Economic evaluations of worksite health promotion programs

Johanna Maria van Dongen

BODY@WORK

Economic evaluations of worksite health promotion programs

The studies presented in this thesis were conducted at the Department of Health Sciences and the EMGO+ Institute for Health and Care Research of the VU University Amsterdam. The EMGO+ Institute participates in the Netherlands School of Primary Care Research (CaRe), which was acknowledged in 2005 by the Royal Netherlands Academy of Arts and Sciences (KNAW). The EMGO+ Institute participates in Body@Work, Research Center on Physical Activity, Work and Health, which is a joint initiative of VU University Medical Center (Department of Public and Occupational Health, EMGO+ Institute for Health and Care Research), VU University Amsterdam, and the Netherlands Organisation for Applied Scientific Research (TNO).

The studies in this thesis were funded by the EMGO+ Institute for Health and Care Research, Amsterdam, Fonds Nuts Ohra (Nuts Ohra Foundation), Amsterdam, and the Workplace Safety and Insurance Board, Ontario. Additional support was provided through a personal travel grant awarded by the EMGO+ Institute for Health and Care Research.

Financial support for the printing of this thesis has been kindly provided by the VU University, the EMGO+ Institute for Health and Care Research, the Department of Health Sciences, and Body@Work, Research Center on Physical Activity, Work and Health.

English title Economic evaluations of worksite health promotion programs

Dutch title Economische evaluaties van leefstijlinterventies op de werkplek

Cover: Dimitri Valentijn (www.dimitrivalentijn.nl)

Lay-out: Gildeprint, Enschede (www.gildeprint.nl)

Printed by: Gildeprint, Enschede (www.gildeprint.nl)

ISBN: 978 94 6108 661 7

VRIJE UNIVERSITEIT

Economic evaluations of worksite health promotion programs

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Aard- en Levenswetenschappen
op vrijdag 13 juni 2014 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Johanna Maria van Dongen
geboren te Haarlem

promotoren:	prof.dr. M.W. van Tulder
	prof.dr. A.J. van der Beek
	prof.dr.ir. P.M. Bongers
copromotor:	dr. M.F. van Wier

CONTENT

1

General introduction

GENERAL INTRODUCTION

Many chronic diseases and various physical and mental health problems are in large part preventable (1). An overwhelming body of evidence indicates that an unhealthy lifestyle (e.g. insufficient physical activity, unhealthy diet, smoking, high levels of alcohol consumption, low levels of relaxation) and other modifiable health risks (e.g. high levels of work stress) play an important role in the development of such conditions (1-7). Nonetheless, the prevalence of modifiable health risks is high (8;9). Moreover, even though the prevalence of some risk factors slightly decreased during the last decades (e.g. smoking), that of others increased dramatically in many developed countries due to changes in daily (working) life (10). For example, an increased availability of larger portion sizes, lower prices of unhealthy food, and the influence of commercials has led to an increased energy intake among the population, while at the same time work-related and leisure-time activity levels decreased (11;12). As a consequence, the prevalence of overweight (Body Mass Index [BMI]≥25 kg m^{-2} and <30 kg m^{-2}), obesity (BMI≥30 kg m^{-2}), and their attributable diseases (i.e. type-2 diabetes, cardiovascular disease, and certain types of cancer) increased dramatically during the last decades (13). In 2008, the World Health Organization estimated that globally more than 1.4 billion adults were overweight or obese (14). In the Netherlands, the combined prevalence of overweight and obesity is estimated to be 43% in adult women and 54% in adult men (15). Also, due to increased work pressure, competition, work pace, and job instability, working life became more emotionally and mentally demanding (16-18). Currently, 36% of Dutch workers "*regularly have to work at a high work pace*" and 30% *"regularly have to work under high time pressure"* (19). As a consequence, workers experience higher levels of work stress as compared to a couple of decades ago (20), which in turn may lead to the development of various stress-related problems (e.g. mental disorders and cardiovascular disease) (18;21-27).

Economic consequences of modifiable health risks

Next to the human suffering associated with modifiable health risks, the economic consequences are considerable. Various studies indicate that such risk factors are associated with increased medical spending (28-30). For example, Goetzel et al. (1998) examined the relationship between ten modifiable health risks (e.g. stress, smoking, unhealthy diet, alcohol abuse, physical inactivity) and medical claim costs among 46,026 United States (U.S.) adult workers. They found that workers with these risk factors had approximately 25% higher healthcare expenditures than workers without these risk factors (29). Several other reports have shown that people with modifiable health risks are also more likely to be absent from work (i.e. absenteeism), are less productive while at work (i.e. presenteeism), and have higher work disability rates (31-35). Boles et al. (2004), for example, found that mean absenteeism and presenteeism rates ranged from 0.0% to 6.3% and 1.3% to 25.9% in workers with zero to eight risk factors, respectively (32;34). Tsai et al. (2005) also demonstrated an association between the number of modifiable health risks and absenteeism from work, with the average number of sickness absence days per year ranging from 4.1 days among workers with zero risk factors to 12.6 days among those with four or more risk factors (34;36). These findings indicate that health promotion programs aimed at preventing and/or reducing (the number of) modifiable health risks may not only be useful to reduce their individual health consequences but also their associated costs.

Rationale for improving health at the workplace

From a public health perspective, implementing health behavior change interventions in the occupational setting offers a number of advantages over approaches to health promotion in other settings. Amongst them are the possibilities to:

1) reach a higher percentage of participants that could benefit from a health promotion program than in, for example, the public health setting,

2) implement multi-level interventions that also address work organizational and environmental/policy variables in addition to individual health behaviours,

3) offer health promotion programs at relatively low costs, because the infrastructure necessary for program implementation is often already available, especially in large enterprises,

4) enhance maintenance of behavior changes, because employees spend many hours at the workplace and organizational and social support can be made easily available (16;31;37;38).

On top of that, employers themselves may (financially) benefit from implementing worksite health promotion programs as they bear most of the financial consequences of increased absenteeism, presenteeism, and work disability rates (16;31;37). In countries with employer-provided healthcare insurance (e.g. the U.S.) they also bear a large part of the medical costs of their workers. In addition, the looming labor shortages associated with the current ageing of the population make it even more important for employers that their workers are vital and healthy so that they can prolong their active labor participation (39).

Worksite health promotion programs

Today, many employers associate poor health with reduced employee performance, safety, and morale (31). Therefore, they increasingly turn to worksite health promotion programs in an effort to manage employee health and costs (31). To be effective, such programs should be developed in close cooperation with employers and have to be tailored to the needs of the employees at hand (10). The latter is critical as the needs of employees seem to vary by age, gender, type of industry, and job category (40). For example, the prevalence of overweight and obesity is particularly high among blue collar workers, whereas high levels of work stress may be a particular concern among white collar workers (41;42).

In recent years, four different research projects were performed at the EMGO^{+} Institute for Health and Care Research, in which several worksite health promotion programs were developed in close cooperation with managers and employees of various participating companies. All programs were systematically tailored to the stakeholders' needs by using the so-called "Intervention Mapping protocol" (43). These programs included:

1) <u>The Vital@Work intervention</u>: A worksite health promotion program aimed at improving physical activity, nutrition, and relaxation, as a potentially effective tool to keep older hospital workers vital and healthy, and thereby contributing to prolonged employability. Vitality is characterized by a perceived high energy level, low levels of fatigue, and feeling fit (44).

2) The Mindful VIP intervention: A mindfulness-based worksite intervention aimed at improving work engagement among knowledge workers (45). Work engagement is defined as *"a positive, fulfilling, work-related state of mind that is characterized by vigor (i.e. vitality), dedication, and absorption"* and was previously found to be negatively associated with burnout, depression, and psychosomatic complaints (46-48).
3) The VIP in Construction intervention: A worksite health promotion program aimed at improving nutrition and physical activity among construction workers. The program was developed in an effort to combat the high prevalence of overweight, obesity, and musculoskeletal disorders among construction workers (49).
4) The Be Active & Relax VIP intervention: A combined social and physical environmental intervention aimed at reducing the need for recovery from work related fatigue in office workers (50). Need for recovery was previously found to be associated with various stress-related problems (e.g. mental disorders and cardiovascular disease) and increased absenteeism (22;24;51).

The impact of worksite health promotion programs

According to *"The conceptual model of health promotion",* the aforementioned interventions may lead to improvements in individual health and various corporate benefits, such as increased corporate reputation and employee retention as well as reduced absenteeism and presenteeism costs. Improvements in individual health outcomes are thought to occur directly from program impact. Corporate benefits, on the other hand, are hypothesized to occur indirectly as the result of individual health improvements or directly from program impact. For example, improved health and/or well-being may lead to lower absenteeism and/or presenteeism costs, while the provision of a worksite health promotion program itself may improve corporate reputation and/or employee retention (Figure 1) (52).

In accordance with this model, various systematic reviews indicated that worksite health promotion programs can be effective in improving employee health and well-being (53-56). Worksite physical activity and/or nutrition programs, for example, were found to be effective in reducing body fat, waist circumference, and body weight (53;54). Worksite stress reduction programs, on the other hand, seem to be effective in reducing the levels of stress, burnout, and/or anxiety among workers (55). Research also indicates that worksite health promotion programs are effective in reducing absenteeism and presenteeism rates, and healthcare utilization (57-60), but evidence on their impact on other corporate benefits, such as improved corporate reputation, remains limited.

