

Monitoring and Evaluation of Worksite Health Promotion Programs - Current state of knowledge and implications for practice

Background paper prepared for the WHO/WEF Joint Event on Preventing Noncommunicable Diseases in the Workplace (Dalian/ China, September 2007)

Author:

Luuk Engbers (PhD, PT)
TNO Quality of Life
Sector Physical activity & Health
P.O. Box 2215, 2301 CE, Leiden
The Netherlands
(www.tno.nl)

With assistance from:

Jacob Sattelmair (M.Sc.)
Doctoral Student, Epidemiology, Harvard School of Public Health (Boston, USA)



© World Health Organization 2008

All rights reserved. Publications of the World Health Organization can be obtained from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int). Requests for permission to reproduce or translate WHO publications – whether for sale or for noncommercial distribution – should be addressed to WHO Press, at the above address (fax: +41 22 791 4806; e-mail: permissions@who.int).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

The named authors alone are responsible for the views expressed in this publication.

Foreword

In response to the growing burden of chronic diseases, the World Health Organization (WHO) developed the Global Strategy on Diet, Physical Activity and Health (DPAS), which was formally adopted by the Fifty-seventh World Health Assembly in May 2004.

The workplace setting is clearly identified in DPAS's 62nd Paragraph as an important area of action for health promotion and disease prevention.

As a further step of DPAS implementation, WHO together with the World Economic Forum (WEF) is hosting an Event on Preventing Chronic Diseases in the Workplace through healthy diet and physical activity promotion, that will be held in Dalian (China) on the 5th - 6th September 2007. Experts from academic institutions, ministries of WHO Member States, private industry and NGOs will participate in this Joint Event.

Purpose

This paper presents a broadly focused review of literature, to describe the current state of knowledge concerning best practice methods of monitoring and evaluation of chronic disease prevention programs in the workplace that address healthy diets and physical activity.

The main purpose of this paper is to provide the background and basis for the Joint Event, and will be a focus point of discussion and information for the meeting participants.

Content

Executive summary	5
1. Introduction.....	9
1.1 <i>Physical (in)activity, diet and health</i>	9
1.2 <i>Environment, overweight and obesity</i>	10
1.3 <i>Health promotion at the workplace</i>	10
1.4 <i>Evaluation and monitoring of WHPPs: Concepts & theories</i>	11
2. Aim and objectives	14
3. Methods.....	15
3.1 <i>Types of studies</i>	15
3.2 <i>Participants</i>	15
3.3 <i>Intervention content</i>	15
3.4 <i>Outcome measures</i>	15
3.5 <i>Exclusion criteria</i>	16
3.6 <i>Search strategy</i>	16
3.7 <i>Case reports</i>	16
4. Included reviews	16
4.2 <i>Study designs of the included studies</i>	17
5. Worksite interventions: a short overview	17
5.1 <i>Behavioural/social approaches</i>	17
5.2 <i>Environmental/ policy strategies</i>	18
6. Indicators for success of WHPPs.....	20
6.1 <i>Self-reported dietary intake</i>	20
6.2 <i>Objective measures of diet</i>	22
6.3 <i>Self-reported PA outcome</i>	22
6.4 <i>Objective PA measurement</i>	25
6.5 <i>Measuring the environment</i>	27
6.6 <i>Determinants of PA or dietary behaviour</i>	29
6.7 <i>Biological disease risk indicators</i>	30
6.8 <i>Process indicators</i>	31
7. Best Practice: Recommendations for M&E of WHPPs	33
8. References.....	35
Appendix: Case reports.....	41

Executive summary

Introduction: There is conclusive statistical evidence that the majority of the population in highly developed (or “high income”) countries, such as the United States, Great Britain, and Australia, are overweight (BMI >25), and that the prevalence of cardiovascular diseases (CDs) and other disabilities related to obesity is also very high (Seidell, 2002). In the U.S., approximately 30% of the population is currently obese and about half (46%) of the Dutch population has overweight and 11% is obese, and these numbers are increasing (Visscher et al., 2001; Haslam et al., 2005). The evidence that restoring energy balance, i.e. sufficient physical activity and a healthy and modest food intake, is beneficiary for good health is well established. Due to current increases in obesity and related diseases and disorders, health promotion to stimulate a healthy lifestyle is necessary, and the worksite is considered to be a practical and effective setting for interventions. However, to satisfactorily obtain evidence regarding the success or failure of worksite health promotion programs (WHPPs), several theories or concepts (described in chapter 1) suggest that quite a few levels of measurement need to be taken into account in order to adequately interpret the findings of such assessments.

Aims and objectives: This paper aims to gain insight into the current state of knowledge with regard to best practice methods of monitoring and evaluation (M&E) of WHPPs that address healthy diets and physical activity (PA). Specific objectives were first to give an overview of frequently used indicators and measures in M&E, second, to comment on the feasibility of using these methods for M&E of non-scientific WHPPs, and third, to formulate recommendations on M&E of WHPPs.

Methods: Strategic collection and synthesis of information from existing reviews of peer-reviewed literature of evaluations of WHPPs (published in the last 10 years), complemented with ‘grey’ literature (e.g. case reports from the private sector--see appendix) on recent evaluations of and initiatives in worksite health promotion

Results: The literature search produced 10 reviews in which the effectiveness of WHPPs on physical activity and/or food habits (nutrition) was evaluated. Three reviews (~57 worksite studies) specifically focused on the effectiveness of worksite PA interventions, five reviews included multicomponent, broadly focused (e.g. Reducing cardiovascular or cancer risk) health promotion programs (~115 studies) aimed at stimulating both PA and/or a healthy diet, and two reviews (18 studies) focused on worksite environmental/policy intervention programs. The following indicators were identified:

Self-reported and objective measurement of diet: outcome indicators on fruit, vegetable or fat intake to measure a change in frequency of the consumption (portions, grams or fat-points per day) were most frequently used in multicomponent WHPPs to reduce health risks (cancer or cardiovascular diseases). Disadvantages are that these questionnaires are subject to many types of bias and most questionnaires are considerable in length. Current developments indicate that shorter and web based food screeners are available which provide new opportunities for M&E of WHPPs. In only a few (mostly environmental) studies objective outcomes were utilised such as sales dates of products in restaurants of vending machines or observation of purchases per lunch tray. These objective methods are less susceptible to bias, and considerably

reduce the burden of measurement on employees, and should therefore increasingly be used.

Self-reported and objective measurement of PA: Self-reported PA was measured using a wide variety of PA dimensions, ranging from frequency only to measures of intensity and duration. Similar to self-report of diet, the measurement PA by self-report is difficult and subject to many types of biases. Moreover, it could not be concluded if validated instruments were used. In the majority of the studies in the included reviews a change in general PA-behaviour outside the workplace was assessed, however it seems only reasonable to measure PA at the worksite as well. Only a limited set of questionnaires is available in literature that includes items on this matter. In addition, research indicates that self-report instruments are not capable of adequately capturing these kinds of routine light-intensity activities.

Based on the studies in the included reviews, objective measures (like pedometers, accelerometers and stair use measurement) to capture these activities are rarely ever used, but should be used more frequently in M&E of WHPPs. Research shows that current technical innovations to these instruments allows feasible and valid measurement of worksite PA.

Environmental measures: Despite the increased attention on environmental determinants of behaviour, none of the worksite studies reported the use an instrument to assess the worksite environment (e.g. availability of healthy food products, means to be active or showers, health policy on active commuting, etc) or the perceived environment by employees as an indicator of success. In literature, only three of such instruments could be found (CHEW, NEMS-R and the WEM). These environmental instruments provide new opportunities in M&E of WHPPs and should be used more frequently.

Determinants of behaviour: Only very few of the studies in the included reviews reported the use of some type of measurement of determinants of behaviour in the outcome assessment. Some constructs such as attitude, social influence and/or stages of changes were identified and are most commonly used in health research. An important determinant that is not usually mentioned in models explaining behaviour, or used in M&E of WHPPs, is habitual behaviour, or habits. As most PA at the worksite occurs habitually, the measurement of this determinant is considered to be important to understand how environmental and worksite behaviour interact. The assessment of intermediate outcomes (i.e. determinants) in M&E can help guide the development of successful WHPPs.

Biological indicators: Although frequently used as indicators of success in more broadly focused multicomponent WHPPs, the effectiveness of these programs on biological indicators like body height/ weight, blood pressure and cholesterol have yet to be demonstrated. WHPPs focused on stimulating PA have shown to be more effective on more direct outcomes like increasing physical fitness. Moreover, extensive measurement of such indicators is expensive and requires trained personnel. A smaller set of biological indicators (finger stick cholesterol and fitness tests) for more practical use in M&E of WHPPs might be sufficient.

Process indicators: It is evident that many process evaluations (e.g. investigating how the program was implemented, by whom, and with what level of effort) of WHPPs do

not get published, nevertheless none of the studies in the reviews included in this paper reported the measurement of process indicators as part of their outcome evaluation, despite the fact that process evaluation is often mentioned as an important part of the evaluation of WHPPs. Only two publications on the finding of process evaluation were found in literature. Several aspects to include in such process evaluations are described. The collection of quantitative and qualitative information on program implementation and delivery should be collected and published more often, because it provides meaningful information on the interpretation of the findings (success or failure) in M&E of WHPPs.

Best practice: Recommendations for M&E of WHPPs in practice

1. To adequately interpret the findings of M&E of WHPPs, outcome indicators should be included on each level of formative, process, intermediate and impact evaluation.
2. Outcome indicators should directly be related to and dependent on intervention components and objectives. They should be logical consequence of decisions made in each level of evaluation as mentioned in point 1.
3. An extensive process evaluation should always be included in M&E of WHPPs; qualitative information on program preparation and implementation is just as important as the impact of WHPPs. Process evaluations will provide useful information on how to make WHPPs more successful.
4. Due to an increasing body of evidence regarding environmental determinants and the effectiveness of such interventions, in M&E of WHPPs at least a measure to quantify the inside (worksite) or outside environment has to be included. An increasing number of such instruments are currently developed for research purposes.
5. To decrease the load of data management for health researchers and to increase the response of subjects to PA or dietary questionnaires, the use of validated and shorter *internet* or *intranet* questionnaires should progressively be used more in M&E of WHPPs.
6. Questionnaires to quantify PA should include items on occupational PA.
7. Technical developments provide new, valid and more feasible means to objectively measure occupational light-intensity PA (tri-axial accelerometers, pedometers and stair use measurement) and dietary habits (specified sales data in restaurants and vending machines). Consequently, the use innovative objective instruments should be stimulated and utilised more frequently in M&E of WHPPs.
8. The inclusion of an extensive set of biological indicators, like the measurement of full spectrum blood lipid analysis, determination of body composition and blood pressure, in WHPPs might not be necessary or feasible. In correspondence with program components and objectives a relatively small set of feasible and less expensive biological indicators (i.e. finger stick cholesterol, fitness or strength

tests and/or body weight and length) might be sufficient in M&E of WHPPs in practice.

9. In combination with the continuous monitoring of sick leave at most worksites, regular (yearly) health check-ups of employees should likewise be incorporated in company health policy. Including the recommended small set of biological indicators and/or a questionnaire, continuous health monitoring:
 - Will give insight into long term in health changes
 - Will automatically measure health changes due to newly implemented policy or intervention elements.
 - Will lead to effective WHP by gaining insight into successful or unsuccessful interventions of policy elements so that health trends can be anticipated.
 - Can be utilised as a bench-mark of company`s health policy
 - Will contribute to the employee`s perception of the commitment of the company to occupational health management.
 - Continuous data flow can make cost-benefit analysis achievable.

1. Introduction

1.1 Physical (in)activity, diet and health

Many of the chronic diseases (cardiovascular diseases, some types of cancer, type II diabetes, etc.), which are faced today are associated with obesity, which is due to an increasingly sedentary modern lifestyle and an unhealthy diet. (Haskell et al., 2007, 1995; Blair and Connelly, 1996; U.S. Department of Health and Human Services, 1996).

