level customization of program offerings that are most applicable based on the particular health needs of a given subset of the population.

Multiple Program Modalities
Because of the diverse employee demographics within the organization, multiple program modalities are utilized to effectively deliver program components and engage employees in health awareness programs. HRAs are offered via Web-based and print modalities to effectively target certain employee groups. The lifestyle management programs guide participants through the process of health behavior change. The program utilizes self-paced mail programs and one-on-one telephonic health coach counseling programs. The program also provides an extensive array of on-site programs including classes, coaching, and health awareness campaigns. This approach builds participation by making it easy for employees to access programs where they live and work.

Population-Based Health Awareness Programs
The Health for Life program implements a wide spectrum of nationwide population-based health awareness programs. Regularly offered programs include health fairs, behavioral change programs, campaigns, and courses. All programs are employee-focused and designed to encourage long-term lifestyle and behavior changes. The program management team monitors all programming and makes sure employee needs are assessed and appropriate interventions are continually delivered.

Comprehensive Communication Strategy
A comprehensive marketing and communication strategy is in place to inform employees about the Health for Life programs and educate them on why wellness is a core strategy for improving both employee and company health. Some of the most successful channels and tactics include: (1) a program management team; (2) Health for Life Ambassadors, a volunteer network of employees that assists with the program; and (3) professional communication pieces, such as brochures, flyers, posters, and other communication elements.

Incentive Models
A variety of incentive strategies are utilized to encourage participation and engagement in employee health management programs. From 2004 to 2008, a $100 monetary cash award was provided to all non-bargaining-group employees who participated in the HRA and biometric health screening, and $25 gift cards were provided to bargaining-unit employees for participating. Differences in the incentive are attributed to permissible rewards based on negotiations with bargaining-unit representatives. In addition, 2008 marked the introduction of the Healthy People Rewards Program. This program allows non-bargaining-group employees to receive an additional $150 for meeting certain biometric criteria. Participants receive $50 for each criterion that is met, including total cholesterol below 200, blood pressure below 130/85, and BMI below 27.5. Employees who have a medical condition that prevents them from meeting criteria are still eligible for rewards based on establishment of Health Insurance Portability and Accountability Act–compliant alternative standards.

In 2009, an integrated incentive model was launched. The model revolves around employee medical plan contributions and integrating that element with various wellness initiatives. Employees who complete an HRA and a biometric screening, and either do not use tobacco products or complete a tobacco cessation program, will qualify for a reduced medical plan contribution rate in 2010. The intended result is significant cost savings for the employee and increased engagement, awareness, and participation in the Health for Life program.

Biometric and Preventive Health Screenings
The Health for Life program also offers biometric and preventive health screenings to assess and educate employees about a variety of topics. Annual biometric measurements include height, weight, blood pressure, total cholesterol, high density lipoprotein cholesterol, and glucose. In addition, preventive health screenings are offered to enhance employee awareness about a variety of topic areas, such as stroke, breast cancer, and osteoporosis.

Dedicated On-Site Staff
The dedicated on-site program management team is responsible for assessing the needs, culture, and management and employee concerns of the worksite to develop an effective delivery plan and implementation schedule. The team provides a physical presence and directly manages all site-level programming. Management ensures that all activities and programs align with corporate principles and are appropriately utilized at the local and national levels. This ensures that the highest standard of services are delivered and maintained.

Integrated Programs
With numerous health plans and claims administrators interfacing with benefit-eligible employees, it is essential to create a comprehensive health management relationship with these relevant partners. The Health for Life program is dedicated to working with these health partners to integrate data and promote cross-referrals for employees to ensure seamless delivery and optimal results in such areas as chronic condition and disease management programs.

What Outcomes Can Be Expected From a Best-Practice Program?

Discussions about program evaluation often are presented in broad theoretical or conceptual terms that can be frustrating for the applied practitioner. This article translates a best-practice evaluation framework into practice by sharing the details behind the results of Volvo’s best-practice program.

Senior executives at Volvo expect to see annual improvements in employee health and health-related costs in order to ensure that their investment yields justifiable results. Year-over-year data comparisons are provided to demonstrate improvement in population-level engagement, end-user satisfaction, health risk status, preventive exam completion rates, and estimated savings in medical and productivity-related costs. In addition to historical measures of improvement, vendor book-of-business and industry benchmarks are provided for comparison. Volvo relies on a variety of different measurement tools and data collection efforts to track such measures, including an annual HRA and on-site biometric health screening, postprogram evaluation surveys, program-level participation tracking, self-reported productivity assessment, medical claims data, and individual participant testimonials.