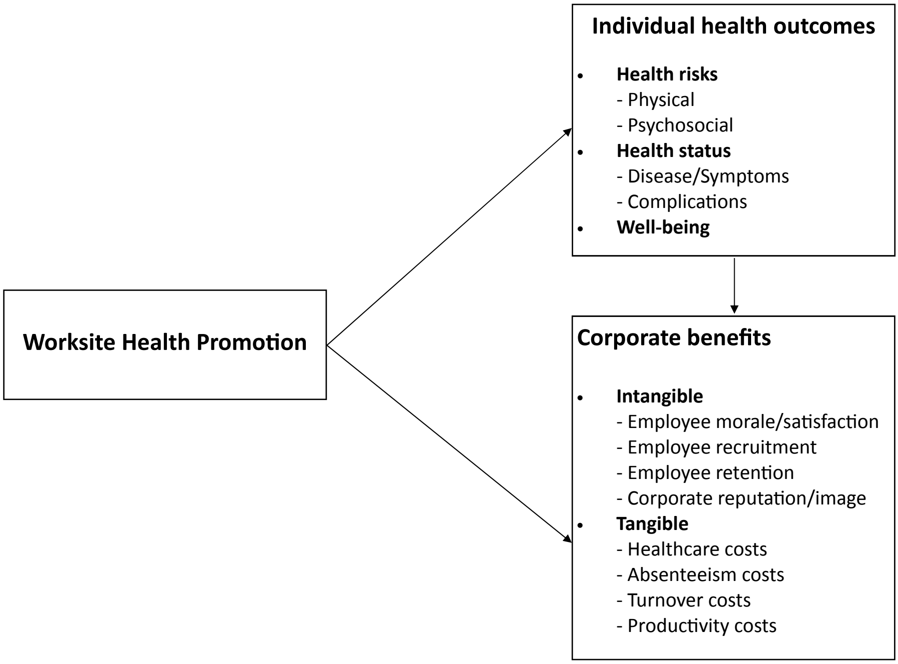

Figure 1: A simplified representation of "The conceptual model of health promotion" (52)

Economic evaluations of worksite health promotion programs

The effectiveness of the aforementioned worksite health promotion programs in comparison to usual practice will be evaluated using a randomized controlled trial (RCT) (44;45;49) or a study with a 2X2 factorial design (50). However, as resources for occupational health are restricted, decisions about investments in such programs are not only guided by evidence of their effectiveness, but also by considerations of their efficiency in terms of their resource utilization (61). To inform such decisions, economic evaluations can be conducted, which are defined as *"the comparative analysis of alternative courses of action in terms of both their costs and consequences"* (62). Economic evaluations inform decision-makers about whether the (financial) consequences of a new program justify their possible additional costs as compared to an alternative strategy (e.g. usual practice) (63). Several kinds of economic evaluations exist. In *cost-effectiveness analyses* (CEAs), incremental costs of alternatives are compared to their incremental effects (expressed in natural units). Such analyses provide insight into the (extra) cost per additional unit of effect gained. *Cost-benefit analyses* (CBAs), also known as return-on-investment (ROI) analyses, provide insight into the net financial benefit or financial return of a program by comparing incremental costs to incremental benefits of alternatives (i.e. program outcomes converted to monetary units). In *cost-utility analyses* (CUAs), the incremental costs of a program are compared to its attributable health improvements measured in utility units (e.g. "Quality Adjusted Life Years") (62;64). In *cost-minimization analyses* (CMAs), only the incremental costs of alternatives are compared when it is assumed that their consequences are similar. CMAs are therefore considered inappropriate if there is uncertainty regarding a possible difference in the magnitude of consequences (62).

Critical elements in the design of an economic evaluation are the choice of the kinds of economic evaluations that are performed as well as the applied analytic perspective(s) (e.g. societal perspective, employer's perspective). When evaluating worksite health promotion programs these choices can be challenging due to the relative complexity of the occupational health decision-making context that generally includes multiple stakeholders (e.g. individual workers, employers, occupational health services, healthcare insurance companies, income insurance companies, public policy makers). A major consideration should be the trade-off and analytic perspective that matters most to the decision-maker(s) at hand (65-67).

Choosing the appropriate kind of economic evaluation

CEAs are of particular interest to occupational health researchers, workers, and public policy makers, particularly if monetary measures do not adequately capture important health outcomes (62;67). CBAs/ROI analyses, on the other hand, are more salient to decision-makers at the company level, since they can give an indication of a program's impact on a company's bottom-line (66).

As decisions to implement worksite health promotion programs are typically made by the company's management, it would be natural to consider focusing exclusively on financial outcomes (i.e. by solely performing a CBA/ROI analysis) (66). This approach, however, has several shortcomings. First, it ignores the fact that the primary objective of worksite health promotion programs is to enhance employee health. Second, it runs the risk of overlooking the fact that costs may be reduced without health improvements (68). Third, various corporate benefits of worksite health promotion programs are hard to monetize (e.g. job satisfaction, corporate reputation) and can therefore not be included in a CBA/ROI analysis. Fourth, it does not provide relevant information to all stakeholders involved. Within this thesis, both CEAs and CBA/ROI analyses will therefore be performed of the aforementioned worksite health promotion programs.

Choosing the appropriate analytic perspective of an economic evaluation

The analytic perspective refers to the "point of view" taken for identifying relevant costs and consequences for inclusion in the evaluation. The chosen perspective may be that of any relevant stakeholder or an aggregate of stakeholders, such as a societal perspective. An item may be considered a cost from one perspective, but not from another (62). In the societal perspective, for example, all costs and consequences are considered irrespective of who pays or benefits, whereas only those borne by employers are taken into account when the employer's perspective is applied.

As mentioned earlier, decisions to implement worksite health promotion programs are typically made by the company's management. Therefore, economic evaluations of worksite health promotion programs are typically performed from the employer's perspective, but other perspectives may also be relevant, such as the societal, worker's, healthcare insurance's, and income insurance's perspective. The societal perspective is particularly useful, because it provides insight into the distribution of costs and benefits between various stakeholders and thereby allows for bargaining between them (62). This is of particular importance in the Dutch situation, in which employers bear most of the costs of worksite health promotion programs, while the government and healthcare insurance companies reap a large part of their possible benefits (i.e. reduced medical spending) (69). Also, the application of the societal perspective better ensures that the societal costs of an intervention are less than the benefits experienced by all stakeholders, rather than simply the company's costs being less than their benefits (68).

As there is no restriction to the number of perspectives that can be taken in an economic evaluation, the aforementioned worksite health promotion programs will be evaluated from both the employer's and societal perspective (65;68).

Transferability of economic evaluation results

Applying economic evaluation results across countries and jurisdictions is hampered by the fact that healthcare and social security systems are organized differently. As a consequence, the source and use of resources for an intervention as well as their costs and benefits may vary between countries (70). On top of that, differences exist regarding the stakeholder(s) that reap the possible benefits of worksite health promotion programs, and this particularly influences the transferability of economic evaluation results when the employer's perspective is applied. To illustrate, in countries with employer provided health insurance (e.g. the U.S.), medical costs are generally included in such analyses, as employers in such jurisdiction bear most of the healthcare costs of their employees. In countries with universal health care coverage (e.g. the United Kingdom) and dual-payer systems (e.g. the Netherlands), on the other hand, this cost category is not included, as these costs mainly accrue to the government and/or health insurance companies (71).

In recent years, various studies have been undertaken to explore how the transferability of economic evaluation results can be improved (70;72;73). Amongst others, these studies recommended the application of the societal perspective, to collect and report resource use data separately from unit costs or prices, as well as the provision of some background information on the (occupational) healthcare and social security system of the country in which the original study has been performed (70;72;73).

The Dutch (occupational) healthcare and social security system

In the Netherlands, healthcare is financed by a dual-payer system. Long-term costs of treatment, nursing, personal care, and support are covered by a universal mandatory social health insurance scheme that is financed by income-related employee contributions, supplemented by an annual State subsidy. Such costs are covered regardless of an individual's financial situation, but for most long-term care services income-related co-payments are required. Other healthcare costs are covered by private health insurance. Dutch citizens are mandated by law to buy a basic package of health insurance from a private health insurance company. Supplementary health insurance packages are voluntary (74;75). Uninsured people are liable to a penalty, but those who cannot afford the monthly insurance premiums get a financial compensation through the tax system. Health insurance companies, on the other hand, must offer the basic package of health insurance to anyone who applies, irrespective of their health or age. They get compensation for taking on high risk individuals from the "Risk Equalization fund", which is financed by income-related employee contributions and individual premiums (76;77). Even though competition in healthcare is increasing, many of the costs are still regulated by the government and therefore based on fixed prices (70).

Dutch employers are required by law to contract either a certified occupational health service or hire a board-certified occupational health and safety expert (often an occupational health physician) to assist them with occupational health and safety and sickness absence management (78). For most Dutch employees, occupational healthcare is supplied by large occupational health services operating from outside the workplace (79). Occupational healthcare is not integrated in the regular healthcare system, prices for occupational healthcare are not regulated, and all occupational healthcare costs are paid by the employers themselves (70). When an employee's sickness absence period exceeds 6 weeks, employers are obliged to seek advice from a certified occupational physician (78). In addition, they are obliged to pay at least 70% of the salary of sick employees for a period of two years, and most of them top up the wage payments from 70% to 100% during the first year of sickness absence. Small and medium-sized companies often take out an insurance contract to cover this risk, whereas larger companies typically pay for these salaries themselves.

After the two-year period, sick employees can apply for a work disability benefit through the "Institute for Employee Benefit Schemes" (UWV). UWV professionals will subsequently determine whether an employee is entitled to receive a work disability benefit, and if so, what his or her benefit level would be (≤75% of his/her last earned wage). Those who are able to work up to a certain level receive a supplement to their wage (80).

Economic evaluations and evidence-based practice in occupational health

Information on the resource implications of worksite health promotion programs seems to play an important role in daily practice when deciding whether or not they should be implemented or continued (69;81). To prevent spending already scarce resources on ineffective and/or inefficient interventions, such decisions should be based on the best available evidence (i.e. evidence-based practice). To ensure that this is the case, it is critical that both methodologically sound evidence exists on the resource implications of worksite health promotion programs and that high quality studies are used in daily practice to inform program implementation and/or continuation decisions.