Statistics indicate without a doubt that the majority of the population in highly developed countries, such as the United States, Great Britain, and Australia, have a BMI of >25 (overweight) and that the prevalence of cardiovascular diseases (CDs) and other disabilities in relation to obesity is also very high (Seidell, 2002). In the U.S. an estimated 30% of the population is currently obese and approximately half (46%) of the Dutch population has overweight and 11% is obese, and these numbers are increasing (Visscher et al., 2001; Haslam et al., 2005). However, obesity is not only a problem in Western industrialised countries; research has concluded that even in large areas of Africa, Asia, and Latin America overweight (BMI > 25) is increasing. In China an alarming increase in the number of overweight and obese people has recently been observed; for instance, 23.9% of the younger population (20-45y) has obesity (Asfaw, 2006; Popkin, 2006 & 2008).

Evidence that restoring energy balance, i.e. sufficient physical activity and a healthy and modest food intake, is beneficiary for good health has been well established. However, in western industrialized countries only about a third of the population is sufficiently physically active (i.e., active at a moderate intensity for 30 minutes per day for at least five days per week). In the US 40% and in 15 EU member states on average only about 31% of the adult population meets this public health recommendation for moderate physical activity (CDC/ACSM physical activity guideline) (Haskell et al., 2007; Pate et al., 1995; EORG, 2003). Moreover, the population attributable risk (PAR) of a sedentary lifestyle for mortality from cardiovascular diseases, colon cancer and type II diabetes, is estimated to be 35%, 32% and 35% respectively. This indicates that on average 35% of the deaths from the abovementioned chronic illnesses could be prevented if the entire population was sufficiently physically active (Ruwaard et al., 1997; Powell et al., 1994).

Besides increasing physical inactivity, the eating patterns in industrialized countries are characterized by high energy intake and over-consumption of (saturated) fat, cholesterol, sugar and salt. Low energy expenditure (physical inactivity) combined with high and unhealthy energy intake will lead in time to overweight and obesity (WHO regional publication, 1988). Besides lowering the intake of energy dense foods, maintaining a healthy diet (i.e., a low saturated fat intake and high fruit and vegetable intake) has also found to be important in the prevention of overweight and consequently chronic health problems (Visscher et al. 2002 & 2001; Seidell et al., 2003). Moreover, in a systematic review of literature (Hu et al., 2005) that evaluated the evidence regarding diet and CVD (cardiovascular disease) prevention, substantial evidence was found that diets including unsaturated fat and an abundance of fruits and vegetables offer protection for CVD. However, the authors mentioned that such diets have to coincide with regular physical activity, no smoking and maintaining a healthy body weight (Haslam et al., 2005).

1.2 Environment, overweight and obesity

The dramatic increase in the prevalence in overweight and obesity (i.e., obesity epidemic) in the US, which started in the beginning of the 1980's (Jeffrey et al., 2003; Haslam et al., 2005) after a period of relative stability, is widely agreed to be the result of a changing environment. Similar trends in overweight and obesity and changes in the environment have since then also occurred in other parts of the world. This 'new' environment is considered to be an important contributor to a sedentary lifestyle.

The short time frame in which the recent increase in the prevalence of obesity has occurred adds substantial support to the concept that the causes for this increase are due to environmental changes rather than biology per se (i.e., in physiological or genetic changes). The underlying idea behind this assumption is that humans were primarily designed for activity (Sparling et al., 2000). Throughout most of human history, physical demands (i.e., household chores, tool making, hunting, farming) were a typical aspect of daily life. In modern times, many of these demands on the human body are no longer part of daily life due to of the mechanization of the society, i.e., an increased use of automobiles, decreased opportunities to walk (e.g., no sidewalks) and bad connectivity (e.g., not being able to walk or bike between home and shopping areas) in many modern living areas, and an increase in televised and computerised entertainment.

The abovementioned changes in physical demands have also occurred at the worksite, as most job descriptions have changed from manual labour to predominantly physically inactive duties (e.g., desk jobs, automated assembly lines, etc.). The following daily routine might be a reality for many employees; travel to work by car, take the elevator to an office and sit behind a desk for the majority of the working day. At the end of the day, many workers commute home by car or train, and then when at home, inactivity may continue, as computerized and television entertainment are popular. This transition from a physically demanding environment to a predominantly mechanical environment focused on convenience, has significantly contributed to the decline decrease in physical activity levels. The following factors are often cited as important components of changing environment that may have had consequences for dietary habits: convenience in food availability (e.g., increase in food stores, vending machines, etc.), increasing availability of energy dense, nutrient poor (ready-to-eat) foods, and greater amount of meals consumed away from home in large portions (i.e., pizza, Mexican food, fast food culture) (Popkin et al., 2005) This 'new' environment can be called the 'obesogenic' environment (Swinburn & Egger, 2004) where the choice to make healthy decisions has become increasingly difficult and more importantly not obvious for most individuals. In this regard, unhealthy behaviour has become a normal response to an abnormal environment.

1.3 Health promotion at the workplace

As described above the unfavourable changes in the worksite environment might significantly contribute to unhealthier diet and sedentary behaviour of employees. On the other hand the workplace is considered to be an excellent setting to attempt behaviour/lifestyle change among workers, because a large proportion of the population can be reached and workers spend about half their waking hours at the workplace. Therefore a large number of reviews can be found in literature in which scientific trials on the effectiveness of worksite health promotion programs (WHPPs) are evaluated.

In general a WHPP can be defined as: "The combination of educational and environmental supports for actions and conditions of living beneficial for health"

(Green et al., 1998). The term ‘environmental supports’ or ‘strategies’ in this definition can be interpreted as: “All strategies that aim to reduce barriers or increase opportunities for healthy choices, e.g. by providing more healthy options, by making healthy choices more accessible and by establishing policies that require healthy choices, or restrict the number of less healthy options” (Glanz et al., 1986). Or as Glanz et al. (1988) defined it more succinctly: “All strategies that do not require the individual to self-select into a defined educational program (i.e., self-help programs, classes or groups)”. In a review of Kahn et al. (2002) three types of (worksite) interventions were identified:

1. Informational approaches to changing knowledge and attitudes about the benefits of and opportunities for physical activity and a healthy diet within a community (worksite or school).
2. Behavioural and social approaches to teach people the behavioural management skills necessary for successful adaptation and maintenance of behaviour change and for social environments that facilitate and enhance behaviour change (i.e. physical activity/ fitness programs).
3. Environmental policy approaches to change the structure of physical and organizational environments to provide safe, attractive, and convenient places for physical activity (i.e. walking tracks, attractive staircases, and fitness facilities).

1.4 Evaluation and monitoring of WHPPs: Concepts & theories

In monitoring and evaluation (M&E) of WHPPs, several levels of measurement should be recognised, as several papers have asserted (Centre for Health Promotion, 2005; Bauman 2006, Nutbeam, 1998; Kwak, 2007). In the paper by Bauman et al., a conceptual model for evaluating WHPPs is outlined (Figure 1) which is roughly equivalent to the concept developed by the Centre for Health Promotion Canada. However, the terminology used to describe similar levels or stages of M&E differs between the two reports.

According to Bauman et al. the first level of evaluation is the preparation/development phase of a health program, which can be assessed by means of a *formative evaluation*. In this phase the responses of the target group regarding the PA and diet messages and program components are collected and summarized. In addition the stakeholders can be interviewed to gain insight into the program’s likelihood of success. After the findings of this phase are interpreted, an adapted and hopefully improved version of the program can be implemented. This updated version of the intervention potentially has a better fit with the specific work culture and organisational details of the targeted worksite.

A second level in the evaluation is the (continuous) *monitoring* of the process and the reach of program components. Aspects in this evaluation that can be monitored are, for example: attendance of the participants; program delivery in comparison to that which was intended, and an assessment of whether environmental changes were implemented at the worksite. According to Nutbeam (1998), it is essential to investigate the manner in which the program was implemented, by whom, and with what level of effort. Only through this type of investigation will more be learned and understood with regard to success or failure in achieving the defined outcomes.

However, it is only through the following levels of measurement that the actual impact or success of the program will be assessed. The third level is an attempt to assess the *proximal* or *intermediate* effects of a program. Outcome indicators that fall within this level are generally referred to as the determinants of behaviour (e.g.

knowledge awareness, self-efficacy, social supports, beliefs, perceived environment). As Nutbeam et al. (1998) states it, ‘these are the cognitive and social processes or skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health’. These determinants are essential to assess and interpret the causal link between the intervention and intended or actual performed behaviour. The ASE-model, social learning theory, and/or the theory of planned behaviour are often used as a theoretical basis for these outcome measures. All these theories are based on the notion that behavioural decision, motivation or intentions are the primary determinants of behaviour.

The actual health impact of a WHPP is considered a fourth level of evaluation and includes measures of *energy balance related behavioural change* (e.g. objective or self-reported PA and/or dietary measures), but may also include biological (cardiovascular) disease indicators such as BMI, blood pressure, percent body fat, or cholesterol levels. In accordance with the definition of WHP provided in the previous paragraph, the fourth level of evaluation also encompasses physical environmental measures which identify the completed modifications in the environment.

The final level of evaluation is the potential long-term outcome of a WHPP. This level includes a measurement of the possible reduction in disease incidence. It may take several years to achieve results on this level, and therefore it is not often measured.

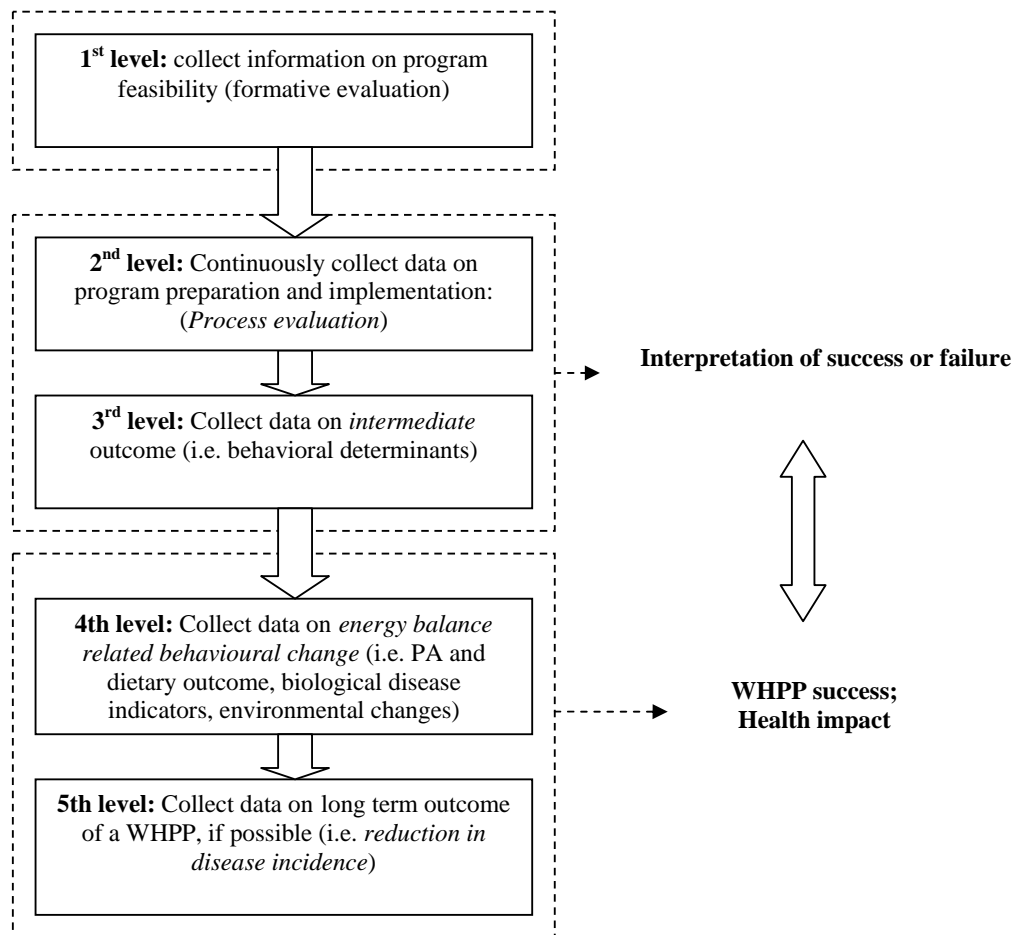


Figure 1. Levels of measurement in the evaluation of a WHPP (based on Bauman et al. 2006)

Another concept which can be helpful in setting up and evaluating WHPPs is called intervention mapping (IM). IM explicates a series of five steps for the development of a health promotion program based on a foundation of theoretical, empirical and practical information (Kwak et al., 2007). Each step is comprised of different tasks, which result in a clear end product, providing the foundation for the next step. The defined steps are, [1] defining proximal program objectives, [2] selecting theory-based interventions and practical strategies, [3] producing and integrating the intervention program, [4] adaptation and implementation plan, [5] monitoring and evaluation plan. In the fifth step decision on which process and indicators of success to include are made. However each decision in the 5th step has to be a logical and direct consequence of the previous four steps, as to ensure that the best indicators of success for a specific program are chosen. Similar to the concept presented by Bauman, the indicators of success can be divided in indirect (i.e. implementation and awareness of intervention components), intermediate (i.e. determinants of behaviour) and outcome indicators (i.e. energy balance related behaviours and indicators). As mentioned earlier in IM the decision to include long term outcome like disease incidence might in many cases not be considered a logical step in the evaluation of WHPPs, as there usually no direct causal (short term) relationships between the developed intervention components and the incidence of chronic disease.