Program Participation
Participation in all program offerings is tracked and entered into one integrated data management system. This includes participation in the on-site fitness centers, group exercise classes, on-site chair massage, brief health chats with trained on-site staff, health education programs, self-directed campaigns, targeted health coaching programs, and special
programs or events. To enable targeted outreach, follow-up, and referral to other programs, data are captured by a unique personal identifier. Participation rates are captured and reported for each activity, along with time-over-time participation trends, repeat participation, and participation across multiple programs.

To date, 80% of eligible Volvo employees have participated in an HRA at least once, and 63% have completed the biometric screening, which is offered concurrently. Of those who completed the HRA at least one time, more than half of them (approximately 63%) have since repeated the assessment, which facilitates tracking of impacts on health status over time. Additionally, of those eligible to participate, 18% have participated in at least one targeted multiple-contact health coaching program, and more than 40% of employees have participated in at least one health education campaign. In terms of fitness center utilization, 44% of those eligible are members of the Mack Headquarters fitness facility, 56% are members of the Volvo Trucks Headquarters facility, and 13% are members of the New River Valley manufacturing facility. Of those who are members, approximately 40% to 66% of eligible members utilize the centers at least once each month. Program completion rates are closely monitored for targeted coaching programs and population-based campaigns. Of the 40% of employees who enrolled in population-based campaigns, more than half are estimated to have completed the 6- to 8-week programs. Additionally, an estimated 91% of those who enrolled in a phone-based, mail-based, or online health coaching program completed the program.

Participant Satisfaction
A variety of methods are utilized to evaluate participant satisfaction with programs. The Health for Life program staff conducts focused individual and group interviews, online evaluations, and paper evaluations of all programs. To supplement this, all on-site Health for Life staff talk with employees on a weekly basis regarding various aspects of the program. When applicable, this dialogue has proven a useful measure to gauge satisfaction. Employee testimonials and stories also are retrieved every year to showcase employee satisfaction with the program. Every year the Health for Life program hosts the “Participant of the Year.” A memo is sent out to all employees asking to hear their stories on how the Health for Life program has affected their lives. The stories and feedback that are submitted provide testimonial that these programs work. Employees have shared that they are now being treated for cancer after a screening at work placed them at risk, or that they are now living an overall healthier lifestyle because of the program. With participant approval, these stories also are published and shared with other employees to showcase the importance of living a healthy lifestyle and being aware of one’s health.

Evaluation is embedded into each program or activity to ensure participants are satisfied with the quality of the offered programs, and to track the immediate outputs from each activity. Outputs include acquisition of new knowledge or skills and allow participants to provide ongoing feedback about new programs or resources they would like Volvo to offer. All vendor staff members also are evaluated annually to ensure that Volvo management provides ongoing feedback to the vendors that serve the program. Participant satisfaction rates are consistently high across the entire spectrum of program deliverables.

Health Risk Reduction
Health risk reduction is measured and tracked using the annual HRA (see Figure 2). Across the entire population of employees who completed at least two HRAs, significant change in health risks has been observed: 84% improved or maintained healthy eating habits, 85% increased or maintained recommended levels of physical activity, 79% decreased or maintained healthy stress management behaviors, 90% decreased weight or maintained a healthy weight, and 98% improved or maintained low depression risk. Population-based net health risk reduction has been significant, resulting in an 8% decrease in nine routinely evaluated health risks since program launch, including alcohol use, back care, safe driving, nutrition, physical activity, tobacco use, stress management, weight management, and mental well-being.

Self-Reported Productivity Assessment
Like many organizations, Volvo relies on a standard paid–time-off bank for a significant portion of its employees. Measures of absenteeism and presenteeism are critical when tracking the successes of a WHP
program. Most companies, including Volvo, do not have sophisticated information technology systems to accurately track absenteeism. Only a relatively few employers have accurate measures of productivity. This is particularly true for companies with a mix of white- and blue-collar jobs. Critical breakthroughs recently have been achieved based on research demonstrating that self-reported time away from work and productivity loss at work are valid measures of absenteeism and productivity. Volvo is now able to track these measures by including self-reported questions in the annual HRA.