The methodological quality of economic evaluations in occupational health

The number of economic evaluations in occupational health is limited (68) and the methodological quality of those that have been performed is generally poor (67;82;83). For practice, the main implication of a poor methodological quality of economic evaluations is that there is a risk that their results are biased. The use of such biased results to guide program implementation and/or continuation decisions may eventually result in inappropriate (business) investments (82). Therefore, it is of utmost importance that the methodological quality of economic evaluations in occupational health is improved. In recent years, some efforts have been undertaken to improve the methodological quality of such studies (66;68;82), but more needs to be done to accomplish this. A possible means to contribute to this cause may be to provide occupational health researchers with a brief overview of the theory and methodology of (trial-based) economic evaluations as well as recommendations for good practice regarding their design, analysis, and reporting.

The use of economic evaluations in the occupational health decision-making process
Up until now, it is unknown to what extent economic evaluations of worksite health promotion programs are used during the occupational health decision-making process. However, as research indicates that results of economic evaluations of healthcare interventions are rarely used among medical decision-makers (84-86), their use among occupational health decision-makers is likely to be limited as well. In order to improve the uptake of economic evaluation results, more insight is needed into the occupational health decision-making process as well as the information needs of decision-makers. By exploring these issues, recommendations can be made as to how occupational health researchers might better frame and disseminate their economic evaluations to ensure uptake in daily practice (87-89).

Objectives of this thesis
The aim of the present thesis is to contribute to the development of a sound evidence base on the resource implications of worksite health promotion programs as well as to improve the uptake of the results of such studies in daily practice. This will be done by summarizing the current literature, generating new evidence, and developing and providing recommendations for good practice when conducting and disseminating economic evaluations in occupational health.

Outline of this thesis
Chapter 2 and *chapter 3* describe two systematic reviews that were conducted to summarize and critically appraise the current literature on the cost-effectiveness and financial return of worksite physical activity and/or nutrition programs, respectively. *Chapter 4* through *chapter 7* contain economic evaluations of the "Vital@Work intervention" (*Chapter 4*), the "Mindful VIP intervention" (*Chapter 5*), the "VIP in Construction intervention" (*Chapter 6*), and the "Be Active & Relax VIP intervention" (*Chapter 7*). As the methodological quality of economic evaluations in occupational health is generally poor, recommendations for good practice regarding their design, analysis, and reporting are provided in *chapter 8*. To improve the uptake of economic evaluations in daily practice, *chapter 9* describes a qualitative study into the occupational health decision-making process and information needs of occupational

health decision-makers. Finally, *chapter 10* presents a general discussion of our main findings, methodological considerations, as well as recommendations for practice and research. This thesis is concluded with both a Dutch and English summary.

REFERENCES

1. Ford ES, Bergmann MM, Kröger J, Schienkiewitz A, Weikert C, Boeing H. Healthy living is the best revenge: findings from the European Prospective Investigation Into cancer and nutrition-Postdam study. *Arch Intern Med* 2009;169(15):1355-1362.
2. Kumanyika SK, Obarzanek E, Stettler N, Bell R, Field AE, Fortmann SP, et al. Population-based prevention of obesity. The need for comprehensive promotion of healthful eating, physical activity, and energy balance. A scientific statement from American Heart Association Council on epidemiology and prevention, Interdisciplinary Committee for Prevention (Formerly the Expert Panel on Population and Prevention Science). *Circulation* 2008; 118(4):428-464
3. Overweight, obesity, and health risk. National Task Force on the Prevention and Treatment of Obesity. *Arch Intern Med* 2000;160(7):898-904.
4. Sasco AJ, Secretan MB, Straif K. Tobacco smoking and cancer: a brief review of recent epidemiological evidence. Lung Cancer 2004 Aug;45, Supplement 2(0):S3-S9.
5. Room R, Babor T, Rehm J. Alcohol and public health. The Lancet 2005;365(9458):519-30.
6. Dimsdale JE. Psychological stress and cardiovascular disease. *Journal of the American College of Cardiology*. 2008;51(13):1237-1246.
7. Tennant C. Life events, stress and depression: a review of recent findings. *Australian and New Zealand Journal of Psychiatry* 2002;36(2):173-182.
8. Troost JP, Rafferty AP, Luo Z, Reeves MJ. Temporal and regional trends in the prevalence of healthy lifestyle characteristics: United States, 1994-2007. *Am J Public Health* 2012;102(7):1392-1398.
9. Hoeymans N, Melse JM, Schoemaker CG. Gezondheid en determinanten: Deelrapport van de VTV 2010 Van gezond naar beter. *Bilthoven: Rijksinstituut voor Volksgezondheid en Milieu*; 2010. Report No.: 270061006.
10. Shaw W, Reme S, Boot C. Health and wellness promotion in the workplace. In: Gatchel RJ, Schultz IZ, editors. *Handbook of Occupational Health and Wellness*. Springer: US, 2012. p. 365-382.
11. Engbers L. 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). WHO Press: Geneva, Switzerland, 2008.
12. Poskitt EME. Countries in transition: underweight to obesity non-stop? *Annals of Tropical Paediatrics: International Child Health* 2009;29(1):1-11.
13. Yach D, Stuckler D, Brownell KD. Epidemiologic and economic consequences of the global epidemics of obesity and diabetes. *Nat Med* 2006;12(1):62-66.
14. World Health Organization. Overweight and Obesity Fact sheet. http://www.who.int/mediacentre/factsheets/fs311/en/
15. Statistics Netherlands. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=81565NED&D1=a&D2=1-2&D3=a&D4=0&D5=30&HDR=T&STB=G1,G2,G3,G4&VW=T.
16. McDaid D. The economics of mental health in the workplace: what dowe know and where do we go? Epidemiol Psichiatr Soc 2007;16:294-298.
17. Siegrist J, Starke D, Chandola T, Godin I, Marmot M, Niedhammer I, et al. The measurement of effort-reward imbalance at work: European comparisons. *Soc Sci Med* 2004;58:1483-1499.
18. World Health Organization. *Mental health and well-being at the workplace - protection and inclusion in challenging times*. Copenhagen; 2010.

19. Koppes LLJ, de Vroome EEM, Mol MEM, Janssen BJM, van Zwieten MHJ, van de Bossche SNJ. *Nationale enquete arbeidsomstandigheden [Netherlands Working Conditions Survey] 2011*. http://www.tno.nl/downloads/rapport_nea_20111.pdf
20. Cartwright S, Cooper CL. Managing workplace stress. Sage Publications, Inc: Thousand Oaks, CA, US, 1997.
21. Virtanen M, Honkonen T, Kivimäki M, Ahola K, Vahtera J, Aromaa A, et al. Work stress, mental health and antidepressant medication findings from the Health 2000 Study. *J Affect Disord* 2007;98(3):189-197.
22. van Amelsvoort LG, Kant IJ, Bultmann U, Swaen GM. Need for recovery after work and the subsequent risk of cardiovascular disease in a working population. *Occup Environ Med* 2003;60(Suppl 1):i83-i87.
23. Tennant C. Work-related stress and depressive disorders. *J Psychosom Res* 2001;51(5):697-704.
24. Sluiter JK, de Croon EM, Meijman TF, Fings-Dresen MH. Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occup Environ Med* 2003;60(Suppl 1):i62-i70.
25. Cooper CL, Baglioni AJ. A structural model approach toward the development of a theory of the link between stress and mental health. *Br J Med Psychol* 1988;61(1):87-102.
26. Choi B, Ostergren PO, Canivet C, Maghadassi M, Lindenberg S, Karasek R, et al. Synergistic interaction effect between job control and social support at work on general psychological distress. *Int Arch Occup Environ Health* 2011;84(1):77-89.
27. World Health Organization. *Mental health and well-being at the workplace - protection and inclusion in challenging times*. Copenhagen; 2010.
28. Anderson DR, Whitmer RW, Goetzel RZ, Ozminkowski RJ, Wasserman J, Serxner S. The relationship between modifiable health risks and group-level health care expenditures. Health Enhancement Research Organization (HERO) Research Committee. *Am J Health Promot* 2000;15(1):45-52.
29. Goetzel RZ, Anderson DR, Whitmer RW, Ozminkowski RJ, Dunn RL, Wasserman J, et al. The relationship between modifiable health risks and health care expenditures: An analysis of the Multi-employer HERO Health Risk and Cost Database. *J Occup Environ Med* 1998;40(10):843-854.
30. Pronk NP, Goodman MJ, O'Connor PJ, Martinson BC. Relationship between modifiable health risks and short-term health care charges. *JAMA* 1999;282(23):2235-2239.
31. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Annu Rev Public Health* 2008;29(1):303-323.
32. Boles M, Pelletier B, Lynch W. The relationship between health risks and work productivity. J *Occup Environ Med* 2004;46(7):737-745.
33. Burton WN, Chen CY, Conti DJ, Schultz AB, Pransky G, Edington DW. The association of health risks with On-the-Job Productivity. *J Occup Environ Med* 2005;47(8):769-777.
34. Goetzel RZ, Carls GS, Wang S, Kelly E, Mauceri E, Columbus D, et al. The relationship between modifiable health risk factors and medical expenditures, absenteeism, short-term disability, and presenteeism among employees at Novartis. *J Occup Environ Med* 2009;51(4):487-499.
35. Henke RM, Carls GS, Short ME, Pei X, Wang S, Moley S, et al. The relationship between health risks and health and productivity costs among employees at Pepsi Bottling Group. *J Occup Environ Med* 2010;52(5):519-527.
36. Tsai SP, Wendt JK, Ahmed PS, Donnelly RP, Strawmyer TR. The illness absence patterns among employees in a petrochemical facility; impact of selected health risk factors. *J Occup Environ Med* 2005;47(8):838-846.