Common to all previously described models or concepts is the concept that program objectives have to be clearly defined from the start, after which appropriate intervention components can be developed based on existing theories and knowledge. Finally and most importantly, when selecting outcome indicators it must be considered that the chosen indicators have to be directly related to the program objectives and components. In addition, when selecting the set of outcome indicators the majority of the stages in the (hypothesised) chain of events leading to the desired outcome of the program should be addressed, i.e. opening the so-called black box. For this reason it can be argued that the long term indicators such as a reduction in chronic disease incidence should only be incorporated as an element of the rationale for the program, but seldom can or should the program be held accountable to the achievement of these indicators (Bauman et al., 2006). Finally, in M&E of WHPPs, outcome and process assessment need to go hand in hand in order to obtain answers on whether or not the program has worked effectively and can be repeated or refined (Nutbeam, 1998).

2. Aim and objectives

As can be concluded from the previous sections; due to current rises in the prevalence of obesity and related chronic diseases and disorders health promotion to stimulate a healthy lifestyle is necessary, and the worksite is considered to be a practical and effective setting for interventions. Moreover, to satisfactorily assess the success of programs several theories suggest that quite a few levels of measurement, comprising a large set of indicators should to be taken into account.

This paper therefore aims to gain insight into the current state of knowledge and practice, regarding best practice methods of monitoring and evaluation (M&E) of WHPPs that address healthy diets and physical activity. The specific objectives are to:

- Strategically collect and synthesise information from existing reviews of peer-reviewed literature of evaluations of WHPPs, complemented with 'grey' literature (e.g. case studies private sector) on recent evaluations of and initiatives in worksite health promotion.
- Give an overview of favourable and frequently used measures/indicators for process, intermediate and health outcome in evaluations of WHPPs.
- Comment on the implications and the feasibility of using these commonly used methods for M&E in non-scientific WHPP's in 'real world' settings.
- Formulate recommendations for non-scientific M&E of WHPPs which feasible in worksite settings.

3. Methods

3.1 Types of studies

This review aims to include all relevant reviews/publications/reports of (quasi)-experiments ((Randomised) Controlled Trials; RCT) regarding chronic disease prevention programs in the workplace, worksite intervention programs or worksite health promotion programs designed to promote physical activity and healthy diet. In this paper all programs will be referred to as WHPPs. The search was limited to systematic literature reviews concerning WHPPs.

3.2 Participants

Only studies focusing on a healthy working population (white-collar and blue-collar workers) were included. Studies investigating a population with pre-determined (chronic) work-related disorders were excluded.

3.3 Intervention content

The intervention had to promote physical activity [PA] (during leisure time or at/around work) and a healthy diet (e.g. fruit, vegetable, fat, fibre intake and related dietary intake) in a worksite setting. The intervention may consist of several components; a mix of education, counselling, incentives, multi-media approaches to raise awareness, changes in the worksite environment (and surroundings). Interventions wherein activities were focused solely outside of the worksite are excluded.

3.4 Outcome measures

The primary outcome must have included changes in self-reported or objective measurements (or determinants) of PA and/or dietary intake. Health and work related outcomes may include measurements of biological cardiovascular risk indicators (body mass index [BMI], blood pressure, body fat, cholesterol [blood lipids], waist circumference etc), sick leave, general health, productivity or work satisfaction.

3.5 Exclusion criteria

- Interventions focused on secondary or tertiary prevention: for example, if a physical activity intervention is focused on a previously existing disorder
- The intervention is aimed at a population with physical disorders/disabilities/diseases (cardiovascular diseases, cancer, type II diabetes), children/adolescents (0-16 years of age) or elderly (> 65 years of age).

3.6 Search strategy

Searches for literature were performed for peer- reviewed articles/ papers; the search strategy was divided into three stages:

1. Reviews published in the last 10 years (between 1997 and July 2007): The following keywords (MeSH and text words) were used: health promotion, intervention, workplace, worksite, work site, work floor, diet, food habit, fruit, vegetable, fat, fibre intake and physical activity.
2. Additional sources of information (examples from case studies shown in appendix 1). This included a web search (Google) for keywords: workplace, worksite, employee, corporate---wellness, physical activity, exercise, health promotion, diet, nutrition. Several websites providing information regarding awards granted to companies for superior workplace wellness programs, specific company websites, news articles, corporate statements, etc. were examined to gather information on functioning examples of WHPPs in the private sector.

The literature search was performed primarily in the Medline database (Entrez Pubmed). First the titles and abstracts of potential papers were checked if it concerned a WHPP to promote healthy diet and/or PA. If abstracts were not available or if titles/abstract were unable to provide the required information, the review was retrieved to screen the full text.

3.7 Case reports

Case Reports were gathered using web searches aimed towards identifying examples of workplace wellness programs in the private sector. Keyword searches included: workplace wellness, worksite wellness, worker wellness, employee health promotion, employee physical activity, employee nutrition, etc. A number of case studies were drawn from reports listed online by the C. Everett Koop National Health Award winners (granted by The Health Project (THP)), from the WEF "Working Towards Wellness" report, and additional information was gathered from specific company web pages.

4. Included reviews

The literature search produced 10 reviews in which the effectiveness of (mostly) worksite health promotion programs (WHPPs) on physical activity and/or food habits (nutrition) was evaluated. Reviews were published between 1997 and 2006. Three reviews (~57 worksite studies) specifically focused on the effectiveness of worksite PA interventions (Dishman et al., 1998; Proper et al., 2003 & 2006a), five reviews included more comprehensive WHPPs (~115 studies); worksite programs aimed at stimulating both PA and/or a healthy diet (Glanz et al., 1996; Heaney et al., 1998; Janer et al., 2002; Engbers et al., 2005; Proper et al., 2006b). Two reviews (18

studies) on worksite environmental/ policy intervention programs (Seymour et al., 2004; Matson-Koffman et al., 2005). Overlap between studies in each category may exist.

Four reviews (Dunn et al., 1998; Kahn et al., 2002; Eakin, 2007; Vandelanotte, 2007) were excluded because their general focus was on wide-ranging PA interventions (community and/or school based) and therefore did not provide sufficient information within the scope of this overview.

4.2 Study designs of the included studies

Generally speaking what can be concluded from the studies in the abovementioned reviews is that for the M&E of WHPPs (randomised) controlled experimental designs are used, including one or more intervention worksites compared to one or more (no or minimal intervention) control worksites. Randomisation of allocation of treatment or control condition is usually performed at the level the worksite in multi-centre trials. In addition, quite a few studies used an uncontrolled pre-post test design to evaluate the effectiveness of WHPP. Although a practical and feasible design for M&E of WHPPs in practice, this design presents serious limitation concerning the interpretation of the findings of a study (i.e. it is not possible to rule out secular trends in populations). Although multicentre randomised research designs (and a large number of participants) might really improve the ‘power’ of a study, the inclusion of several worksites to serve as control or intervention worksites is in most cases not possible. Second, all sorts of random or systematic bias by means of randomisation or matching of worksites can more easily be prevented. As said for M&E of WHPPs in practice this type of powerful randomised research design is not always a feasible option. However a high level of methodological rigour has to be pursued at all time to adequately interpret the success or failure of WHPPs, especially when using a weaker research design

In the studies in the included reviews, in general the evaluation period (time between the baseline and the follow-up measurements) in both intervention and control worksites is about 3-6 months for the interim follow-up and 12 to months for secondary follow-up. This evaluation period can be considered sufficient to determine changes in most intermediate and energy balance related behavioural outcomes in WHPPs.

The following chapter will give a description of two major categories of intervention content (behavioural/social and environment/policy approaches) as described in the included reviews.

5. Worksite interventions: a short overview

5.1 Behavioural/social approaches

The majority of the WHP-interventions described in the included reviews, are comprehensive multicomponent interventions ((Proper et al., 2006a,b; Janer et al., 2002; Engbers et al., 2005; Heaney et al., 1998; Glanz et al., 1996) focusing on increasing general health status for the prevention of overweight, cardiovascular diseases or cancer. Besides stimulating PA and/or a healthier diet, the intervention goals and intervention components in these studies were broadly focused on risk reduction in general for example; smoking cessation, stress reduction and/or mental health. Three other reviews described worksite interventions (Dishman et al., 1998;

Proper et al., 2003 & 2006a) specifically aiming at increasing PA at or outside the worksite. In general the majority of interventions can be characterised by a combination of:

- Education or counselling (individual or group sessions; personal diaries; computer tailored intervention or interactive websites)
- Mass media/informational strategies (food labelling, brochures, leaflets, access to websites)
- Group or individual based exercise/fitness sessions to increase cardiac fitness and/or strength at the worksite or outside.
- Some interventions also included environmental/social supports (increased availability or access to healthy options)

5.2 Environmental/ policy strategies

A review by Engbers et al. (2005) specifically aimed at assessing the effectiveness of multicomponent WHPPs including environmental modifications in the intervention. In this review thirteen (randomized) controlled trials were included. All studies focused on facilitating a healthy diet, but only three studies also aimed to stimulate PA. In these three studies PA was facilitated by constructing walking tracks around the worksite, providing new or upgrading existing fitness facilities and by making the staircases more attractive. Significant effects on self-reported exercise were found in two of the three studies.

A similar study but broader in scope, was conducted by Matson-Koffman et al. (2007). These authors performed a site-specific literature review of policy and environmental interventions that promote PA and nutrition. Data was synthesized by topic, type of intervention and setting. Consequently, this review also included a section pertaining to the effectiveness of worksite PA programs. All existing studies (i.e. non-, quasi- and full experiments [RCT's]) with an environmental intervention were included. Based on this review a number of effective (worksite) environmental approaches were identified:

- Point-of-decision prompts (motivational texts) to increase stair use, which included signs, banners, music and food (5 studies), these interventions components led to an increase in at least one worksite stair use study.
- Enhanced access to places and opportunities for PA, including onsite shower and locker room facilities, walking routes/groups, fitness facilities (6 studies) these components led to increases in the number of employees walking or biking to work.
- Comprehensive worksite approaches: employee and peer support, incentives and access to exercise facilities (5 studies).