Measures of work days and productivity lost because of poor health were embedded into the HRA to enable an estimate of productivity savings that could be attributed to participation in the Health for Life program. For individuals who took the HRA at least twice, self-reported productivity measures assessed change in time away from work (absenteeism) and productivity impairment due to the employee’s poor health (presenteeism). Self-reported absenteeism decreased from an average of 2.04 days to an average of 1.87 days per program participant, representing an 8% decrease in absenteeism, or 1087 days. Self-reported productivity improved from an average of 93.65% to an average of 94.04%, representing a 0.4% increase in on-the-job productivity, or 6009 days. When converted to days of time lost, absence and on-the-job productivity loss declined by a total of 7096 days for those with repeat data. Productivity outcomes were monetized using methods reported elsewhere, with productivity-related cost savings of $242 per participant per year demonstrated.

Medical Claims Data Analysis
Because Volvo employees are served by a variety of health plans and multiple third-party administrators, data integration is challenging and costly. Rather than bear the cost of one integrated data management system, Volvo commissioned a custom data analysis 2 years after program implementation. The data vendor integrated all medical claims and health plan enrollment data into one repository along with all program participation data. Data were integrated at the person-centric level to allow assessment of cost trends for each person in the analytic data file. After claims were adjusted for inflation and changes in health benefit plan design, claims were logged to ameliorate the skewness of the data. All individuals who had claims data from 2 years prior to program launch through 2 years after program launch were included in the analysis and separated into participant and nonparticipant groups. Cost savings were calculated using a “differences in differences” approach described elsewhere. This approach represents a quasi-experimental study design, comparing participant health care costs to nonparticipant health care costs. Multivariate regression models were used to control for group differences in age, gender, health plan enrollment, union status, and baseline health care utilization. After adjusting for health plan design changes and inflation, medical costs for employees who did not participate in the program increased by 5%, whereas medical costs for program participants decreased by 3%. The study yielded a medical-only cost savings of $249 per participant (see Figure 3). In addition to the quasi-experimental study, senior management also is interested in assessing total impact on health care cost trends. Prior to program launch, Volvo’s health care spending increased at double-digit rates each year. Total health care trend was at 10% in 2002 and 11% in 2003. In 2004 (the year Health for Life was launched), the total health care cost trend decreased to 9%. The health care trend in 2005 was negative at −8.6%; in 2006, the trend was positive at 6.8%; in 2007, it was 5.1%; and in 2008, it was 5.5%.

Return on Investment
Based on a medical savings of $249 per participant and productivity savings of $242 per participant, the Health for Life program was associated with a total savings of $3.1 million. When compared against all costs invested in the program, including those associated with financial incentives, biometric screenings, on-site staff, classes, and remotely delivered programs, the company achieved the level of savings needed to recoup its costs after only 2 years of implementation.

How Does a Best-Practice Evaluation Framework Support Assessment of Economic Impacts?
Although each of the measures previously described has its limitations, the overall evaluation approach provided the level of credibility and evidence of program success that garnered strong and continued support for the Health for Life program from Volvo’s executive leadership. Just as the Best-Practices Evaluation Framework suggests, any one measure alone is insufficient to meet industry benchmarks of WHP quality. However, providing this full dashboard of success indicators not only
documents program effectiveness, but also serves as a robust planning tool guiding future investment decisions.

Current literature specific to WHP evaluation is suggestive of the priorities that should be afforded to select success measures within such a dashboard. Recent articles, for example, suggest that program participation can serve as a leading measure of success; however, there is little consensus concerning whether certain program components should be weighted as having more population health impact than others. The Health Enhancement Research Organization's Best Practices Scorecard, V2—an industry expert panel-reviewed index—that indicates a need for a best-practice approach to WHP evaluation is consistent with our findings from other research on the benefits of WHP. In one example of an ROI calculation derived from health risk reduction, researchers at Thomson Medstat concluded that achievement of 1% net risk reduction per year would translate to a “large program impact” after 10 years. In program year 1, the Health for Life program achieved a 6.5% reduction in health risks for Mack employees, which exceeded expectations by 3.5%. At Volvo Trucks, the program yielded a 4.5% reduction in health risk, which exceeded expectations by 1.5%. Note that expectations were based on the existence of previous health management programming at Mack prior to the more strategic Health for Life program launch. In subsequent years, consistent risk reduction was realized each year despite the extraordinary stresses on the auto industry along with the aging of the overall eligible population. In the first 2 years alone, population average health risks decreased by 6% from 3.7 health risks in 2004 to 3.5 health risks in 2006.