37. Goetzel RZ, Juday TR. What's the ROI? A systematic review on return of investment (ROI) studies of corporate health and productivity management initiatives. *AWPH's Worksite Health* 1999;(6):12-21.
38. Bull SS, Gillette C, Glasgow RE, Estabrooks P. Work site health promotion research: To what extent can we generalize the results and what is needed to translate research to practice? *Health Educ Behav* 2003;30(5):537-549.
39. Eurofound. Changes over time - First findings from the fifth European Working Conditions Survey. Résumé. http://www.eurofound.europa.eu/publications/htmlfiles/ef1074.htm
40. Sparling PB. Worksite health promotion: Principles, resources, and challenges. *Prev Chronic Dis* 2010;7(1):A25.
41. Arbouw. Bedrijfstaksatlas 2012. http://www.arbouw.nl/pdf/tools/bedrijfstak-atlas-2012.pdf
42. Proper KI, Hildebrandt VH. Overweight and obesity among Dutch workers: differences between occupational groups and sectors. *Int Arch Occup Environ Health* 2009;83:61-68.
43. Bartholomew LK, Parcel GS, Kok G, Gottlieb NH. Planning health promotion programs: an intervention mapping approach. Jossey-Bass: San Francisco: CA, 2006.
44. Strijk JE, Proper KI, van der Beek AJ, van Mechelen W. The Vital@Work Study. The systematic development of a lifestyle intervention to improve older workers' vitality and the design of a randomised controlled trial evaluating this intervention. *BMC Public Health* 2009;9(1):408.
45. van Berkel J, Proper KI, Boot CRL, Bongers PM, van der Beek AJ. Mindful "Vitality in Practice": an intervention to improve the work engagement and energy balance among workers; the development and design of the randomised controlled trial. *BMC Public Health* 2011;11(1):736.
46. Alarcon GM, Lyons JB. The relationship of engagement and job satisfaction in working samples. *J Psychol* 2011;145(5):463-480.
47. Schaufeli WB, Taris TW, Van Rhenen W. Workaholism, burnout, and work engagement: Three of a kind or three different kinds of employee well-being? *Applied Psychology* 2008;57(2):173-203.
48. Schaufeli WB, Bakker AB, Van Rhenen W. How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *J Organiz Behav* 2009;30(7):893-917.
49. Viester L, Verhagen EA, Proper KI, van Dongen JM, Bongers PM, van der Beek AJ. VIP in construction: systematic development and evaluation of a multifaceted health programme aiming to improve physical activity levels and dietary patterns among construction workers. *BMC Public Health* 2012;12(1):89.
50. Coffeng JK, Hendriksen IJ, Duijts SF, Proper KI, van Mechelen W, Boot CR. The development of the Be Active & Relax "Vitality in Practice" (VIP) project and design of an RCT to reduce the need for recovery in office employees. *BMC Public Health* 2012;12(1):592.
51. de Croon EM, Sluiter JK, Frings-Dresen MHW. Need for recovery after work predicts sickness absence. A 2-year prospective cohort study in truck drivers. J Psychos Res 2003;55(4):331-339.
52. Anderson DR, Sexner SA, Gold DB. Conceptual framework, critical questions, and practical challenges in conducting research on the financial impact of worksite health promotion. *Am J Health Promot* 2001;15(5):281-288.
53. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Lifestyle-focused interventions at the workplace to reduce the risk of cardiovascular disease--a systematic review. *Scand J Work Environ Health* 2010;36(3):202-215.

54. Verweij LM, Coffeng J, Van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. Obes Rev 2011;12(6):406-429.
55. Ruotsalainen J, Serra C, Marine A, Verbeek J. Systematic review of interventions for reducing occupational stress in health care workers. Scand J Work Environ Health 2008;34(3):169-178.
56. Richardson KM, Rothstein HR. Effects of occupational stress management intervention programs: a meta-analysis J Occup Health Psychol 2008;13(1):69-93.
57. Aldana SG, Pronk NP. Health promotion programs, modifiable health risks, and employee absenteeism. *J Occup Environ Med* 2001;43(1):36-46.
58. Kuoppala J, Lamminpää A, Husman P. Work health promotion, job well-being, and sickness absences-A systematic review and meta-analysis. *J Occup Environ Med* 2008;50(11):1216-1227.
59. Jensen JD. Can worksite nutritional interventions improve productivity and firm profitability? A literature review. *Perspect Public Health* 2011;131(4):184-192.
60. Cancelliere C, Cassidy JD, Ammendolia C, Cote P. Are workplace health promotion programs effective at improving presenteeism in workers? a systematic review and best evidence synthesis of the literature. *BMC Public Health* 2011;11(1):395.
61. Burdorf A. Economic evaluation in occupational health--its goals, challenges, and opportunities. *Scand J Work Environ Health* 2007;33(3):161-164.
62. Drummond MF, Sculpher M.J., Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.
63. Krol M. Productivity costs in economic evaluations Erasmus. Thesis (PhD), Universiteit Rotterdam, 2012.
64. Stone PW. Return-on-investment models. *Appl Nurs Res* 2005;18(3):186-189.
65. Uegaki K. Economic evaluation of interventions for occupational health. Thesis (PhD), Vrije Universiteit Amsterdam, 2010.
66. Tompa E, Culyer AJ, Dolinschi J. *Economic Evaluation of Interventions for Occupational Health and Safety: Developing Good Practice*. Oxford University Press; New York, 2008.
67. Tompa E, Dolinschi R, de Oliveira C. Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil* 2006;16(3):367-392.
68. Tompa E, Verbeek J, van Tulder MW, de Boer A. Developing guidelines for good practice in economic evaluation of occupational health and safety intervention. *Scand J Work Environ Health* 2010;36(4):313-318.
69. Downey AM, Sharp DJ. Why do managers allocate resources to workplace health promotion programmes in countries with national health coverage? *Health Promot Int* 2007;22(2):102-111.
70. Verbeek J, Pulliainen M, Kankaanpaa E, Taimela S. Transferring results of occupational safety and health cost-effectiveness studies from one country to another - a case study. *Scand J Work Environ Health* 2010;36(4):305-312.
71. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011;12(12):1031-1049.
72. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: recommendations for the design, analysis, and reporting of studies. *Int J Technol Assess Health Care* 2005;21(2):165-171.

73. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. *Value health* 2009;12(4):409-418.
74. Daley C, Gubb J. Healthcare systems: The Netherlands. http://www.civitas.org.uk/nhs/download/netherlands.pdf
75. World Health Organization. The new Dutch health insurance scheme: challenges and opportunities for better performance in health financing. Discussion Paper. http://www.who.int/health_financing/documents/dp_e_07_3-new_dutch_healthinsurance.pdf
76. Van de Ven WPMM, Schut FT. Managed competition in the Netherlands: still work-in-progress. *Health Econ* 2009;18(3):253-255.
77. Westert GP, Burgers JS, Verkleij H. The Netherlands: regulated competition behind the dykes? *BMJ* 2009;339:b3397
78. Hulshof CTJ, Frings-Dresen MHW. International OH systems. Part 2: occupational health delivery in the Netherlands. O*ccupational Health at Work* 2010;6(5):19-23.
79. De valk MA, Oostrom C, Schrijvers AJP. An assessment of occupational health care in the Netherlands (1996-2005). *Occ Med (Lond)* 2006;56(7):475-479.
80. OECD. Sickness and Disability Schemes in the Netherlands: Country memo as a background paper for the OECD Disability Review. http://www.oecd.org/social/soc/41429917.pdf
81. Miller P, Haslam C. Why employers spend money on employee health: Interviews with occupational health and safety professionals from British Industry. *Safety Science* 2009;47(2):163-169.
82. Uegaki K, de Bruijne MC, Lambeek L, Anema JR, van der Beek AJ, van MW, et al. Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. S*cand J Work Environ Health* 2010;36(4):273-288.
83. Niven KJM. A review of the application of health economics to health and safety in healthcare. *Health Policy* 2002;61(3):291-304.
84. Christiane H. The influence of economic evaluation studies on decision making.: A European survey. *Health Policy* 2000;52(3):179-192.
85. Eddama O, Coast J. A systematic review of the use of economic evaluation in local decision-making. *Health Policy* 2008;86(2-3):129-141.
86. Eddama O, Coast J. Use of economic evaluation in local health care decision-making in England: A qualitative investigation. *Health Policy* 2009;89(3):261-270.
87. Nutbeam D. Achieving 'best practice' in health promotion: improving the fit between research and practice. *Health Educ Res* 1996;11(3):317-326.
88. Sackett DL, Rosenberg WCM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312.
89. Sackett DL, Straus SE., Richardson WS, Rosenberg W, Haynes RB. *Evidence-based Medicine: How to Practice and Teach it.* Churchill Livingstone: New York, NY, 2010.

2

A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs

Johanna M van Dongen
Karin I Proper
Marieke F van Wier
Allard J van der Beek
Paulien M Bongers
Willem van Mechelen
Maurits W van Tulder

Scand J Work Environ Health 2012, 38(5): 393-408

ABSTRACT

Objective: The aim of this study was to appraise and summarize the evidence on the cost-effectiveness of worksite physical activity and/or nutrition programs.

Methods: We searched EMBASE, MEDLINE, SportDiscus, PsycInfo, NIOSHTIC-2, NHSEED, HTA, and Econlit for studies published up to 14 January 2011. Additionally, we searched for articles by reviewing references, searching authors' databases, and contacting authors of included studies. Two researchers independently selected articles. Articles had to include a cost-effectiveness and/or cost-utility analysis comparing a worksite physical activity and/or nutrition program to usual care or an abridged version of the program. Data were extracted on study characteristics and results. Two researchers independently assessed the risk of bias using the Consensus on Health Economic Criteria list (CHEC-list).