In the reviews by Matson-Koffman et al. (2006) and Engbers et al. (2005) several (R)CT's focusing on WHPPs were also identified which used environmental modifications (among other methods) to stimulate a healthy diet. Another review with similar purpose, by Seymour et al. (2004), was identified which had the purpose of reviewing the effectiveness of worksite interventions that aimed to stimulate a healthy diet using exclusively environmental strategies. In this review another 10 (R)CT's covering worksite based interventions were identified; three studies used information strategies only, one study labelled the energy content of food items and another listed low fat items on a sign in front of the cafeteria and placed a 'heart' symbol next to each low fat entrée on the serving line. Some overlap in studies in the above-mentioned reviews exists, however, based on the significant effects on food intake in

the desired direction in most of the studies reviewed (in all reviews), the following promising environmental modifications could be identified:

- Increased availability of health products in company canteens and vending machines.
- Pricing strategies: reducing the prices of healthy products while increasing the prices of unhealthy products.
- Information strategies: cards or food labelling (i.e. point of decision prompts) displaying caloric content and/or coloured stickers (e.g. red and green) distinguishing between healthy or unhealthy products.
- Marketing strategies: food placement, for example by placing healthy products more in sight.
- Provision of incentives; number of healthy (labelled) products purchased can result in awards or discounts.

6. Indicators for success of WHPPs

6.1 Self-reported dietary intake

As could be concluded from the studies presented in (Table 1) the included systematic literature reviews a large variety of predominantly self-reported outcome measures which pertain to one or more aspects of dietary intake were used to assess the success of mostly comprehensive multi-component WHPPs.

Table 1. Included reviews with self-reported dietary outcome measures

Author	Included WHPPs with dietary outcome
Proper et al. (2006b):	14 (R)CT's
Janer et al. (2002);	7 (R)CT's
Engbers et al. (2005) :	13 (R)CT's
Heaney et al. (1998)	8 (R)CT's & pre-post test studies
Glanz et al. (1996)	12 (R)CT's

Depending on the intervention goals and components; the following outcome indicators to measure a change in the frequency of consumption (mostly fruit, vegetable and fat intake) and/or the related caloric content of certain pre-selected food items were described in the included reviews (Table 1).

Table 2. Outcome variables used in the included studies of the reviews

Outcome variable
<ul style="list-style-type: none">• The number of portions per day or per week (fruit/ juice, vegetables)• Energy% obtained from fat in products• (Saturated) fat intake (gram/day; points per day); fat intake from separate groups of products e.g. spread fillings and butter; milk products, snacks etc)• Fibre intake (gram/day)• Food diaries (e.g. 7- day or 24-h recall); a specific description of all food items eating and/or nutrient intakes within a specific time frame• Dietary changes/ practices; unknown method• Eating patterns (frequency and choices)

In addition a few studies included nutritional knowledge (n = 5) as an outcome, but no specification of the content and how it was measured could be determined based on the descriptions provided in the reviews.

6.1.1. Reasons for concern and opportunities

Although widely used and perhaps often the only feasible outcome indicator for the evaluation of WHPPs on food intake, self-reported data on food habits are generally known to be susceptible to many forms of bias. Examples include unavoidable recall bias and bias due to social desirability. These forms of biases have been found to be even more pronounced in specific subgroups of people such as those who are obese as well as various ethnic, social-cultural and low-income groups. A large amount of detailed information (i.e. diet) can be difficult to recall for anybody, and obese people (especially those who are on a diet) may for example be much more aware of what they eat than people with normal weight, but may not necessarily report it accurately. As most WHPPs focus on weight management, the measurement of overweight or obese subjects seems to be an important point of concern.

In a paper by Lissner et al. (2001), several studies are described that found that obese people systematically underreport their general food intake. Moreover, specific unfavourable food items (i.e. high in fat or sugar content) are generally underreported

by obese but also subjects of healthy weight. Validation studies concluded that a Dutch short-form fat, fruit and vegetable questionnaire is both valid and reliable (Van Assema, 2001 & 2002) and at least comparable to other generally accepted food frequency questionnaires. However, these validation studies concluded that *even* healthy subjects systematically overestimate fruit and vegetable intake and underestimate fat intake. In consequence, when measuring the food intake of obese subjects or other specific subgroups at risk, who are quite often the target group of WHPPs, general and food-specific underreporting (i.e. selective omissions) has to be taken into account.

As previously described, it is difficult to assess general food habits. However, quite a number of studies focusing on WHP included in various literature reviews were in fact able to detect significant changes in fruit-vegetable, fat and fibre intake using food frequency questionnaires. These food-specific intake measures are also the most common tools used to assess food habits. Quite a number of studies in these reviews used the 'Block' diet history questionnaire (Block, 1990) to measure changes in food patterns. As most food frequency questionnaires, the 'Block' is an extensive food-frequency questionnaire (36 pages), which also includes questions on foods obtained from six types of restaurants. The inclusion of these specific items might make the questionnaire more sensitive to detect changes. Similar to the majority of food-frequency questionnaires, the considerable length (number of items) of the 'Block' is a serious concern and may threaten the accuracy of participant responses (McClelland et al., 2001). This use of extensive and lengthy questionnaires might be less suitable for use in M&E of WHPPs in practice. The response issue especially becomes a concern when, in addition to food habits, other self-report outcomes are also assessed, as is the case in most studies in the included reviews. For this reason, an increasing number of shorter (electronic) but valid and reliable food frequency questionnaires or food screeners are currently being developed (Rocket et al., 2003; Vandelanotte et al., 2004; Block et al., 2000 www.nutritionquest.com). Other examples are the Dutch food frequency questionnaire on fat and/or fruit and vegetable intake (Van Assema et al., 2001 & 2002) or the rapid food screener (Block et al., 2000). The Dutch Questionnaire has the advantage that it is relatively short in length (4 pages). However, in contrast to the above-mentioned Block-questionnaire, this Dutch questionnaire does not include items about foods consumed in (worksite) restaurants. The rapid food screener is a one page instrument which is able to quickly monitor a person's diet by identifying high fat and low fruit and vegetable intake. Shorter food-frequency questionnaires which are able to give valid insight into a person's food intake are likely to increase the response rate and possibly the accuracy of responses.

More importantly, short questionnaires provide (less expensive) opportunities to offer questionnaires over the internet or via a company's intranet. Internet questionnaires reduce man power related to data management and may also increase the response rate. Since subjects can be effectively routed through questionnaires because they are able to skip questions which are not applicable to them, the questionnaire effectively becomes shorter. Moreover, these internet based questionnaires can be incorporated into web or email based intervention strategies, and these relatively inexpensive/cost-effective interventions are increasingly rapidly in popularity within WHPPs.

Another useful initiative, which is currently being developed in the Netherlands, is a computer-based tool used to generate and process food frequency questionnaires in order to provide valid assessments of dietary intake (De Vries et al, 2000). This tool

provides the opportunity to quickly renew and update questionnaires, e.g. to make them suitable for assessment of dietary components of novel interest or specifically matched to WHPP intervention component and objectives. Another promising initiative can be found in computer adaptive testing; these programs use statistical techniques (iterative algorithms and item response theory) on databases with questions that optimize the number of propositions presented to each respondent. In many cases sufficient information is obtained from the respondents by asking just one third of the questions (Green, 2000; Thissen & Mislevy, 2000). A last but important point of concern needs to be mentioned; food patterns (more than physical activity behaviour) are culturally defined. Since many questionnaires are designed in the US they might not be suitable for outcome measurements in working population in other parts of the world (e.g. Europe, Asia, etc.). For example, the Block questionnaire was specifically designed to measure only American food patterns.

6.2 Objective measures of diet

Actual general sales data in company restaurants, although scarcely used as outcome measures in the M&E of multi-component WHPPs, are quite often used for the evaluation of environmental (worksite) interventions. In a review by Seymour et al. (2004) of environmental nutrition interventions, ten worksite studies published between 1970 and 2003 were identified that used an objective method to assess dietary habits. In three other reviews only in four studies (Heaney et al., Glanz et al., 1996; Proper et al., 2006b) an objective instrument was identified. Most of these studies used sales data to assess the success of an intervention. Group level monthly changes in the consumption/purchases of specific healthy or unhealthy (e.g. high or low fat) food items as highlighted in the intervention were analysed. A number of uncontrolled studies found in the review by Seymour et al. (2004) also included the collection of data on changes in sales of healthier (reduced price) products in vending machines. Another rather intensive and time consuming method which is used in relatively few studies is the manual observation of purchased specific food items or the caloric content per purchased lunch tray. This method requires trained observers with specific food inventory forms, and therefore may not be a feasible method for evaluation of many non-scientific WHPPs.

To conclude, assessing objective food/ product sales data appears to be a feasible and practical method for M&E of WHPPs in practice. The use of this automatically generated data does not contribute to time invested in non-work related activities by employees and no burden of measurement is put upon them. Moreover, in an increasing number of company restaurants, chipcards or company access cards (with unique numbers) are used for payment of food purchase. These purchases are often electronically saved into a (cash register system) database; consequently data on an individual level might in theory be collected by the push of a button.

6.3 Self-reported PA outcome

Depending on the program content and focus, the measurement of PA is generally considered to be one of the most important outcome indicators for the success of WHPPs. In contrast to health related long-term outcomes (body composition, cholesterol levels, etc) a change in self-reported PA-behaviour can be achieved in a relatively shorter amount of time. Although quite a few of the studies in the included reviews of this paper had PA as an outcome, only three systematic literature reviews (Dishman et al. 1997, Proper et al., 2003 & 2006a) of PA interventions were screened on PA outcome measures. In Proper et al. (2006a) 11 studies, in Proper et al. (2003) 8

studies and in Dishman et al. (1998) 13 studies used one or more types of self-report PA assessment as an outcome indicator. A large variety of mostly self-reported outcome variables to measure a change in individual PA-behaviour were described in the included reviews (Table 3).

Table 3. PA related outcome variables as described in the reviews

Outcome variable
<ul style="list-style-type: none"> • 3-day activity diary to calculate energy expenditure • Energy expenditure; unknown method (kcal/ week) • Exercise habits (unknown method) • 1-week activity recall • Frequency of intensive exercise per week • Minutes of moderate intensity exercise per week • Leisure time physical activity • Proportion of subject participating in regular exercise • Amount of PA (MET/ minutes/ week) • Number of steps measured with a personal pedometer filled out in a questionnaire or personal website • Total minutes of PA per week • Worksite PA (<i>method not specified</i>) • A sport index (<i>method not specified</i>) • Proportion of subjects meeting the CDC/ ACSM PA-guideline (US surgeon general, 1996) • Minutes per week of inactivity; or the proportion of subjects inactive

In the review by Dishman et al. (1998) it was only reported that studies had used either validated self-report of PA or just self-report of PA, two studies registered attendance and three studies used activity logs to assess PA.

This wide variety of outcome variables which measure several dimensions of PA can, according to Bauman et al. (2006), be categorised using five components; frequency, duration, intensity, type and the domain or setting wherein PA was performed. Based on the general description provided in the reviews it cannot be concluded whether all of the dimensions of PA were measured, however, most available validated questionnaires assess at least one of the aspects described in Table 4.

Table 4. Categories of PA dimensions used in most available PA questionnaires.

Component	Description
Frequency of PA	Recall of a certain time frame; last week or month, or a subject is asked to take in mind a usual week
Duration	Duration of PA per session (total time per day) in hours or minutes or an average of a certain time frame.
Intensity	Based or estimated on self-perceived intensity or energy expenditures (MET) associated with certain activities.
Type of activity	A subject is asked about what specific PA or sport he or she performed; or the subject is inquired about PA in broader categories such as moderate and vigorous activities (complemented with examples of PA's within these categories).
Domains or settings of PA	Description of the location where the PA was carried out: <ul style="list-style-type: none"> - organised or non- organised leisure time PA (LTPA) - Occupational PA (at work) - Domestic setting (at home)

- Active commuting
- Incidental PA (stair use, elevator use)
- Sedentary behaviours (time spent sitting at work, watching TV, computer use etc)

LTPA = Leisure Time Physical Activity; PA = Physical Activity; MET = Metabolic Equivalent

6.3.1. Reasons for concern and opportunities

Like the measurement of dietary habits, the measurement of something as complex as PA by means of self-report is extremely difficult and remains a challenge in health research. Similar problems such as cultural differences, recall bias and social desirability, as described in the paragraph addressing the measurement of dietary habits, are also to be expected in the measurement of PA. Consequently, to make adequate assumptions about the effectiveness of WHPP, it is important to at least select standardized instruments that are tested on validity and reliability. However, based on the general descriptions of the outcome indicators used in the included reviews, the specific PA questionnaire being used could not always be determined, nor could it be assessed whether or not these questionnaires were validated. To facilitate the selection of PA-instruments in a paper of Rennie and Wareham (1998) a checklist is given. These items are quite helpful in selecting suitable validated physical activity instruments for M&E of WHPPs (Table 5). These authors state that “the claim that a particular method has been ‘validated’ ought really to be questioned by anyone proposing to use that method”.