In the present study, cost-savings estimates were based only on cost trends observed in medical claims. Pharmacy data were not available for the entire baseline measurement period because of a change in the pharmacy benefits administrator just prior to program launch. This and the application of regression analysis controlling for mitigating variables makes the cost-savings estimate conservative. Some common criticisms of quasi-experimental methods are the threats of self-selection bias and regression to the mean. Self-selection bias is difficult to eliminate without the use of a randomized study design or a case-matching protocol. However, the use of multivariate regression procedures that included an analysis of differences in baseline claims costs partly mitigates selection bias concerns. Similarly, because individuals were not selected for participation in this study based on high health care utilization costs, the chance that our findings can be explained by regression to the mean is mitigated by our findings that noncost variables were ameliorated along with cost variables.

The lack of a valid external control group remains a consistent limitation in WHP research, but randomized controlled trials (RCTs) are not a practical or feasible recommendation for evaluation of WHP programs. In the present case study, the program was available to everyone, and baseline health care costs were not extraordinarily different; participant and nonparticipant groups differed by approximately $300 at baseline. Thus, the favorable cost trends that occurred for participants after the program was launched can reasonably be attributed to the program, particularly given the risk reduction that occurred among participants in the program. Indeed, one of the noteworthy findings in this study relates to higher initial health care costs of program participants. This finding affirms the effectiveness of the Health for Life program recruitment model that intentionally directed the higher risk eligible participants to the higher intensity interventions such as health coaching. That the program participants at the advent of the interventions were higher health care users also counters the long-held notion that wellness programs tend to cater to those who are already healthy.

Productivity measurement was based on the self-reported HRA data and was available only for employees who completed the HRA at least twice. Certainly productivity measurements may have decreased because factors not associated with the Health for Life program, and the lack of other data sources to confirm the self-reported changes in productivity is another limitation of this case study. Nevertheless, when asked about program benefits, satisfaction surveys and employees interviewed about the program indicate that the program helps employees to feel that the organization values their health and well-being. The programs are associated with positive feelings of well-being, and individuals state that they feel more energy and vitality as a result of their participation. Furthermore, the productivity questions used for this evaluation have been demonstrated to be highly correlated with the health risk measures that also have improved.

ROI calculations invariably are subject to heightened scrutiny, so researchers included every documentable cost associated with the Health for Life program during the claims measurement period. We found that even after including incentives and biometric screening costs, the program still yielded sufficient savings to break even on the amount invested. Even the cost of the research to conduct claims analysis was included as part of the cost basis for the ROI calculation. We found no ROI studies in the published literature that included research costs as a direct cost related to program delivery.

Our findings in this case study build on a growing body of research that indicates a need for a best-practice approach to WHP evaluation and a dashboard of indicators to guide a program’s success. The present study results compare favorably to other evaluation-driven programs that have successfully engaged employee participants in taking HRA’s, enrolling in health education offerings, and in supporting employees in improving their health habits and reducing their health risks. The ROI reported here also is consistent with the findings of others that ROI is achievable within a few years of program implementation, assuming that the program is comprehensive and provided within the context of a healthy and supportive culture. More research is needed to better determine the interaction between the intensity of interventions offered and the duration of a program as it relates to the level of ROI that can be expected. More practitioner experiences also need to be shared via the literature regarding the feasibility and benefits of organizing the full dashboard of measurements for success that are recommended in this best-practice evaluation framework. Tenets of continuous quality improvement apply to such a framework; that is, improving it is a function of testing it.

How Can the Best-Practice Evaluation Framework Be Used to Foster and Maintain Stakeholder Support for WHP Programs?

In the final assessment, there are two keys to Volvo’s success. The first is having a program champion who is highly placed among company leadership. Volvo’s Corporate Medical Director has had primary oversight of the program since its initiation and has remained highly involved in all aspects of program operations. He serves as a liaison between senior company executives and program staff. The second is placement of the program budget within the health care budget rather than as a standalone “C” suite-funded initiative. If the program had been forced to compete for limited dollars along with manufacturing and engineering departments, funding might not have been sufficient to develop a best-practice program that would serve all North America-based employees as comprehensively. Because of the initial positioning of the program as a long-term investment in employee health and productivity, the program had 3 to 4 years to prove its worth to its
The strength of the Best-Practices WHP Evaluation Framework is that it can be used to set expectations with an organization’s leadership concerning the value of various components of the program, the attendant success metrics related to each component, and the role the full dashboard plays in substantiating the program’s ROI. The framework provides a logical cascade of intermediate events that can be tracked once the program is launched, and the attributes captured in the dashboard will undoubtedly evolve in sophistication over time.