Results: Ten studies (18 programs) were included. More than 50% of the studies fulfilled 11 (58%) of the 19 CHEC-list items. From various perspectives, worksite nutrition and worksite physical activity and nutrition programs (N=6) were more costly and more effective in reducing body weight than usual care. When only intervention costs were considered, most worksite nutrition (N=4/5) and worksite physical activity and nutrition programs (N=5/6) were more costly and more effective in reducing cholesterol level and cardiovascular disease risks, respectively.

Conclusions: The cost-effectiveness of more costly and more effective programs depends on the "willingness to pay" for their effects. It is unknown how much decision-makers are willing to pay for reductions in body weight, cholesterol level, and cardiovascular disease risks. Therefore, conclusions about the cost-effectiveness of worksite physical activity and/or nutrition programs cannot be made. There is substantial need for improvement of the methodological quality of studies and particular emphasis should be placed on the handling of uncertainty.

INTRODUCTION

Regular physical activity and healthy dietary habits are considered important in preventing overweight, obesity, and their attributable diseases [eg, diabetes type 2, cardiovascular disease (CVD) and certain cancers] (1–5). Nevertheless, many adults do not meet public health recommendations for nutrition and physical activity (6–11). Currently, 34% of United States (US) adults are overweight [body mass index (BMI) ≥25 kg/m2 and <30 kg/m2] and 34% are obese (BMI ≥30 kg/m) (12). In Europe, the combined prevalence of overweight and obesity ranges from 38–61% among women and 52–69% among men (13). In addition to the toll that overweight and obesity take on the health and wellbeing of individuals, they impose considerable financial burdens in terms of increased productivity-related and healthcare costs (14–16). Therefore, health promotion programs aimed at increasing physical activity and/or improving nutrition are warranted.

The worksite provides a useful setting for implementing such programs; since employees spend up to 60% of their waking hours at the worksite, organizational and social support can easily be made available, and large enterprises often have the infrastructure to offer such programs at relatively low costs (17–19). Employers themselves may also benefit from implementing worksite health promotion (WHP) programs, as healthier workers are expected to be more productive and miss fewer days of work (17).

WHP programs aimed at increasing physical activity and/or improving nutrition were found effective in reducing body fat and body weight (20–22). For example, a recent systematic review found worksite physical activity and nutrition programs to significantly reduce body weight by 1.2 kg, BMI by 0.3 kg/m^2, and body fat percentage by 1.1% during the first years after implementation (≤3 years) (22).

Budgets for occupational healthcare are restricted. Decisions about investments in WHP programs may, therefore, not only be guided by the evidence on their effectiveness, but also by considerations of their costs in relation to these effects (23–25). For this reason, cost-effectiveness analyses (CEA) and cost-utility analyses (CUA) are conducted to gain insight into the (additional) costs of an intervention per additional unit of effect gained. These analyses not only give insight into the cost savings of an intervention, like return on investment (ROI) analyses, but also provide details on the price of achieving a particular goal if an intervention produces better outcomes at additional costs (eg, costs per kilogram body weight loss) (21). Although ROI results are likely to be most frequently used within companies to describe the financial aspects of a business case for occupational health initiatives (26, 27), CEA may be of interest for corporate decision-makers as well. A recent systematic review on the financial return of worksite physical activity and/or nutrition programs indicated that they may not pay for themselves in terms of reduced medical and/or absenteeism costs during the first years after implementation (28). Nevertheless, a significant lag between health improvements and reductions in medical and/or productivity-related cost may exist. Therefore, reporting on cost-effectiveness in terms of intermediate outcome measures that might be associated with long-term cost savings (eg, body weight loss) (29), may also give useful information to aid implementation decisions (17). Furthermore, investments in WHP programs may be motivated not only by making a profit but also by obtaining positive health effects and/or by the wish to be a caring employer. In that case, their anticipated effects are worth having and the question is to determine the most cost-effective way to achieve it (ie, least costly per unit of effect) (30).

Up until now, various reviews have been conducted on the cost-effectiveness of WHP programs (20, 31–39). One of them (36), for example, concluded that the literature provided “guarded cautious optimism” about their cost-effectiveness. However, these reviews were limited to studies published up until 2008 and most of them looked at the cost-effectiveness of WHP programs in general (ie, also including disease and stress management, and smoking cessation programs), instead of worksite physical activity and/or nutrition programs in particular. Furthermore, although the quality of the design and execution of economic evaluations should be considered when judging the validity of their findings, none of the reviews used an internationally accepted instrument for assessing their risk of bias (40, 41). This raises questions about the credibility of their conclusions. Therefore, the aim of the present study was to critically appraise and summarize the current evidence on the intermediate and long-term cost-effectiveness of worksite physical activity and/or nutrition programs compared to usual care or an abridged version of the program.

METHODS

Search strategy

A systematic search was conducted to identify studies evaluating the cost-effectiveness and/or cost-utility of WHP programs aimed at improving nutrition and/or increasing physical activity. Eight databases (EMBASE, MEDLINE, SPORTDiscus, PsycINFO, NIOSHTIC-2, NHSEED, HTA, and Econlit) were searched for studies published up to 14 January 2011.

An information specialist of the VU University Medical Center was consulted to develop and run the search strategy. Databases were searched with the following keywords: participant/setting type (eg, "Workplace", "Employee", "Workforce"), intervention type (eg, "Health Promotion", "Lifestyle"), intervention aim (eg, "Exercise", "Physical Activity", "Nutrition", and "Diet"), and study design (eg, "Cost-Effectiveness Analysis", "Cost-Utility Analysis", and "Economic Evaluation"). In addition to the present study, a systematic review on the financial return of worksite physical activity and/or nutrition programs was conducted (28). Therefore, a broad search strategy was used so that the search results could be used for both studies simultaneously. As an example, the complete search strategy for EMBASE can be found in Appendix 1. In addition to the electronic search, reference lists of relevant review articles (17, 18, 20, 21, 31–39, 42) and those of the retrieved fulltext were searched. Articles were also identified from the authors' own literature databases. To identify unpublished studies, authors of included studies published during the last decade were contacted. During the search, a "search diary" was maintained, including keywords used, searched databases, and search results.

Study selection

Titles and abstracts of retrieved studies were stored in an electronic database using Reference Manager 11.0 (ISI Research Soft Inc, Berkeley, California). Two reviewers independently assessed whether these studies met the following inclusion criteria: (i) the study included a CEA and/or CUA, (ii) participants were part of the adult working population, (iii) the intervention under study was a WHP program aimed at improving nutrition and/or increasing physical activity, (iv) the intervention was compared to usual care (including no intervention) or an abridged version of the program, (v) outcome measures included a behavioral measure (eg, physical activity and dietary intake), a health-related measure (eg, BMI, waist circumference, body fat percentage, musculoskeletal symptoms, cardiorespiratory fitness, and health risk profiles), or a work-related measure (eg, productivity and work satisfaction), and (vi) the study was reported in English, German, French, or Dutch.

For the purpose of this review, analyses could be performed from all perspectives (eg, employer's perspective and societal perspective). Furthermore, no limitations were set as to program format [eg, (self-)assessment, counselling, and exercise program], worksite characteristics (eg, age, gender, occupation, proportion of full-time employees, and number of employees), length of the intervention, and follow-up duration. Studies aimed at long-term sick-listed employees, employees with chronic conditions (eg, diabetes type 2 and CVD), retirees, and children were excluded. If studies met the inclusion criteria, or if uncertainty remained about inclusion, fulltexts were retrieved. All fulltexts were read and checked for eligibility. To resolve disagreements between the two reviewers, a consensus procedure was used. A third reviewer was consulted when disagreements persisted; this was necessary on one occasion.

Risk of bias assessment

Two reviewers independently assessed the risk of bias of included studies. If one of the reviewers was a (co-)author of a given study, another reviewer acted as the second reviewer. Risk of bias was assessed using the Consensus Health Economic Criteria list (CHEC-list), which was developed for systematic reviews of economic evaluations using a Delphi consensus procedure involving 23 international experts in economic evaluations (43). The test–retest reliability of the CHEC-list was shown to be good (intra-class correlation coefficient: 0.97, 95% CI 0.73–0.98) (44). Items were scored as negative in case of an inadequate performance of an item or if insufficient information was available in the article or related materials (43). If a study presented its results in multiple articles, those articles were scored as one study. A consensus procedure was used to resolve disagreements between the two reviewers. When disagreements remained, a third reviewer was consulted; this was necessary on two occasions.

Data extraction, data analyses, and applied classification schemes

Data were extracted on: (i) study details (eg, perspective, primary study design, setting, and follow-up duration), (ii) characteristics of the study population (eg, participant and job characteristics), (iii) program focus (ie, improving nutrition, increasing physical activity, or both), (iv) program format [ie, (self-)assessment, educational/informational, behavioral, exercise, environmental, and incentive components], (v) measurement and valuation methods of costs, (vi) measurement methods of effects, and (vii) study results [reported costs, effects, and incremental cost-effectiveness ratios (ICER)]. One reviewer extracted data using a pre-designed data extraction form. Ten percent of the extracted data was checked by a second reviewer, which did not reveal any errors. If articles did not contain sufficient information on study results, authors were contacted for missing data.

If an incremental cost-effectiveness analysis was not performed, an ICER was calculated per reported outcome measure as the incremental difference in costs relative that in effects (30). Costs and ICER were standardized to 2010 US dollars (USD) using consumer price indices (45) and purchasing power parities (46). For this, their reference year was needed. If their reference year was not stated, the year of publication was used. For data analyses and presentation, studies were grouped according to their program focus (ie, improving nutrition, increasing physical activity, or both), (stated) perspective, and outcome measures.