Table 5. Checklist for the validation or selection of physical activity instruments (according to Rennie & Wareham, 1998)

Checklist items
1 Has the dimension of physical activity that the instrument is purported to measure been clearly defined?
2 Does the validation method chosen measure the true exposure of interest and has it been applied in the same time frame of reference?
3 Has correlated error between the validation method and the physical activity instrument been avoided as far as possible?
4 Is there a close relationship between the validation method and the appropriate ‘gold standard’ instrument?
5 Is the sample chosen representative of the population to whom the physical activity instrument will be administered?
6 Have appropriate statistical techniques been employed to assess the validity of the physical activity instrument?

In most studies found within the included reviews a change in general PA-behaviour outside the workplace was assessed. In concurrence with the idea that the selection of outcome indicators should always directly be related to program objectives, it is remarkable that these items are barely measured. Moreover to quantify the effect of a worksite PA program, it seems only reasonable to measure PA at the worksite as well, because an increase in short bouts of (occupational) PA accumulated over longer periods of time can increase energy expenditure and might contribute to the prevention of overweight (Dunn et al., 1998) and may also serve as an intermediate to an increase in overall PA. Therefore, it would seem beneficial to include measurements of items such as occupational walking and stair climbing. Nevertheless, a major problem with most self-report instruments is that these spontaneous, habitual activities are mostly unconscious, and for that reason difficult to

recall for the majority of employees. To date, few or none of the existing self-report measures can adequately capture the type of low intensity PA which is characteristic of sedentary (worksite) populations. In this regard, 'floor effects' might be a common threat to the validity of self-report methods; wherein the lowest scores or answer categories available in a questionnaire are too high for some respondents. Due to these questionnaire limitations many participants may be forced to report incorrectly by reporting higher scores on low-intensity activities that may not represent their true activity levels. Most questionnaires only inquire about brisk or powerful walking, but this type of PA is captured least reliably by recall (Kriska et al., 1990; Richardson et al., 1993; Ainsworth et al., 1993), and does infrequently occur at the worksite. In addition, questions regarding walking speed and intensity are often poorly understood by respondents. As will be explained in the next paragraph, the use of pedometers or accelerometers can support self-report methods in this regard, by objectively measuring walking at the workplace (Tudor-Locke, 2001).

As stated the measurement of occupational PA, though important, is not included in most questionnaires. In a special supplement to the Journal of the American College of Sports Medicine (Jacobs et al., 1992), and also in a paper by Bauman et al. (2006) a list of (validated) questionnaires which can be used to assess (changes in) PA behaviour within intervention studies is provided. Nevertheless, based on the descriptions of the questionnaire content, few of the questionnaires (in Bauman et al., just three questionnaires) seem to address occupational activities. Examples of instruments that included items on occupational PA are: the Paffenberger Physical Activity Questionnaire (Ainsworth et al., 1993), Seven-day Physical activity Questionnaire (PAR) (Taylor et al., 1984) and the Scottish Physical Activity questionnaire (SPAQ) (Lowther et al., 1999). Another recently developed and useful and short PA-questionnaire is the Short QUEStionnaire to ASsess Health-enhancing physical activity (SQUASH) which assesses worksite PA (WPA) (i.e., occasional walking, lifting objects and stair climbing) as well as total PA (i.e., household, work, commuting and leisure time) (Wendel-Vos et al., 2003).

6.4 Objective PA measurement

6.4.1 Stair use measurement

As mentioned in a review by Matson-Koffman et al. (2005) the potential to increase stair use by means of the environment seems to be promising. Adequate measurement of stair use on the other hand is quite difficult. Merely one out of seven worksite studies used measurement of stair use as an indicator for PA at the worksite. Other identified stair use studies in this review (n = 7) were community based; and though these studies showed significant increases in stair use, they were conducted in public settings (shopping malls, train stations, etc). These mostly uncontrolled pre post-test studies used methods as:

1. *Observation*; head counting of persons entering or exiting a stair case by means of observers or by placing cameras.
2. (*Infra red*) *counting devices*, placing devices at the entrance of staircases.

Infrared measurement and in particular stair use observation is relatively expensive and time consuming, therefore the use of self-reported stair use might be more feasible for use in determining the effectiveness of WHPPs on stair use.

To gain a first insight into the use of self-reported stair use, in a study by Engbers et al. (2007a) stair use was measured in two office buildings as a part of a larger trial (Engbers et al., 2007b) using an objective measuring system and as well as just two questions as part of a larger questionnaire. The objective system consisted of registration devices (RFID technology) placed on every floor directly behind the door to the staircase. In order to register stair use special chip cards with a unique code had to be carried around by the subjects. This way the direction, number of floors and time between floors could be registered. With each outcome measure the number of stair use flights per day and number of floors per day could be determined at the level of the individual. Results of this study showed that self-reported stair use was systematically twice as high as the objective measurement, which might indicate recall bias or socially desirable answers. Nevertheless, the correlations (ICC's) between objective and self-reported stair use found in this study could be interpreted as acceptable. The lack of a previously validated gold standard for measuring stair use makes it difficult to draw solid conclusions about the validity of self-reported stair use, but for use in the evaluation of WHPPs in practice it might be an acceptable and feasible option for quantifying stair use at worksites. Based on the significant lower stair use in one worksite compared to the other, this study revealed some important insights in the how differences in architectural design and elevator location between the included worksites might have had a strong influence in stair use. Therefore, it might be argued that the architecture at worksites influenced sedentary employee behaviour.

In a review by Eves et al. (2006) a few recommendations are made regarding the measurement of stair use in worksite settings. In this review six environmental worksite stair use studies (mostly no control pre-posttest studies) were identified. Based on these studies it was concluded, that if stair use is measured in worksite settings, that in the analysis a clear distinction between stair climbing and descent has to be made. Stair climbing uses three times the energy as does the ascent, 8.6-9.6 opposed to 2.9-3.2 METS (Metabolic Equivalents) (Ainsworth et al., 2000) respectively. Eves states that only when individuals are increasing their stair climbing (number of flights/ day) they are able to meet the standard CDC/ ACSM PA-recommendation. This recommendation is the equivalent to 40 minutes of stair *descent* a day, which seems impossible in a worksite setting. Moreover, several authors also mention other issues such as the number of stories in a building, building design and the location of stairs and elevators. These factors all need to be considered when measuring stair use (Eves et al., 2006; Engbers et al., 2006). Research indicates that when the number of stories in a building increases, then stair use will decrease, and about four stories is the maximum that subjects are willing to climb (Kerr et al., 2002).

A last point worth mentioning is that unlike the abovementioned studies only controlled trials provide opportunities to interpret changes in stair use. In the literature no randomized controlled study on stair use in a workplace setting can be found to date. Randomised controlled stair use interventions should be explored future research in WHP.

6.4.2 Pedometers and accelerometers

As previously mentioned, spontaneous low intensity activities such as occupational walking are least reliably captured by self-report. However, pedometers and the more expensive accelerometers offer good opportunities to fill this gap.

Accelerometers measure (vertical) motion, and also record the time and intensity of movement (Welk, 2002). Pedometers count only the number of steps taken by subjects, which in turn have to be reported back to the researchers. Nevertheless, in the outcome assessment of studies in the included reviews, pedometers and accelerometers are barely used. The small use of monitors such as accelerometers in the measurement of WHPPs might be attributed to the following explanations:

- Accelerometers are relatively costly (€100- €350); therefore they cannot be used to measure a large study population. A subgroup of subjects has to be determined and this might be a threat to the generalisability of the data obtained from these subjects.
- The use of accelerometers requires a great deal of logistics and organisation to give them away and retrieve them again from subjects. This is often considered too time consuming within the short time frame in which the implementation and evaluation of a WHPP has to take place.
- Analysis and interpretation of the data obtained with accelerometer is also extremely time consuming.

Mechanical pedometers on the other hand are easily manipulated, though pedometers (and accelerometers) are only able to adequately capture walking, whereas activities such as biking or stair climbing are not satisfactorily captured. Using pedometers requires less logistic and organisational burden to researchers, and analysis (of number of steps) is less complicated, however, in the recent past support for the use of relatively cheap pedometers for research has been abandoned due to a large amount of criticism (poor accuracy and reliability) by researchers. Though, Tudor-Locke (2001) asserts that newer more sophisticated electronic pedometers are currently available, which are capable to more accurately record walking-related activities (Basset et al., 1996, Slootmaker et al., submitted). Research indicates that only three days of wearing a pedometer is sufficient to estimate a person's walking behaviour, which represents only a small burden for the respondent (Tudor-Locke et al., 2005).

Accelerometers have also recently improved to more adequately capture walking as well as other kinds of PA such as biking and stair-climbing. Research and development initiatives have produced an increasing number of more sophisticated and relatively cheap models of uni-axial and even tri-axial accelerometer models. It is hypothesised that the latter would be more sensitive for PA other than walking, due to the movement registration in both vertical and horizontal directions (Rowlands et al., 2004; Levine et al., 2001).

To conclude, the contribution of pedometers and/or accelerometers to facilitate the accurate measurement of the success of a WHPP might have implications for general occupational settings. The measurement of occupational walking (or even stair climbing) by means of accelerometers and/or pedometers is a feasible option which needs to be more frequently used in the evaluation of WHPPs.

6.5 Measuring the environment

Given that environmental influences are increasingly and widely recognised as important, if not the main contributors to weight gain (French et al., 2001; Glanz et al., 2005; Sallis, 2002), it seems viable that the measurement (of changes) of the worksite environment should become a standard outcome indicator in the evaluation of WHPPs. However, none of the studies found within the included reviews used such a measure, mainly because few of these kinds of measures exist to date. Three relatively recent measures are available:

- Checklist of Health Promotion Environments at Worksites (CHEW)

- Worksite Environment Measure (WEM)
- Nutrition Environment Measures Study Restaurant assessment (NEMS-R)

The CHEW (Oldenburg et al., 2002) was developed to ‘objectively’ observe and check the worksite environmental influences that relate to a wide range of health-related behaviours. The CHEW includes a broad range of behaviours (healthy eating, physical activity, alcohol consumption, and smoking) and has high inter-rater reliability. However, the WEM was developed in response to the unavailability of the CHEW’s sensitivity to intervention effects and the associations between the CHEW objective environmental variables and individual-level perceived environmental variables (Shimotsu et al., 2007). The WEM aims to objectively assess the number, condition and presence/absence of specific items of food, physical activity, weight management, media and social environment. However, this instrument was designed and tested for the assessment of the inside and outside environment of an automobile garage, so it might need to be altered to fit the specific environments of other workplaces. The WEM was found to be a reliable instrument with which to measure the environment at worksites. It also provides an opportunity to assess the perceived and social environment from the employee perspective, and can be utilized as a process measure, which is important when evaluating WHPPs.

A third (and most recently developed) instrument used to measure environments is the NEMS-R. This instrument measures factors which are believed to contribute to food choices in restaurants, including availability of more healthy foods, facilitators and barriers to healthful eating, pricing, and signage/promotion of healthy and unhealthy foods (Sealens et al., 2007). Although this instrument had very good results on intra-, inter-, and test-retest observer reliability, its use as an indicator for success in the evaluation of WHPPs has yet to be demonstrated.

The abovementioned instruments collect data on the environment inside the worksite; however the direct surroundings of a worksite (i.e. walkability, availability of shops, parks and/or green spaces, etc.) might also play an important role in determining the dietary or PA behaviour of employees. In an (unpublished) study on lunch-time walking among employees in the Netherlands, it appeared that a forced change (due to a move) of inside and outside environment of a worksite significantly reduced the proportion of employees that walked during lunch time (see also: Hendriksen, 2000; or www.lunchwandelen.nl/contact.html). The inside and outside environment in this study was assessed using the ANGELO-model (Analysis Grid for Environments Linked to Obesity) developed by Swinburn et al. (1999). ANGELO is a matrix with two axes; on one axis the size of the environment is shown from micro (e.g. places where people meet and interact, like workplaces) to macro (e.g. more anonymous places like municipalities or even countries) environment, and on the other axis four types of environments (physical, economic, political and social-cultural) are shown (see: Figure 2).