Although it would be easy to criticize the relative weaknesses of each of the measures alone, when combined, they tell a compelling story of cause and effect. When all major benchmarks are in alignment, as in this case study, the resulting ROI is much more credible and defensible. More importantly, systematic and well-planned evaluation provides the substance needed for program quality improvement. However, the framework described here does not replace the RCTs that are so clearly needed for advancing the science behind WHP. Moreover, many more RCTs are imperative to bolster the policy and the business case for prevention in the workplace, and to realize the tremendous potential WHP should bring to national health care reform strategies. At the same time, there is a movement afoot in the health services research community to better inform such research with hard-earned lessons from practitioners. To this end, practitioners not only must put a comprehensive evaluation framework in place where they work, but they also must work to be sure that the results of their best-practice evaluations are being used to inform researchers and program planners in the spirit of continuous ongoing quality improvement.

Jessica Grossmeier, MPH; Paul E. Terry, PhD; and Aldo Capriotti are with Staywell Health Management, Minneapolis, Minnesota. Jeffrey E. Bartaine, MD, is with Highmark BCBS, Pittsburgh, Pennsylvania. Jeffrey E. Bartaine, MD, was with Volvo at the time this study was conducted.

References

Beyond the Evidence-Based Practice Paradigm to Achieve Best Practice in Rehabilitation Medicine: A Clinical Review.


OBJECTIVE: Best practice is a practice that, on rigorous evaluation, demonstrates success, has had an impact, and can be replicated. It is differentiated from its constituent parts, evidence-based practice and knowledge translation, by its general meaning and global purview. The purpose of this clinical review is to provide transparency to the concept and achievement of best practice in the context of rehabilitation medicine. The authors will review and analyze the roles of evidence-based practice and knowledge translation in rehabilitation medicine as they work to support best practice. Challenge areas will be discussed, including an evidential hierarchy in need of update, a lack of “high-level” research evidence, and delays in translating evidence to practice. Last, the authors will argue that rehabilitation medicine is well-positioned to effect change by promoting inclusion of emerging research methodologies and analytic techniques that better capture context-specific rehabilitation evidence, into the evidential hierarchy. Achieving best practice is dependent on this, as well as alignment of all key stakeholders, ranging from the patient, researchers and clinicians, to policymakers, payers, and others.


Developing a “Best Practice” Influenza Vaccination Program for Health Care Workers—An Evidence-Based, Leadership-Modeled Program.

Hood J, Smith A.

Transmission of influenza among health care workers is a well-documented problem. Influenza vaccination is an effective intervention to reduce the influenza burden; however, vaccination rates remain low among health care workers. The challenge for occupational health nurses is how to increase health care workers’ vaccination rates. This article describes the key components of a successful influenza program at a large integrated health care system. A multidisciplinary team developed and implemented an evidence-based, leadership-modeled program that led to improvement in health care workers’ vaccination rates from 66% to 77% in year one and from 77% to 84% in year two.


Kleining TJ, Kimber TE, Thompson PD.

OBJECTIVE: To identify and quantify current deficiencies in primary and secondary stroke prevention, as well as potential gains from optimal employment of thrombolysis. DESIGN, PARTICIPANTS AND SETTING: Observational study of 259 consecutive patients admitted to a tertiary hospital stroke unit from 24 January 2006 to 10 January 2007, with retrospective assessment of prestroke risk factors and therapies to determine stroke preventability, based on relative risk reductions from published meta-analyses of preventive therapies. MAIN OUTCOME MEASURES: Numbers of strokes preventable by optimal risk factor modification and numbers of strokes with preventable disability through optimal thrombolysis; characteristics of patients with preventable strokes; contributions of each risk factor to stroke preventability. RESULTS: 183 patients had a disabling or fatal stroke; 135 patients had at least one suboptimally managed risk factor. On the basis of prespecified stroke preventability weightings, 70 strokes were preventable. The younger the patient, the more likely that the stroke was potentially preventable (relative risk [RR] for age < 60; > or = 80 years, 3.10; 95% CI, 1.96–4.92). Smoking, inadequate control of hypertension and suboptimal anticoagulation accounted for nearly 90% of preventable strokes. Patients with target systolic blood pressures of 130 mmHg or lower were more likely to have inadequately controlled hypertension (RR, 4.27; 95% CI, 2.58–7.05). By comparison, disability could have been prevented in four strokes through optimal thrombolysis. CONCLUSIONS: A significant proportion of stroke remains preventable, especially in younger patients, by optimal modification of risk factors, particularly smoking, blood pressure and anticoagulation. Only a small proportion of patients will benefit from best-practice thrombolysis.