To summarize results, and thereby draw conclusions about the cost-effectiveness of the included programs, their incremental costs and effects were explored. Programs that were less costly and more effective than the control condition were considered cost-effective (ie, the program dominates the control condition). For programs that were more costly and less effective, the opposite was true. Programs that were either more costly and more effective or less costly and less effective were only considered cost-effective if their ICER was respectively lower or higher than the “willingness to pay” (ie, the maximum amount of money decision-makers are willing to pay per unit of effect gained) (30).

RESULTS

Literature search and study selection

The electronic search yielded 3230 unique references. After screening their abstracts and titles, we retrieved 47 fulltexts. Thirty-one additional fulltexts were retrieved after screening references of relevant review articles and those of the retrieved fulltexts. After reading those 78 fulltexts, 9 articles were identified that met the inclusion criteria (Figure 1). Additionally, three unpublished articles were identified by searching the authors' own literature databases. Contacting authors of included studies did not yield any additional results. Most studies were excluded because they did not include an economic evaluation or because they only evaluated the financial return by comparing intervention costs to their financial consequences. Eventually, 12 articles, including 10 original studies (47–58), were included in the review.

Study characteristics

A description of the study characteristics can be found in Table 1. Worksite nutrition programs (N=7) were evaluated in four studies (47, 51, 55, 56) and worksite physical activity and nutrition programs (N=11) in six studies (48–50, 52–54, 57, 58). None of the studies evaluated a WHP program solely aimed at increasing physical activity. In general, interventions consisted of a (self-)assessment, educational/informational, behavioral, exercise, environmental, and/or incentive component. All interventions were compared to usual care, consisting of no intervention or a (self-)assessment, educational/informational, and/or environmental component. The number of participants in the studies ranged from 66–1883. The length of the interventions ranged from 12 weeks to 3 years. Four studies (51, 54–56) evaluated the short-term effectiveness of the programs (follow-up ≤6 months) and six studies (47–50, 52, 53, 57, 58) evaluated the long-term effectiveness (follow-up >6 months). No studies had a follow-up >3 years. Seven studies (47–50, 52–54, 57, 58) were conducted alongside a randomized controlled trial and three (51, 55, 56) alongside a non-randomized study. Five studies (47–49, 51, 54, 56) were conducted in the US, three (50, 53, 57, 58) in the Netherlands, one (52) in Australia, and one (55) in Denmark. All studies conducted a CEA, and one (58) also conducted a CUA.

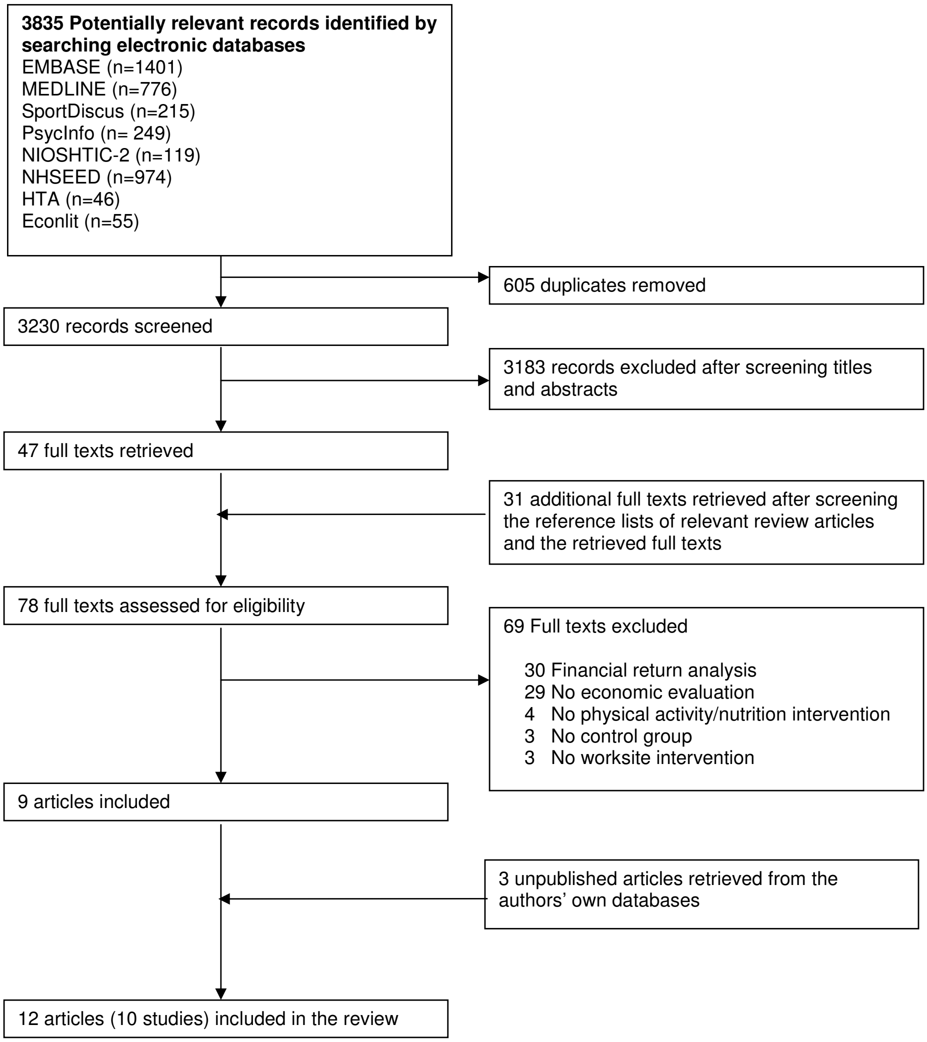

Figure 1: Flow chart for inclusion of studies

Table 1: Characteristics of the included studies (n = 10)

Study	Study Details	Population	Intervention and control conditions	Intervention purpose, format	Costs	Outcomes
			Worksite nutrition program			
Katcher et al.(51)	Perspective: NS Primary study: NRS Setting: U.S. 2007-2008 Length intervention: 22 weeks Follow-up: 22 weeks	113 *overweight* employees and/or employees with an *elevated type 2 diabetes* risk from an insurance company UC: 45 I: 68	UC: - I: Weekly (vegan) diet instructions (22 sessions), daily low-fat vegan options in cafeteria, daily multivitamin	Purpose: Improving health related quality of life and work productivity Format: UC: - I: b, e	<u>Intervention costs</u> Micro-costed, valuation method not stated	<u>Body weight</u> Digital scale, light clothing, without shoes(68) <u>Eating behavior</u> The Eating Inventory <u>Health related quality of life</u> SF-36 <u>Work productivity</u> WPAI questionnaire
Siggaard et al.(55)	Perspective: NS Primary study: NRS Setting: Denmark (year not stated) Length intervention: 12 weeks Follow-up: 12 weeks	66 insurance company employees UC:16 I: 50	UC: - I: Nutrition education (12 sessions), weekly weightings	Purpose: Weight loss Format: UC: - I: a, b	<u>Intervention costs:</u> Micro-costed, valuation method not stated	<u>Body weight:</u> Digital scale, light clothing, sober for 3 hours <u>Daily carbohydrate and fat intake:</u> 4-day dietary record
Wilson et al.(56)	Perspective: NS Primary study: NRS Setting: U.S. (year not stated) Length intervention: 1 to 3 months Follow-up: 6 months	652 manufacturing workers with an *elevated CVD risk.* UC: 146 I-1: 180 I-2: 161 I-3: 112 I-4: 53	UC: Cholesterol screening, health education materials I-1: Cholesterol screening, health education materials, 1-month educational program I-2: Cholesterol screening, health education materials, 1-month educational program, incentives I-3: Cholesterol screening, health education materials, 3-month educational program I-4: Cholesterol screening, health education materials, 3-month educational program, incentives	Purpose: CVD risk reduction Format: UC: a, b I-1: a, b I-2: a, b, f I-3: a, b I-4: a, b, f	<u>Intervention costs:</u> Micro-costed, valued using tariffs and depleted sources	<u>Cholesterol level reduction</u>: Percentage of participants with 10% (or more) cholesterol level reduction, measured using the Reflotrol dry chemical analyzer

Study	Study Details	Population	Intervention and control conditions	Intervention purpose, format	Costs	Outcomes
Byers et al.(47)	Perspective: NS Primary study: RCT Setting: U.S. (year not stated) Length intervention: 1 month Follow-up: 12 months	846 employees with an *elevated CVD risk* (variable companies) UC: 463 I: 383	UC: Cholesterol screening, health education materials I: Cholesterol screening, health education materials, nutrition education sessions (2 hours), 30 minutes videocassette on nutrition	Purpose: CVD risk reduction Format: UC: a, b I: a, b	Intervention costs: Micro-costed, valued using tariffs and depleted sources	Cholesterol level reduction: Percentage cholesterol level reduction, measured using the Reflotrol dry chemical analyzer or the Kodak system
Worksite physical activity and nutrition programs						
Rasu et al.(54)	Perspective: Implementing agency Primary study: RCT Setting: U.S. 2003-2006 Length intervention: 6 months Follow-up: 6 months	442 *overweight* U.S. Air Force employees UC: 215 I: 227	UC: HRA, standard face-to-face nutrition/ PA programs I: HRA, standard face-to-face nutrition/PA programs, in person treatment orientation meeting + 24 week behavioral internet therapy, manual for weight control, telephone counseling (2 times)	Purpose: Weight loss Format: UC: a, b I: a, b, c	Intervention costs: Micro-costed, valued using tariffs and depleted sources	Body weight: Calibrated scale, indoor clothes, without shoes Waist circumference: Gulick tape at the umbilicus Weight efficacy lifestyle: WEL questionnaire
Groeneveld et al.(50)	Perspective: Societal Primary study: RCT Setting: NL 2007-2009 Length intervention: 6 months Follow-up: 12 months	573 construction workers with an *elevated CVD risk* UC: 280 I: 293	UC: HRA, health education materials I: HRA, health education materials, 6-month counseling program (3 face-to-face sessions / 4 telephone contacts)	Purpose: CVD risk reduction Format UC: a, b I: a, b, c	Intervention costs Micro-costed, valued using tariffs and depleted sources Absenteeism costs Self-reported days missed due to health complaints, multiplied by an average wage rate Medical costs Self-reported health care utilization valued using Dutch standard costs	Body weight: Digital scale, no shoes, no jacket