	Micro-environment	Macro-environment
Physical environment <i>(Characteristics of the environment to stimulate or reduce healthy behaviour; availability of fruits/vegetables or walk- bike tracks)</i>		
Economic environment <i>(The costs and consequently the accessibility of healthy choices).</i>		

Political environment (Laws and legislation influencing healthy behaviour; e.g. smoking laws, company lease bikes)		
Socio-cultural environment (Social and cultural norms, support and pressure on healthy behaviour)		

Figure 2. The ANGELO grid (based on Swinburn and Egger., 1999)

The ANGELO grid can help to outline and determine gaps in the large variety of environmental determinants useful for or addressed in intervention studies. This example on lunchtime walking provides an idea of how an instrument used for the measurement of the environment can play a part in M&E of WHPPs.

To conclude, it is highly recommended to increase the use and development of tools that reliably assess changes in the worksite environment, but also serve as indicators of the successfulness of a WHPP.

6.6 Determinants of PA or dietary behaviour

Only very few of the studies in the included reviews reported the use of some type of measurement of determinants of behaviour in the outcome assessment. The following outcome variables regarding the psychosocial determinants of behaviour are considered to be important as intermediate outcomes in the evaluations of WHPPs:

- Attitude, self-efficacy, and/or social support towards certain behaviours such as fruit, vegetable and fat consumption, and PA.
- Intention of a subject to change behaviour.
- Percentage of subjects transferring from pre-action (contemplation) phase to action/maintenance-phase (the stages of change model)
- Perceived benefits/barriers: what are the benefits or barriers with respect to eating less fat and more fruits and vegetables, and being more physically active?
- Awareness: is the subject aware of the intervention content and related objectives.

Most of the abovementioned variables are derived from the Attitude-Social influence-Self-efficacy (ASE) model (De Vries et al., 1988; Brug et al., 1995). This is a useful model that specifies intrapersonal factors and can be helpful in creating items in questionnaires for use in M&E of WHPPs. In this model, behaviour is determined by an individual's attitude towards a certain behavioural pattern, and also by social influences and self-efficacy. The first determinant in the ASE-model is 'attitude', which can be described as the general feeling of a person (good or bad, favourable/unfavourable) towards, for example, physical activity. This feeling is determined by the perceived positive or negative consequences of being physically active. Social influence, the second determinant, can be described as the influence and the expectations of significant others (friends, family, colleagues), but also as, for example, the level of physical activity of these significant others. Finally, self efficacy is the self-perception of an individual's capability to perform a certain activity. These three determinants predict the intention, which in its turn is believed to predict physical activity and dietary behaviour. Most of the outcome variables regarding determinants of behaviour are usually measured on Likert-scales. Separate items or groups of items referring to the abovementioned concepts are measured using a five or seven point Likert-scale. Each subject has to fill out to what degree he/she agrees with

a number of statements regarding eating less fat or more fruits and vegetables and/or being physically active.

An important determinant that is not usually mentioned in models explaining behaviour, nor used in M&E of WHPPs is habitual behaviour or habits. Habits as an intermediate to predict behaviour are not mentioned in the ASE-model. However, an understanding of habitual behaviour may give additional insight in how the worksite environment intervenes on the behaviour of the individual. Habits are the result of automated unconscious cognitive processes (Verplanken et al., 1998; Aarts et al., 1997). Habits are usually formed when certain behaviours are repeated enough. Thus, (new) habit formation is more likely to occur when the evaluation of a certain action was satisfactory to the individual. Habits are also strongly dependent on situational and environmental factors or obstacles. Therefore, it is hypothesized that if habits are a strong determinant for a certain behaviour, then the introduction of new motivational materials or cues in an existing environment, may change or momentarily interrupt routine or habitual behaviour (e.g., taking the elevator). If this interruption of a certain 'unwanted' unhealthy behaviour occurs on several occasions, a new healthy habit can be formed, under the assumption that this new behaviour was deemed satisfactory. In addition, information processing is unlikely to occur when behaviour is guided by strong habits. In other words, behavioural patterns that are determined by strong habits are less guided by cognitive/conscious processes. For that reason, in many WHPPs there is a shift from providing information (or counselling) to changing the environment in order to break down unhealthy habits and achieve significant behavioural changes among the target population. Assessing changes in (the perception of) habitual behaviour might be an important determinant to add as an indicator of success of WHPPs.

In conclusion, measuring the change on variables regarding behavioural determinants might indicate to what extent a subject is likely to perform and/or maintain a certain desired behaviour. Therefore the possible changes on the scores of determinants are only an intermediate indicator of the success of a program, but their measurement can help guide the development of future WHPPs.

6.7 Biological disease risk indicators

In the majority of studies reviewed, one or more biological disease risk indicator is assessed in order to measure the success of a WHPP. As can be expected, the included worksite PA intervention studies reviewed by Dishman et al. (1998) and Proper et al. (2006a) measured an estimate of VO₂max (e.g. walking and or bike tests) as a part of their outcome assessment in 11 and 4 studies respectively. However, in contrast to the review by Dishman (1998), the more recent review by Proper (2006a) reported that 8 RCT's also collected data on two or three other biological outcome indicators (e.g. body composition) as described in Table 6.

Table 6. Frequently used biological in outcome indicator in PA interventions

Outcome variable
Weight (kg) or body mass index (BMI: weight/ length ²)
Systolic and diastolic blood pressure (mmHg)
Cholesterol (total) and/or LDL/ HDL cholesterol (mmol/l)
Body composition: percentage of body fat or the sum of four skin folds thickness (mm)
Physical fitness; sub-maximal bike/ walking tests (VO ₂ max, sub maximal heart rate)
Muscle strength (leg or back muscles)

On the other hand, more comprehensive studies with broader intervention goals (health risk reduction) that were included in other reviews (see: Table 1) also collected data on one or occasionally two of the listed biological outcome indicators; though predominantly cholesterol complemented with measurements of blood pressure and/or BMI.

6.7.1 Reasons for concern and opportunities

The collection of data on biological indicators is usually very time consuming, expensive and requires a trained research assistant. Consequently, when the program becomes more comprehensive, the burden, in terms of outcome assessment, on the subjects increases significantly. In this case, well thought-out decisions regarding which outcome indicators to include or exclude have to be made. For M&E of most WHPPs in practice (as opposed to research), the measurement of a small number of biological indicators such as total cholesterol (finger stick cholesterol method), physical fitness (walking or bike test) and/or only body weight/height might be most feasible and sufficient. The measurement of these variables can be incorporated in an annual health check which is currently already provided by an increasing number of employers. In this way, longitudinal data are collected and guaranteed, and long-term trends can be observed.

However, based on the conclusions from the studies within all of the included reviews, convincing evidence for an effect of WHPPs on blood pressure, BMI or cholesterol have yet to be found. Mostly because of the poor methodological quality in many of the existing studies, the feasibility of a causal relationship between the WHPPs and these outcome indicators remains unanswered. Due to the poor methodological rigour (e.g. the absence of a controlled design), performing extensive physical measurements on subjects might also contribute to the poor success of WHPPs on biological indicators. In a study of Van Sluijs et al. (2006), it was concluded that performing a measurement (e.g., a physical examination) is an intervention in and of itself. Because of such measurements, changes in biological outcome variables especially in the desired direction might occur in both the intervention and the control condition, thus a possible effect is diluted.

The measurement of biological risk indicators is essential from a clinical perspective in order to gain insight into the direct physical health (changes) of employees. However, due to the cost of these measurements (i.e. personal and financial burden) as well as the limited success of WHPPs on these variables, well-thought decisions have to be made on which biological indicators (related to program objectives) to include in M&E of WHPPs in practice.

6.8 Process indicators

Though many process evaluations of WHPPs do not get published, not one of the studies in the reviews included in this paper reported the measurement of process indicators as part of their outcome evaluation, despite the fact that process evaluation is mentioned as an important part of the evaluation of WHPPs (Centre for Health Promotion, 2005; Bauman 2006, Nutbaum, 1998; Kwak, 2007). To our knowledge only two publications exist in the literature wherein process evaluations of WHPPs can be found (Steenhuis et al., 2004; Beresford et al., 2000). Beresford et al. (2000) evaluated the Seattle 5-a-Day Work-Site Project and compared aspects like program material delivery between small or large worksites and the influence of gender distributions in these companies on intervention success. Steenhuis et al. (2004)

conducted semi-structured interviews with 21 managers of supermarkets and worksite cafeterias where the programs were implemented. In this study two environmental programs in worksite cafeterias of seventeen worksites were evaluated. The main topic of the ‘after the intervention’ interviews in the study by Steenhuis were; general impression of the intervention, opinion on various program materials, difficulties with implementation, perceived benefits of the intervention, reaction of customers and perceived effects on buying and eating habits. In Table 7, important and relatively easy to monitor aspects of a process evaluation are described.

Table 7. Indicators to include in process evaluation of WHPPs

Process indicators
<ul style="list-style-type: none"> • Did the WHPP reach all of the target population • How were WHPP components and services provided • Was the program acceptable for the target population? • What problems were encountered in implementing the program? How were they resolved? • If objectives, plans, or timetables were revised, why was this necessary? • What costs (materials, measurements, management) were incurred? Did they exceed initial projections? • Were possible co-interventions (commercial health services, questionnaires) in progress at the worksite during the intervention? • How did the data-collection take place; how were the data collected?

Interviews and observations or questionnaires for stakeholders and participants to collect the abovementioned information usually takes place after the intervention but it is recommended to keep track of several indicators during both the preparation and implementation phase of the WHPP.

To conclude, the collection of quantitative but mostly qualitative information on program implementation and delivery is just as important as the M&E of impact of WHPPs. The findings of process evaluations ought to be published more frequently because it provides meaningful information on the interpretation of the findings (success or failure) in M&E of WHPPs.

7. Best Practice: Recommendations for M&E of WHPPs

1. To adequately interpret the findings of M&E of WHPPs, outcome indicators should be included on each level of formative, process, intermediate and impact evaluation.
2. Outcome indicators should directly be related to and dependent on intervention components and objectives. They should be logical consequence of decisions made in each level of evaluation as mentioned in point 1.
3. An extensive process evaluation should always be included in M&E of WHPPs; qualitative information on program preparation and implementation is just as important as the impact of WHPPs. Process evaluations will provide useful information on how to make WHPPs more successful.
4. Due to an increasing body of evidence regarding environmental determinants and the effectiveness of such interventions, in M&E of WHPPs at least a measure to quantify the inside (worksite) or outside environment has to be included. An increasing number of such instruments are currently developed for research purposes.
5. To decrease the load of data management for health researchers and to increase the response of subjects to PA or dietary questionnaires, the use of validated and shorter *internet* or *intranet* questionnaires should progressively be used more in M&E of WHPPs.
6. Questionnaires to quantify PA should include items on occupational PA.
7. Technical developments provide new, valid and more feasible means to objectively measure occupational light-intensity PA (tri-axial accelerometers, pedometers and stair use measurement) and dietary habits (specified sales data in restaurants and vending machines). Consequently, the use innovative objective instruments should be stimulated and utilised more frequently in M&E of WHPPs.
8. The inclusion of an extensive set of biological indicators, like the measurement of full spectrum blood lipid analysis, determination of body composition and blood pressure, in WHPPs might not be necessary or feasible. In correspondence with program components and objectives a relatively small set of feasible and less expensive biological indicators (i.e. finger stick cholesterol, fitness or strength tests and/or body weight and length) might be sufficient in M&E of WHPPs in practice.
9. In combination with the continuous monitoring of sick leave at most worksites, regular (yearly) health check-ups of employees should likewise be incorporated in company health policy. Including the recommended small set of biological indicators and/or a questionnaire, continuous health monitoring:
 - Will give insight into long term in health changes
 - Will automatically measure health changes due to newly implemented policy or intervention elements.