Informing Best Practice With Community Practice: The Community Change Chronicle Method for Program Documentation and Evaluation.

Scott SA, Proescholdbell S.

Health promotion professionals are increasingly encouraged to implement evidence-based programs in health departments, communities, and schools. Yet translating evidence-based research into practice is challenging, especially for complex initiatives that emphasize environmental strategies to create community change. The purpose of this article is to provide health promotion practitioners with a method to evaluate the community change process and document successful applications of environmental strategies. The community change chronicle method uses a five-step process: first, develop a logic model; second, select outcomes of interest; third, review programmatic data for these outcomes; fourth, collect and analyze relevant materials; and, fifth, disseminate stories. From 2001 to 2003, the authors validated the use of a youth empowerment model and developed eight community change chronicles that documented the creation of tobacco-free schools policies (n = 2), voluntary policies to reduce secondhand smoke in youth hangouts (n = 3), and policy and program changes in diverse communities (n = 3).


Orsted HL, Searles GE, Troncelli H, Shapera L, Miller P, Rahman J.

PURPOSE: To provide the specialist in skin and wound care with evidence-based guidelines for care of the person with a diabetic foot ulcer. TARGET AUDIENCE: This continuing education activity is intended for physicians and nurses with an interest in wound care and related disorders. OBJECTIVES: After reading this article and taking this test, the reader should be able to: 1. Describe the pathophysiology, assessment, and diagnostic techniques related to diabetic foot ulcers. 2. Identify current, evidence-based preventive and treatment options for the diabetic foot ulcer.


Methods of Defining Best Practice for Population Health Approaches With Obesity Prevention as an Example.

McNeil DA, Flynn MA.

Childhood obesity has reached a crisis stage and has become a population health issue. The few traditional systematic reviews that have been done to identify best practice provide little direction for action. The concept of evidence-based practice has been adopted in health care, and in medicine in particular, to determine best practice. Evidence-based medicine has its origins in the scientific method and for many researchers this concept means strict adherence to standards determining internal validity in order to justify a practice as evidence based. Practitioners addressing population health face challenges in identifying criteria for determining evidence, in part because of the nature of population health with its goal of shifting the health of whole populations. As well, the type of evidence provided by more traditional critical appraisal schema is limiting. Expanded approaches in finding and defining evidence have been proposed that use: expert panels; broad and inclusive search and selection strategies; appraisal criteria that incorporate context and generalizability. A recent synthesis of 147 programmes addressing childhood overweight and obesity provides a concrete example of using a broader approach to identify evidence for best practice (Flynn et al. 2006). Incorporating evaluation and population health frameworks as criterion components in addition to traditional methodological rigour criteria, this synthesis has identified programmes that provide contextual information that can be used to populate what Swinburn et al. (2005) have described as the ‘promise table’. Using this approach a range in ‘certainty of effectiveness’ and a range in ‘potential for population impact’ are integrated to identify promising strategies. The exercise can provide direction for agencies and practitioners in taking action to address obesity.


In this edition of The Art of Health Promotion the authors use Volvo as an employer that has applied a “best-practices” approach to program evaluation. The commitment of Volvo to excellence in worksite health promotion (WHP) programming is apparent and encouraging, as is the effort expended by Staywell staff as the primary program vendor.

The Selected Abstracts included in this edition of The Art of Health Promotion, in a similar fashion, also underscore the large amount of methodological and applied “best-practices” work that is currently underway in many other health-related fields.

On the eve of a likely major event in the continuing saga of health care reform, we can only hope that the political commitment to evidence-based “best practices” will grow and flourish. The recent visibility of comparative effectiveness research under the federal stimulus package represents a similar desire to find the objective center of “best practices.” For those of us who have been around the field of WHP for two-plus decades this seems long overdue.

However, one of the most important unanswered questions that must be addressed soon is, what range of prevention “targets” and accompanying interventions need to be included in our “best-practice” version of WHP and wellness programs? Should all worksite wellness programs address injury prevention, including home, vehicular and recreational injury prevention in addition to work-related injury prevention? Should all WHP programs include some level of access to coaching for all participants? Should all programs address low back pain prevention and medical self-care? What role should consumer health education play in WHP programs?