Study	Study Details	Population	Intervention and control conditions	Intervention purpose, format	Costs	Outcomes
Van wier et al.(58)	Perspective: Societal Primary study: RCT Setting: NL 2004-2006 Length intervention: 6 months Follow-up: 24 months	1386 *overweight* employees (variable companies) UC: 460 I-Phone: 462 I-Internet: 464	UC: Health education materials + oral instructions I- Phone: Ring binder containing 10 health education modules, telephone counseling, pedometer, health education materials + oral instructions I-Internet: Website containing 10 health education modules, email counseling, pedometer, health education materials + oral instructions	Purpose: Weight loss Format: UC: b I-Phone: b, c I-Internet: b, c	<u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources <u>Absenteeism costs</u> Self-reported number of days missed due to health complaints multiplied by an average wage rate <u>Medical costs</u> Self-reported health care utilization valued using Dutch standard costs	<u>Body weight:</u> Digital scale, light clothes, no shoes *When digital weight measurements were missing, self-reported weight measurements were used* <u>Quality Adjusted Life Years</u>: EuroQol scores valued using Dutch tariffs and multiplied with the amount of time spent in a specific health state
Gussenhoven et al.(57)	Perspective: Employer Primary study: RCT Setting: NL 2004-2005 Length intervention: 6 months Follow-up: 12 months	See Van wier et al.(58)	See Van wier et al.(58)	See Van wier et al.(58)	<u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources <u>Absenteeism costs</u> Number of days missed because of a health condition based on disability absence data multiplied by an average wage rate	<u>Body weight:</u> Self-reported body weight *When self-reported body weight data was missing, digital weight measurements were used*

Study	Study Details	Population	Intervention and control conditions	Intervention purpose, format	Costs	Outcomes
Oldenburg et al.(52)	Perspective: NS Primary study: RCT Setting: Australia (year not stated) Length intervention: 12 months Follow-up: 12 months	431 ambulance workers UC: 130 I-RFE: 82 I-BC: 124 I-BCI: 95	UC: HRA I-RFE: HRA, standardized healthy lifestyle advice, educational resource manual and videotape I-BC: HRA, standardized healthy lifestyle advice, educational resource manual and videotape, counseling sessions for participants with a high CVD risk, self-instruction lifestyle change manual I-BCI: HRA, standardized healthy lifestyle advice, educational resource manual and videotape, counseling sessions for participants with a high CVD risk, self-instruction lifestyle change manual, incentives	Purpose: CVD risk reduction Format: UC: a I-RFE: a, b I-BC: a, b, c I-BCI: a, b, c, f	<u>Intervention costs:</u> Micro-costed, valued using tariffs and depleted sources	<u>Changes in CVD risk:</u> Composite score, including; cholesterol level, blood pressure, number cigarettes smoked, body mass index, maximum oxygen uptake
Erfurt et al.(48)	Perspective: NS Primary study: RCT Setting: U.S. 1985-1988 Length intervention: 3 years Follow-up: 3 years	1607 manufacturing workers with an *elevated CVD risk (one or more of 3 CVD risks)* UC: 420 I-HE: 432 I-HEC: 408 I-HECE: 347	UC: HRA, periodic health improvement classes, fitness centre I-HE: HRA, periodic health improvement classes, wellness committee, health education strategies I-HEC: HRA, periodic health improvement classes, wellness committee, health education strategies, counseling and health improvement programs for high risk employees. I-HECE: HRA, periodic health improvement classes, wellness committee, health education strategies, counseling and health improvement programs for high risk employees, health communication networks, peer support groups, health promotion groups and organizational strategies to encourage/support health improvements	Purpose: CVD risk reduction Format: UC: a, b, d I-HE: a, b, e I-HEC: a, b, c, e I-HECE: a, b, c, e, d	<u>Intervention costs:</u> Micro-costed, valued using tariffs and depleted sources	<u>Changes in CVD risk:</u> Percentage of three risk factors (high blood pressure, overweight and smoking) that were moderately and highly reduced or prevented

Study	Study Details	Population	Intervention and control conditions	Intervention purpose, format	Costs	Outcomes
Erfurt et al. 2(49)	See Erfurt et al.(48)	1883 manufacturing workers with an *elevated CVD risk (one or more of 4 CVD risks)* UC: 493 I-FC: 505 I-HEC: 482 I-HECE: 403	UC: I-HE Erfurt et al.(48) I-FC: "UC" Erfurt et al.(48) I-HEC: "I-HEC" Erfurt et al.(48) I-HECE: "I-HECE" Erfurt et al.(48)	Purpose: CVD risk reduction Format: UC: a, b, e I-FC: a, b, d I-HEC: a, b, c, e I-HECE: a, b, c, e, d	See Erfurt et al.(48)	<u>Changes in CVD risk:</u> Percentage of four risk factors (high blood pressure, overweight, smoking and lack of exercise) that were moderately and highly reduced or prevented
Proper et al.(53)	Perspective: Employer Primary study: RCT Setting: NL 2000-2002 Length intervention: 9 months Follow-up: 9 months	299 civil servants UC: 168 I: 131	UC: Health education materials I: Health education material, 9 month counseling program (7 face-to-face sessions)	Purpose: Improving physical fitness Format: UC: b I: b, c	<u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources <u>Absenteeism costs</u> Number of days missed because of a health condition based on disability absence data multiplied by an average wage rate	<u>Meeting PA recommendations:</u> Self-reported <u>Energy expenditure:</u> Structured interview (7-day PA recall) <u>Sub-max heart rate</u>: Bicycle ergometer test <u>Upper-extremity symptoms</u>: Nordic questionnaire

NS: Not stated
NRS: Non-Randomized Trial
RCT: Randomized Controlled Trial
U.S.: United States
NL: The Netherlands
CVD: Cardiovascular disease
UC: Usual care
I: Intervention group
HRA: Health Risk Assessment
PA: Physical activity
Program format(s): a: (Self-)assessment, b: Education/Information, c: Behavioral, d: Exercise, e: Environment, f: Incentives
SF-36: Short Form-36
WPAI questionnaire: Work Productivity and Activity Impairment questionnaire
WEL questionnaire: Weight Efficacy Lifestyle questionnaire

Risk of bias assessment

Reviewers initially disagreed on 40 (21%) of the 190 items (10 studies multiplied with 19 CHEC-list items). Most disagreements were due to reading errors and different interpretations of the CHEC-list items and were solved during the consensus meeting; for four disagreements a third reviewer was consulted. Of the 19 CHEC-list items, 11 (58%) were fulfilled by more than 50% of the studies and 7 items (37%) by more than 75%. The economic perspective was specifically stated by four studies (50, 53, 54, 57, 58), including: the societal perspective, the employer's perspective, and that of an implementing agency. The latter (54) appropriately collected costs to the chosen perspective by only including intervention costs. Studies performed from the societal and employer's perspective also included absenteeism and/or medical costs and were all conducted in The Netherlands. Costs were measured in physical units [ie, individual items of an intervention were measured (30)] in four studies, (52–54, 57, 58). One of them (54) also appropriately valued costs by calculating them based on depleted sources [ie, based on the value of the forgone benefits because the resources were not available for their best alternative use (30)] and stating their reference year. Three studies (47, 51, 56) presented both costs and effects, but did not conduct an incremental cost-effectiveness analysis. Just over half of the studies conducted a sensitivity analysis to test the robustness of their results (Table 2).

Table 2: Risk of bias assessment of included studies using the CHEC-list

CHEC-list items	Studies scoring "Yes" [No. (%)] (n = 10)
1) Study population	7 (70)
2) Competing alternatives	9 (90)
3) Research question	5 (50)
4) Study design	10 (100)
5) Time horizon	9 (90)
6) Perspective	3 (30)
7) Costs identified	1 (10)
8) Costs measured	4 (40)
9) Costs valued	1 (10)
10) outcomes identified	9 (90)
11) Outcomes measured	8 (80)
12) Outcomes valued	10 (100)
13) Incremental analysis	7 (70)
14) Discounted	7 (70)
15) Sensitivity analysis	6 (60)
16) Conclusions	10 (100)
17) Generalizability	1 (10)
18) Conflict of interest	2 (20)
19) Ethical and distributional issues	0 (0)

Cost effectiveness analysis

Worksite nutrition programs.

All four studies (47, 51, 55, 56) evaluating WHP programs aimed at improving nutrition only included intervention costs in their cost estimates (Table 3). Two of them (51, 55) evaluated cost-effectiveness by comparing intervention costs to the effect on body weight reduction. Both interventions were more costly and more effective than usual care at a cost of \$43 and \$20 per kilogram body weight loss (see also table 4). One of those (55) was also more costly and more effective in reducing daily fat intake and increasing daily carbohydrate intake. The other intervention (51) was also more costly and more effective in improving physical functioning, general health, vitality, mental health, impairment at work, and impairment with daily activities. However, the intervention was more costly and less effective in reducing restraint, disinhibition (ie, overeating in response to stress or other cues), and hunger. Two other studies (47, 56) evaluated cost-effectiveness by comparing intervention costs to the effect on cholesterol level reduction. However, both studies used different outcomes for assessing the degree of cholesterol level reduction, which limits their comparability. The first study (56) evaluated four different interventions: (i) I-1: 1-month program without incentives, (ii) I-2: 1-month program with incentives, (iii) I-3: 3-month program without incentives, and (iv) I-4: 3-month program with incentives. The least intensive program (ie, 1-month program without incentives) was more costly and less effective than usual care (ICER: $-110 per 1% of participants reducing their cholesterol level by ≥10%). The other interventions were more costly and more effective (ICER I-2: $0.1; I-3: $4; and I-4: $54). The nutrition intervention evaluated by the second study (47) was also more costly and more effective than usual care at a cost of $11 per 1% cholesterol level reduction.