- Will lead to effective WHP by gaining insight into successful or unsuccessful interventions of policy elements so that health trends can be anticipated.
- Can be utilised as a bench-mark of company's health policy
- Will contribute to the employee's perception of the commitment of the company to occupational health management.
- Continuous data flow can make cost-benefit analysis achievable.

8. References

- Aarts H, Paulussen T, Schaalma H. Physical exercise habit: on the conceptualization and formation of habitual health behaviours. *Health Education Research*. 1997;12(3): 363-74
- Ainsworth B., Haskell, White C. et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med. Sci. Sports Exerc*. 32(suppl.):S498–S504, 2000.
- Ainsworth BE, Leon AS, Richardson MT, et al.. Accuracy of the college alumnus physical activity questionnaire. *Journal of Clinical Epidemiology* 1993; 46: 1403-11. In : Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine*. 2001 Feb;31(2):91-100.
- Ainsworth BE, Leon AS, Richardson MT, Jacobs DR, Paffenbarger RS Jr. Accuracy of the College Alumnus Physical Activity Questionnaire. *Journal of Clinical Epidemiology*. 1993 Dec; 46(12):1403-11.
- Asfaw A. The effects of obesity on doctor-diagnosed chronic diseases in Africa: empirical results from Senegal and South Africa. *J Public Health Policy*. 2006;27(3):250-64.
- Bauman A, Phongsavan P, Schoeppe S, Owen N. Physical activity measurement--a primer for health promotion. *Promot Educ*. 2006;13(2):92-103.
- Beresford SA, Shannon J, McLerran D, Thompson B. Seattle 5-a-Day Work-Site Project: process evaluation. *Health Education and Behavior*. 2000 Apr;27(2):213-22.
- Blair S, Connelly J, How much physical activity should we do? The case for moderate amount and intensities of physical activity. *RQES* 1996; 67: 193-205.
- Block G, Gillespie C, Rosenbaum EH, Jenson C. A rapid food screener to assess fat and fruit and vegetable intake. *American Journal Preventive Medicine*. 2000 May;18(4):284-8.
- Block G, Woods M, Potosky A, Clifford C. Validation of a self-administered diet history questionnaire using multiple diet records. *Journal of Clinical Epidemiology* 1990; 43(12): 1327-1335.
- Brug J, Lechner L, De Vries, H. Psychosocial determinants fruit and vegetable consumption. *Appetite* 1995; 25, 285-296.
- Centre for Health Promotion; University of Canada. Evaluating Comprehensive Workplace Health Promotion, Version 1. March 15, 2005
- De Bourdeaudhuij I, Stevens V, Vandelanotte C, Brug J. Evaluation of an interactive computer-tailored nutrition intervention in a real-life setting. *Annals of Behavioral Medicine*. 2007 Feb;33(1):39-48.
- De Vries JHM, Dagnelie PC, Jansen MCJF, Ocke MC, V.d. Brandt PA, Van 't Veer P. A computer-based tool to generate and process food frequency questionnaires for valid assessment of dietary intake in the Netherlands. Local project, Wageningen University, The N
- De Vries H, Dijkstra M, Kuhlman P. Self-efficacy: The third factor besides attitude and subjective norm as a predictor of behavioral intentions. *Health Education Research*. 1988; 3: 273-282.
- Dishman RK, Oldenburg B, O'Neal H, Shephard RJ. Worksite physical activity interventions. *American Journal Preventive Medicine*. 1998 Nov;15(4):344-61.
- Dunn AL, Andersen RE, Jakicic JM. Lifestyle physical activity interventions. History, short- and long-term effects, and recommendations. *American Journal Preventive Medicine*. 1998 Nov;15(4):398-412.

- Eakin EG, Lawler SP, Vandelanotte C, Owen N. Telephone interventions for physical activity and dietary behavior change: a systematic review. *American Journal Preventive Medicine*. 2007 May;32(5):419-34.
- Engbers LH, van Poppel MN, van Mechelen W. Measuring stair use in two office buildings: a comparison between an objective and a self-reported method. *Scandinavian Journal of Medicine Science and Sports*. 2007a Apr;17(2):165-71.
- Engbers LH, van Poppel MN, van Mechelen W. Modest effects of a controlled worksite environmental intervention on cardiovascular risk in office workers. *Preventive Medicine*. 2007b Apr;44(4):356-62.
- Engbers LH, van Poppel MN, Chin A Paw MJ, van Mechelen W. Worksite health promotion programs with environmental changes: a systematic review. *American Journal Preventive Medicine*. 2005 Jul;29(1):61-70.
- EUROSTAT, European Opinion Research Group. Physical Activity. Special Eurobarometer 183-6/ Wave 58.2. Brussel: Europese Commissie, 2003 (http://europa.eu.int/comm/public_opinion/archives/ebs/ebs_183_6_en.pdf.)
- Eves FF, Webb OJ. Worksite interventions to increase stair climbing; reasons for caution. *Preventive Medicine*. 2006 Jul;43(1):4-7.
- French SA, Story M, Jeffery RW. Environmental influences on eating and physical activity. *Annual Review Public Health*. 2001; 22:309-35.
- Glanz K, Seewald-Klein T. Nutrition at the worksite: an overview. *Journal of Nutrition Education* 1986; 15: 395-415.
- Glanz K, Mullis RM. Environmental Intervention to Promote Healthy Eating: A Review of Models, Programs and Evidence. *Health Education Quarterly* 1988; 15(4): 395-415.
- Glanz K, Sorensen G, Farmer A. The health impact of worksite nutrition and cholesterol intervention programs. *American Journal of Health Promotion*. 1996 Jul-Aug;10(6):453-70.
- Glanz K, Sallis JF, Saelens BE, Frank LD. Healthy nutrition environments: concepts and measures. *American Journal of Health Promotion*. 2005 May-Jun;19(5):330-3.
- Green, B.F. System design and operation. In Wainer, H. (Ed.) *Computerized Adaptive Testing: A Primer*. Mahwah, NJ: Lawrence Erlbaum Associates. 2000
- Green LW, Kreuter MW. Health promotion planning: an educational and environmental approach. Mountainview, California: Mayfield, 1991. In: Sallis JF, Bauman A, Pratt M. Environmental and policy interventions to promote physical activity. *American Journal Preventive Medicine*.1998;15(4): 379-397.
- Haslam DW, James WP. Obesity. *Lancet* 2005, 366(9492): 1197-1209.
- Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, Macera CA, Heath GW, Thompson PD, Bauman A. Physical Activity and Public Health. Updated Recommendation for Adults From the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007 Aug 1; [Epub ahead of print]
- Heaney CA, Goetzel RZ. A review of health-related outcomes of multi-component worksite health promotion programs. *American Journal of Health Promotion*. 1997. Mar-Apr;11(4):290-307.
- Hendriksen IJ, Zuiderveld B, Kemper HC, Bezemer PD. Effect of commuter cycling on physical performance of male and female employees. *Medicine Science in Sports and Exercise*. 2000 Feb;32(2):504-10.
- Hu F, Willett W. Optimal diets for prevention of coronary disease. *JAMA* 2003; 27: 2569-2578.

- Jacobs DR, Ainsworth BE, Hartman TJ, et al.. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Medicine Science in Sports and Exercise*. 1992; 25 (1): 81-91
- Janer G, Sala M, Kogevinas M. Health promotion trials at worksites and risk factors for cancer. *Scandinavian Journal of Work Environmental Health*. 2002 Jun;28(3):141-57.
- Jeffery RW, Utter J. The changing environment and population obesity in the United States. *Obesity Research* 2003; 11 Suppl. 12S-22S.
- Kahn EB, Ramsey LT, Brownson RC, Heath GW, Howze EH, Powell KE, Stone EJ, Rajab MW, Corso P. The effectiveness of interventions to increase physical activity. A systematic review. *American Journal of Preventive Medicine*. 2002 May;22(4 Suppl):73-107.
- Kerr J, Kerr, Eves F and Carroll D. Can posters prompt stair use in a worksite environment?, *Journal of Occupational Health* 43 (2001), pp. 205–207.
- Kriska AM, Knowler WC, LaPorte RE, et al.. Development of questionnaire to examine relationship of physical activity and diabetes in Pima Indians. *Diabetes Care* 1990; 13 (4): 401-11 In: Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine*. 2001 Feb;31(2):91-100.
- Lowther M, Mutrie N, Loughlan C, McFarlane C. Development of a Scottish physical activity questionnaire: a tool for use in physical activity interventions. *British Journal of Sports Medicine*. 1999 Aug;33(4):244-9.
- Levine JA, Baukol PA, Westerterp KR. Validation of the Tracmor triaxial accelerometer system for walking. *Medicine Science in Sports and Exercise*.. 2001 Sep;33(9):1593-7.
- Lissner L. Measuring food intake in studies of obesity. *Public Health Nutrition*. 2002 Dec;5(6A):889-92.
- Matson-Koffman DM, Brownstein JN, Neiner JA, Greaney ML. A site-specific literature review of policy and environmental interventions that promote physical activity and nutrition for cardiovascular health: what works? *American Journal of Health Promotion*. 2005 Jan-Feb;19(3):167-93.
- McClelland JW, Keenan DP, Lewis J, Foerster S, Sugerman S, Mara P, Wu S, Lee S, Keller K, Hersey J, Lindquist C. Review of evaluation tools used to assess the impact of nutrition education on dietary intake and quality, weight management practices, and physical activity of low-income audiences. *Journal of Nutritional Education*. 2001;33 Suppl 1:S35-48.
- Oldenburg B, Sallis JF, Harris D, Owen N. Checklist of Health Promotion Environments at Worksites (CHEW): development and measurement characteristics. *American Journal of Health Promotion*. 2002 May-Jun;16(5):288-99.
- Pate R, Pratt M, Blair S, et al.. Physical activity and public health: a recommendation from the centers for disease control and prevention and the American College of Sports Medicine. *JAMA* 1995; 273: 402-407.
- Popkin, BM. "Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases." *The American Journal of Clinical Nutrition* 2008; 84, 289-98
- Popkin, BM, et al "Measuring the full economic costs of diet, physical activity and obesity-related chronic diseases." *The International Association for the Study of Obesity*; 2006; 7, 271-93.
- Popkin BM, Duffey K, Gordon-Larsen P. Environmental influences on food choice, physical activity and energy balance. *Physiology and Behavior* 2005; 86(5): 603-613.
- Powell KE, Blair SN. The public health burdens of sedentary living habits: theoretical but realistic estimates. *Medicine Science in Sports and Exercise*. 1994; 26(7): 851-856.