It would seem that this particular question related to “best-practice” prevention “targets” and accompanying interventions needs to find objective resolution as soon as possible if the field of WHP is to find its appropriate place in the societal infrastructure of our 21st-century worksite.

Larry S. Chapman, MPH, is Editor of The Art of Health Promotion.
Why Limit Employers to Low-impact Smoking-Cessation Programs—
Comments on “Best Practices for Smoking Cessation”

The Best Practices article by Musich, Chapman and Ozminkowski1 begins with the importance and prevalence of smoking and the fact that there has been all too little reduction in the prevalence over the past decade. The implication is that best practices will be evaluated on the potential impact in reducing the prevalence of smoking in population smokers. But the authors soon made a series of claims that constrained them to focus only on low impact programs. First, they made the incredible claim that “Smoking cessation programs are low impact because there is a population of smokers who are ready or planning to quit smoking and are in the preparation stage of change.” So, they have limited their recommendations to employers that are designed for the small minority of smokers ready to quit, and excluded the vast majority.

But they seek to exclude the most smokers. They exclude because of their apparent belief that there are no smoking cessation programs designed for all smokers, the less than 20% in prevention who are ready to quit, the approximately 40% getting ready in the contemplation stage and the 40% in the precontemplation stage. One would think this article was written more than 15 years ago when smoking cessation programs for all smokers in each stage of change were just becoming available. But today there are a growing array of programs designed for all smokers, including the original state-matched for contemplation stage by Transtheoretical Model (TTM) computer tailored interventions (CTIs) programs with and without counseling. There also are coaching only programs, such as telephonic motivational interviewing and smoking reduction counseling2 and there are free lines for all smokers widely disseminated by the American Cancer Society.

By excluding such programs from consideration and thereby excluding most smokers, what would be the effects on the impacts of programs? The impact equation for a single behavior like smoking in Impact = Reach x Effect (I = R x E). If a program designed for only 20% of smokers who were stage of change by stage producing 30% abstinence or efficacy, the impact would be 1.5% (I = 10% reach x 30% efficacy = 1.5%). This means the prevalence of smoking would be reduced slowly only 5% and would have minimal impact on the smoking related health risk and prevalence of the overall population of smokers.

The authors make another limiting claim that is less clear. They state, “There is no evidence that current programs can systematically move people through the stages of change from precontemplation to contemplation to preparation to decision to quit smoking may be unplanned and spontaneous; nevertheless, the motivation to quit smoking must result from an intrinsic decision for there to be a substantial chance of success associated with participation in programs” (TAHP-3). First, it is unclear how practitioners can better identify smokers who are ready for planning to quit or are unplanned. All of the “best” practices they recommend require planning.

Further, the single study they cited (Aveyard et al., 2003) did not study whether smoking decisions to quit were planned or unplanned. What was studied, was whether a TTM CTI produced greater stage progression than a comparison treatment. But, what they discovered was that their comparison treatment was also stage matched. The authors conceded that a clearly stage-matched comparison treatment would be a better standard.

The statement of Musich et al. that “There is no evidence that current programs can systematically move people through the stages of change ...” (TAHP-3) could be interpreted to mean that programs designed for all smokers in each stage of change would not be effective. Again, they seem to cite the single study of Aveyard et al.4 Unfortunately, they did not cite this author’s invited commentary that accompanied the Aveyard et al. article.5 This commentary made it perfectly clear that there are smoking cessation programs designed for all smokers in each stage of change. It also briefly reviewed some of the evidence of how effective such programs can be across a variety of populations. First, a study by Aveyard’s team compared counseling alone with pregnant smokers in all stages of change to counseling plus CTMs and found that adding the brief TTM CTMs produced 2.2 times greater impact.6

TTM CTIs have produced significant and robust results with representative populations of smokers in the 24% abstinence range at long-term follow-up (usually 18-24 months).7 With 65% of adolescent smokers proactively recruited from primary care, TTM CTIs plus counseling (mostly telephone) produced 24% abstinence compared to 11% in control.8 This is particularly important given that the Surgeon General’s report on adolescent smokers said to forget adolescent smokers since they won’t participate and if they do, they won’t quit.9 Depressed smokers proactively recruited from depression clinics who received TTM CTI counseling and NFT had significant abstinence rates of 24.9%.10 This is particularly important given results with the types of cessation programs included by Musich et al. “Even very low levels of depressive symptoms at the onset of smoking cessation predict dramatically lower long-term abstinence rates” (TAHP-3). It is even more important to note that nearly 50% of all cigarettes are bought by smokers with mental illness.11

Further analyses across five populations have found that men and women smokers were essentially identical in abstinence rates at each follow-up even though it was believed for some time that women are less successful with cessation than are men.12 The oldest smokers over 65 did the best, with over 30% abstinence, in spite of the stereotype that you can’t teach old dogs new tricks. African-American smokers had abstinence rates a few percentage points higher than Caucasians and Hispanic-American smokers did even better.

The point here is that there are stereotypes that some populations don’t have the same ability to change. These results suggest that the issue is not ability to change, but accessibility to appropriate programs.

Another important population are those with multiple health risk behaviors, since they are the highest risk and highest cost population.13 Until recently, there was no programmatic research demonstrating significant effects when treating populations for multiple risk behaviors.14 And, smokers are among the highest risk populations, with only about 10% of smokers in our studies having only one of three or four risk behaviors.15 But, when simultaneously treating a population of parents for three risk behaviors with TTM CTIs, significant impacts were found for all three behaviors with abstinence rates for smokers comparable to programs treated for just smoking.16 Similar results were found with a population of patients from primary care treated for one, two or three behaviors.17 What are the potential impacts of programs designed for all smokers not just the small minority currently ready or motivated to quit? In the population trials cited, the average recruitment rate was about 75%. Multiplying 75% reach times 25% efficacy would produce a 15% impact which would be ten times greater than the impact that could be expected from the “best” practices recommended by Musich et al. Applying a more recent impact equation, would produce even greater impact: I = (Reach x Effect x Σ Number of behaviors changed). By applying multiple behavior change programs designed for at-risk individuals in each stage of change, there is the potential to produce unprecedented impacts not only on smoking but on other major health risk behaviors that plague not only smokers but the majority of employed populations. If we are to keep raising the bar, we need to draw on more inclusive research that supports more inclusive programs that can provide more inclusive care for health risk behaviors that are major causes of chronic disease, increased health care costs and decreased productivity.

James O. Prochaska, PhD
Cancer Prevention Research Center
University of Rhode Island

dedicated to the memory of the late Dr. Vincent P. Prochaska

References


Letter to the Editor
A fusion of the best of science and the best of practice—
together, to produce the greatest impact.

Definition of Health Promotion

“Health Promotion is the art and science of helping people
discover the synergies between their core passions and optimal
health, enhancing their motivation to strive for optimal health,
and supporting them in changing lifestyle to move toward a state
of optimal health. Optimal health is a dynamic balance of
physical, emotional, social, spiritual and intellectual health.
Lifestyle change can be facilitated through a combination of
learning experiences that enhance awareness, increase motivation,
and build skills and most importantly, through creating
opportunities that open access to environments that make positive
health practices the easiest choice.”

(O’Donnell, American Journal of Health Promotion, 2009, 24,1,iv)

“• The American Journal of Health Promotion provides a forum for that rare commodity
— practical and intellectual exchange between researchers and practitioners. ”

Kenneth E. Warner, PhD
Avedis Donabedian Distinguished University Professor of Public Health
School of Public Health, University of Michigan

“• The contents of the American Journal of Health Promotion are timely, relevant, and
most important, written and reviewed by the most respected researchers in our field. ”

David R. Anderson, PhD
Vice Programs and Technology, StayWell Health Management

Stay on top of the science and art of health promotion with
your own subscription to the American Journal of
Health Promotion.

Subscribe today...

ANNUAL SUBSCRIPTION RATES: (Available 1/1/10. Good through 12/31/10)

<table>
<thead>
<tr>
<th></th>
<th>Individual</th>
<th>Institution</th>
<th>Institution</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.</td>
<td>$139</td>
<td>$177</td>
<td>$359</td>
<td>$359</td>
</tr>
<tr>
<td>Canada and Mexico</td>
<td>$148</td>
<td>$186</td>
<td>$359</td>
<td>$368</td>
</tr>
<tr>
<td>Other Countries</td>
<td>$157</td>
<td>$195</td>
<td>$359</td>
<td>$377</td>
</tr>
</tbody>
</table>

CALL 800-783-9913 (U.S. ONLY) or 818-760-8520
OR FIND US ON THE WEB AT
http://www.HealthPromotionJournal.com