- Proper KI, Koning M, van der Beek AJ, Hildebrandt VH, Bosscher RJ, van Mechelen W. The effectiveness of worksite physical activity programs on physical activity, physical fitness, and health. *Clinical Journal of Sport Medicine*. 2003 Mar;13(2):106-17.
- Proper KI, Bergstra B, Bakker I, Mechelen van W, De effectiviteit van interventies ter stimulering van beweging [The effectiveness of intervention to stimulate physical activity]. *Tijdschrift voor bedrijfs- en verzekeringsgeneeskunde*. 2006a, 14; 103-109.
- Proper KI, Bakker I, Overbeek K, Verheijden M, Mechelen van W, De effectiviteit van interventies ter stimulering van gezonde voeding [The effectiveness of intervention to stimulate healthy nutrition]. *Tijdschrift voor bedrijfs- en verzekeringsgeneeskunde*. 2006b, 14; 243-248.
- Rennie KL, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutrition*. 1998 Dec;1(4):265-71.
- Richardson MT, Leon AS, Jacobs DR, et al.. Comprehensive evaluation of the Minnesota Leisure Time Physical Activity Questionnaire. *Journal of Clinical Epidemiology* 1993; 47: 271-81 In: Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine*. 2001 Feb;31(2):91-100.
- Rowlands AV, Thomas PW, Eston RG, Topping R. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Medicine Science in Sports and Exercise*. 2004 Mar;36(3):518-24.
- Rockett HR, Berkey CS, Colditz GA. Evaluation of dietary assessment instruments in adolescents. *Current Opinion in Clinical Nutrition and Metabolic Care*. 2003 Sep;6(5):557-62.
- Ruwaard D, Kramers PGN. Volksgezondheid toekomst verkenning 1997. De som der delen. [Public Health; exploration of future developments 1997. Som of the parts.] Utrecht/ Amsterdam: Rijksinstituut voor Volksgezondheid en Milieu. Elsevier/ De Tijdstroom, 1997.
- Sallis JF, Glanz K. The role of built environments in physical activity, eating, and obesity in childhood. *Future Child*. 2006 Spring;16(1):89-108.
- Sallis JF, Kraft K, Linton LS. How the environment shapes physical activity: a transdisciplinary research agenda. *American Journal of Preventive Medicine*. 2002 Apr;22(3):208.
- Saelens BE, Glanz K, Sallis JF, Frank LD. Nutrition Environment Measures Study in restaurants (NEMS-R): development and evaluation. *American Journal of Preventive Medicine*. 2007 Apr;32(4):273-81.
- Seidell JC. Prevalence and time trends of obesity in Europe. *Journal of Endocrinological Investigation* 2002; 25: 816-22.
- Seidell JC, Visscher TL. [Nutrition and health--obesity] *Ned Tijdschr Geneesk*. 2003; 147(7): 281-286. Erratum in: *Ned Tijdschr Geneesk*. 2003; 147(52): 2612. *Ned Tijdschr Geneesk*. 2003; 147(10): 472.
- Seymour JD, Yaroch AL, Serdula M, Blanck HM, Khan LK. Impact of nutrition environmental interventions on point-of-purchase behavior in adults: a review. *Preventive Medicine*. 2004 Sep;39 Suppl 2:S108-36.
- Shimotsu ST, French SA, Gerlach AF, Hannan PJ. Worksite environment physical activity and healthy food choices: measurement of the worksite food and physical activity environment at four metropolitan bus garages. *The International Journal of Behavioral Nutrition and Physical Activity*. 2007 May 11;4:17.
- Sparling PB, Owen N, Lambert EV, Haskell WL. Promoting physical activity: the new imperative for public health. *Health Education Research* 2000 Jun; 15(3), 367-376.

Slootmaker S, Chin A Paw, M, Koppes L, Schuit A, Van Mechelen W. Concurrent validity of the PAM accelerometer relative to the MTI 1 Actigraph using oxygen consumption as a reference (submitted).
Steenhuis I, van Assema P, Reubsaet A, Kok G. Process evaluation of two environmental nutrition programs and an educational nutrition programme conducted at supermarkets and worksite cafeterias in the Netherlands. *Journal of human nutrition and dietetics*. 2004 Apr;17(2):107-15.

Swinburn B, Egger G. The runaway weight gain train: too many accelerators, not enough brakes. *BMJ* 2004; 329(7468): 736-739.

Swinburn B, Egger G, Raza F. Dissecting obesogenic environments: the development and application of a framework for identifying and prioritizing environmental interventions for obesity. *Preventive Medicine*. 1999 Dec;29(6 Pt 1):563-70.

Taylor CB, Coffey T, Berra K, Iaffaldano R, Casey K, Haskell WL. Seven-day activity and self-report compared to a direct measure of physical activity. *American Journal of Epidemiology*. 1984 Dec;120(6):818-24.

Thissen, D., & Mislevy, R.J. Testing Algorithms. In Wainer, H. (Ed.) Computerized Adaptive Testing: A Primer. Mahwah, NJ: Lawrence Erlbaum Associates. (2000).

Tudor-Locke C, Burkett L, Reis JP, Ainsworth BE, Macera CA, Wilson DK. How many days of pedometer monitoring predict weekly physical activity in adults? *Preventive Medicine*. 2005 Mar;40(3):293-8.

Tudor-Locke CE, Myers AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Medicine*. 2001 Feb;31(2):91-100.

U.S. Surgeon General. Physical activity and health; a report of the surgeon general. Atlanta GA: U.S. Department of Health and Human Services, CDC, National Center for Disease Control and Prevention, USA, 1996.

Van Assema P, Brug J, Ronda G, Steenhuis I, Oenema A. A short Dutch questionnaire to measure fruit and vegetable intake: relative validity among adults and adolescents. *Nutrition and Health* 2002; 16(2): 85-106.

Van Assema P, Brug J, Ronda G, Steenhuis I. The relative validity of a short Dutch questionnaire as a means to categorize adults and adolescents to total and saturated fat intake. *Journal of human nutrition and dietetics*. 2001; 14(5): 377-390.

Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. *American Journal of Preventive Medicine*. 2007 Jul;33(1):54-64.

Vandelanotte C, Matthys D, De bourdeaudhuij I. Reliability and validity of a computerized questionnaire to measure fat intake in Belgium. *Nutrition Research*, Volume 24, Issue 8, Pages 621-631

Van Sluijs EM, Van Poppel MN, Twisk JW, Van Mechelen W. Physical activity measurements affected participants' behavior in a randomized controlled trial. *Journal of Clinical Epidemiology* 2006; 59(4): 404-411.

Verplanken B, Aarts H, Van Knippenberg A, Moonen A. Habit versus planned behaviour: a field experiment. *British Journal Social Psychology* 1998; 37(Pt 1): 111-28.

Visscher TLS, Seidell JC. The public health impact of obesity. *Annual Review of Public Health* 2001; 22: 355-357.

Visscher TL, Kromhout D, Seidell JC. Long-term and recent time trends in the prevalence of obesity among Dutch men and women. *International Journal of Obesity Related Metabolic Disorders*. 2002; 26(9): 1218-1224.

Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *Journal of Clinical Epidemiology*. 2003 Dec;56(12):1163-9.

WHO Regional office for Europe. Healthy Nutrition. Preventing nutrition related diseases in Europe. WHO regional publications 24, 1988. In: Braeckman L, Maes L, et al.. Worker participation in a nutrition education programme. *Archives of Public Health* 1998; 56: 275-289.

References of the case reports

Case report 1:

C. Everett Koop National Health Awards. The Health Project. 2000.
(<http://healthproject.stanford.edu/koop/caterpillar/description.html>)

Caterpillar. Health Benefits. Healthy Balance Program. 2007.
(<http://cathealthbenefits.cat.com/cda/layout?m=46929&x=7>)

Case report 2:

Virgin Life Care. Health Rewards Program. 2007. (<http://www.virginlifecare.com/>)

Case report 3:

Vancouver International Airport Authority (British Columbia). Case Study. March 2002. Clarence Lochhead. <http://www.clbc.ca/files/CaseStudies/vancouverairport.pdf>

Fitness and Balance: Creating a Healthy Workplace at Vancouver International Airport Authority.
(www.clbc.ca/files/Presentations/Beckett_presentation.ppt)

Appendix: Case reports

1. Caterpillar

Name of Organization: Caterpillar (USA) (large manufacturer of industrial equipment)

Name of Wellness Program(s): Healthy Balance Program & Healthtrac® Program

Program Description: Salient components of the Caterpillar Healthy Balance Program include confidential health assessments, personalized health education messages, phone counselling, healthy cafeteria and vending machine options, walking paths, onsite exercise and stress management classes, reimbursement for fitness club membership, a toll free health information line, etc.

Description of M&E: Ongoing evaluation has been essential to this program's success, and data regarding employee health, medical claims, absenteeism, etc. is continually gathered and integrated to track and improve upon health promotion programs. To facilitate this process, Caterpillar's Healthy Balance Program is administered by an external organization named Healthtrac®, who, in conjunction with the Caterpillar corporate Medical Director, keeps track of information such as program participation rates, employee retention, self-reported risk factors and health indicators (e.g. weight, BMI), the use of various healthcare services (e.g. doctor visits), employee satisfaction, and baseline vs. current medical claims expenses. This allows Caterpillar to monitor vital employee health-related statistics, and to adapt its programs accordingly in order to optimize both health and financial outcomes. Caterpillar emphasizes the importance of ongoing M&E of its health initiatives to its capacity to make continuous improvement, which translates to maximized health and cost savings.

Comments: Caterpillar provides an example of a combined in-house/external team working together to coordinate M&E efforts. Moreover, the integration of outcome indicators to inform improvements to ongoing wellness programs may help achieve more effective programs.

2. Virgin Life Care

Name of Organization: Virgin Life Care (USA) (provides corporate wellness solutions)

Name of Wellness Program: Health Miles Health Rewards Program

Program Description: Virgin Life Care, a member of the Virgin family, provides corporate wellness solutions to organizations seeking to promote healthy lifestyles and behaviours among employees. Virgin Life Care incentivizes healthy behaviours with "HealthMiles", which can be redeemed by employees who engage in physical activity and improve biometrics (such as blood pressure and BMI) for popular retail goods and services.

Description of M&E: Monitoring and evaluation of employee progress is fundamental to the success of Virgin Life Care's business. Three technological solutions help facilitate this process. First, employees are equipped with "GoZone Pedometers", which record physical activity parameters such as steps taken, distance travelled, and calories burned. Information is then uploaded via computer into online employee "LifeZone" folders, where accumulated activity and progress is tracked by Virgin. Employees also regularly visit a "HealthZone" health-kiosk (which is typically located on-site) in order to automatically check their weight, blood pressure, and body fat. These data are also uploaded to their online "LifeZone". Employees are then able

to monitor their activity levels and biometrics, as well as the number of accumulated HealthMiles at their online Lifezone. In addition, they can create and follow exercise programs, read health and fitness news updates, and make purchases using their HealthMiles. In this way, companies and employees alike are able to efficiently track employee activity levels and vital health statistics.

Comments: By offering both the incentives to engage in healthy behaviours, as well as M&E solutions to track employee health progress, the Virgin Life Care Health Rewards Program is one way that companies can initiate employee wellness programs with minimal investment in on-site infrastructure or in-house M&E. Anecdotal evidence suggests high participation rates, significant reductions in the prevalence of overweight/obesity and hypertension, and cost savings for participating companies. However, no controlled studies were identified which assess the effectiveness of this program.

3. Vancouver International Airport Authority

Name of Organization: Vancouver International Airport Authority (Canada)

Name of Wellness Program: Employee Wellness Program

Program Description: The Vancouver International Airport Authority (YVR) is a community based non-profit organization with over 300 employees. The YVR initiated a comprehensive employee wellness program in 2001, which was planned by an inter-departmental committee comprised of union and management employees, and focuses on voluntary, honour-based incentive programs wherein employees are rewarded for making healthy choices. A multifaceted approach to health promotion is emphasized. However, one program in particular, entitled "Fitness and Balance", encourages employees to make lifestyle choices that are right for them. Physical activity is one such component, as are meditation, community involvement, environmentally conscious projects, etc. Several elements of this program include seminars, workshops, lunch & learn sessions, health-related newsletters, fitness club discounts, annual wellness fairs, and outdoor activity days for employees and their families.

Description of M&E: The results of M&E of this and similarly purposed health promotion programs at YVR provide some indications that the company and the employees are benefiting. In 2003, for example, over one third of employees were reportedly enrolled in the wellness program. In the first several years since the implementation of the wellness program, total absenteeism fell from 4.1% (1999) to 2.6% (2002), and workers compensation board (WCB) injuries and "days lost" dropped from 22 down to 6 and from 223 down to 24 per year, respectively, over the same time. Moreover, 75% of employees agree that YVR is doing a "very good job" at addressing employee health and safety. Although these are important data describing distal health-related economic outcomes for YVR, no such data were identified that address more direct measures of health risk factors or behaviours of employees, such as employee diet, physical activity, weight, or relevant biomarkers.

Distinguishing Characteristics: YVR's in-house M&E has helped to establish financial and qualitative indicators for their employee wellness program, which may justify further support for investment. In one respect, with such a large number of programs ongoing, evaluating distal outcome measures, such as absenteeism, may be advantageous in that it captures collective contributions of various programs. However, adding more proximate indicators such as employee health-related behaviours or biomarkers would add another component to evaluation.