Worksite physical activity and nutrition programs

Six studies evaluated the cost-effectiveness of WHP programs aimed at increasing physical activity and improving nutrition (48–50, 52–54, 57, 58). Three of them (48, 49, 52, 54) only included intervention costs in their cost estimates. When costs were considered from a broader perspective (50, 53, 57, 58), intervention costs were partially offset by a reduction in absenteeism and/or medical costs (Table 3). Three studies (50, 54, 57, 58) evaluated the cost-effectiveness in terms of body weight reduction from various perspectives. All interventions were more costly and more effective than usual care. When only intervention costs were considered, the additional costs per kilogram body weight loss were $26. When analyses were performed from the employer's perspective those costs were $75 and $1534, and from the societal perspective $174, $20, and $1282 (see also table 4). One of those interventions (54) was also more costly and more effective in reducing waist circumference. Two other studies (48, 49, 52) evaluated the cost-effectiveness by comparing intervention costs to the effect on CVD risk reduction. Both studies, however, used different composite scores to estimate the level of CVD risk reduction, which limits their comparability.
The first study (52) evaluated three different interventions: (i) I-RFE: risk factor education, (ii) I-BC: behavioral counseling, and (iii) I-BCI: behavioral counseling plus incentives. All interventions were more costly and more effective than usual care (ICER I-RFE: $10, I-BC: $24, and I-BCI: $363 per CVD risk unit reduced). The other study

presented its results in two articles (48, 49), which differed in the number of CVD risk factors included in the composite score (ie, three risk factors in the first article versus four in the second article) as well as their control condition. Furthermore, in the first article (48), they did not include all intervention costs in their cost estimates (ie, fitness centre costs were missing). In the second article (49), they evaluated three different interventions: (i) I-FC: fitness centre, (ii) I-HEC: health education & follow-up counseling, and (iii) I-HECE: health education, follow-up counseling & environmental strategies. I-FC was more costly and less effective than usual care. The other interventions were more costly and more effective (ICER I-HEC: $2 and $2 I-HECE: $3 and $3 per 1% of CVD risks, respectively highly or moderately reduced). Another study (53) evaluated the cost-effectiveness from the employer's perspective using its effect on physical activity-related outcome measures. The intervention was more costly and more effective than usual care in increasing energy expenditure, and decreasing sub-maximal heart rate. However, the intervention was more costly and less effective in increasing the number of participants meeting physical activity recommendations.

Cost-utility analysis

One study (58) evaluated the cost-utility of both an internet- and a phone-based nutrition and physical activity program. Analyses were conducted from the societal perspective. After 24 months, the cost-utility of the internet-based intervention was $1698 per quality adjusted life year (QALY) gained and that of the phone-based intervention $311 523 per QALY gained.

Table 3: Costs, outcomes, and incremental cost-effectiveness ratios (Dollars, 2010) of included studies (n=10)

Study	Costs Mean [incremental]	Outcomes Mean [incremental]	Incremental cost-effectiveness ratio
		Worksite nutrition programs	
Katcher et al.(51)	UC: 0 I: 226 [226]	Body weight reduction (kg) UC: -0.1; I: 5.1 [5.3]*	43 per kg body weight loss#
		Eating behavior (points)	Costs per point decrease on the Eating Inventory subscales#
		Restraint UC:-1.1; I: 0.5 [1.6]*	Restraint -141
		Disinhibition UC: 0.1; I: 1.7 [1.7]*	Disinhibition -133
		Hunger UC: -0.7; I: 1.6 [2.3]*	Hunger -98
		Health related quality of life (points)	Costs per point improvement on the SF-36 subscales#
		Physical functioning UC: 0.23; I: 9.1 [8.9]*	Physical functioning 25
		General health UC: 2.3; I: 12.6 [10.3]*	General health 22
		Physical role limitations UC: -2.3; I: 8.5 [10.7]	Physical role limitations 21
		Emotional role limitations UC: 6.8; I: 8.7 [1.9]	Emotional role limitations 119
		Bodily pain UC: -0.3; I: 6.4 [6.7]	Bodily pain 34
		Vitality UC: -0.3; I: 10.8 [11]*	Vitality 21
		Social functioning UC: 1.7; I: 6.2 [4.4]	Social functioning 51
		Mental health UC: -1.5; I: 5.1 [6.6]*	Mental health 34
		Work productivity (points)	Costs per point decrease on the WPAI questionnaire subscales#
		Impairment at work UC: 2.0; I: -6.1 [-8.1]*	Impairment at work 28
		Impairment with daily activities UC: 0.0; I: -9.8 [-9.8]*	Impairment with daily activities 23
		Overall work impairment UC: 1.4. I: -6.2 [-7.6]	Overall work impairment 30

Study	Costs Mean [incremental]	Outcomes Mean [incremental]	Incremental cost-effectiveness ratio
Siggaard et al.(55)	UC: NS I: NS [NS]	<u>Body weight reduction (kg)</u> UC: 0.8; I: 4.2 [3.4]*	20 per kg body weight loss
		<u>Reduction in overweight (%)</u> UC: 0.9; I: 5 [4.1]*	17 per 1% reduction in overweight
		<u>Daily carbohydrate intake (g)</u> UC: 3.1; I: 36.0 [32.9]*	2 per gram increase in daily carbohydrate intake
		<u>Daily fat intake (g)</u> UC: -4.0; I: -27.5 [-23.5]*	3 per gram decrease in daily fat intake
Wilson et al.(56)	UC: 140 I-1: 217 [77] I-2: 141 [1] I-3: 204 [64] I-4: 461 [321]	<u>Participants with a cholesterol level reduction of 10%</u> UC: 18.5 I-1: 17.8 [-0.7] I-2: 28 [9.5] I-3: 33 [14.5] I-4: 24.5 [6] *Levels of significance unknown*	Costs per 1% of participants with a cholesterol level reduction of ≥10%# I-1: -110 I-2: 0.1 I-3: 4 I-4: 54
Byers et al.(47)	UC: 33 I: 72 [39]	<u>Cholesterol level reduction (%)</u>	
		6 months UC: 0.4 I: 1.2 [0.8]	*6 months* 48 per 1% cholesterol level reduction#
		12 months UC: 3 I: 6.5% [3.5]*	*12 months* 11 per 1% cholesterol level reduction#

Study	Costs Mean [incremental]	Outcomes Mean [incremental]	Incremental cost-effectiveness ratio
Worksite physical activity and nutrition programs			
Rasu et al.(54)	UC: NS I: NS [49]	<u>Body weight reduction (kg)</u> UC: -0.6; I: 1.3 [1.9]*	26 per kg body weight loss
		<u>Waist circumference reduction (cm)</u> UC: 0.4; I: 2.1[1.7]*	29 per cm waist circumference reduction
		<u>Participants with a weight reduction of ≥5% (%)</u> UC: 6.8; I: 22.6 [15.8]*	3 per 1% of participants with a weight reduction of ≥5%
		<u>Weight efficiency lifestyle (points)</u> Social pressure UC: 1.3; I: 2.5 [1.3] Positive activity UC: 0.8; I: 2.3 [1.5] Availability UC: 2.2; I: 3.2 [1.0] *Levels of significance unknown*	Costs per point improvement on the WEL questionnaire subscales# Social pressure 38 Positive activity 33 Availability 49
Groeneveld et al.(50)	UC: 5048 I: 5399 [351]	<u>Body weight reduction (kg)</u> UC: -1.1; I: 1 [2]*	174 per kg body weight loss from the societal perspective
Van Wier et al.(58)	UC: 3150 I-1: 3597 [447] I-2: 3168 [18]	<u>Body weight reduction (kg)</u> UC: 1.1 I-phone: 1.5 [0.3] I-Internet: 1.9 [0.9]	Costs per kg body weight loss from the societal perspective I-phone: 1282 I-Internet: 20
		<u>QALY</u> UC: 1.85 I-phone: 1.85 [0.001] I-Internet: 1.86 [0.01]	Costs per QALY gained from the societal perspective I-phone: 311523 I-Internet: 1698
Gussenhoven et al.(57)	UC: 4100 I-1: 4469 [369] I-2: 4161 [61]	<u>Body weight reduction (kg)</u> UC: 1.6 I-phone: 1.9 [0.2] I-Internet: 2.4 [0.8]	Costs per kg body weight loss from the employer's perspective I-phone: 1534 I-Internet: 75

Study	Costs Mean [incremental]	Outcomes Mean [incremental]	Incremental cost-effectiveness ratio
Oldenburg et al.(52)	UC: 111 I-RFE: 146 [35] I-BC: 274 [163] I-BCI: 285 [174]	<u>CVD risk unit reduction</u> *6 months* UC: 1.04 I-RFE: 3.99 [2.96]* I-BC: 8.13 [7.09]* I-BCI: 4.16 [4.01]* *12 months* UC: -0.76 I-RFE: 2.79 [3.55] I-BC: 6.10 [6.86]* I-BCI: -0.28 [0.48]	Costs per CVD risk unit reduced *6 months* I-RFE: 12 I-BC: 23 I-BCI: 43 *12 months* I-RFE: 10# I-BC: 24 I-BCI: 363#
Erfurt et al.(48)	UC: 5 I-HE: 28 [23] I-HEC: 50 [45] I-HECE: 61 [56]	<u>CVD risk reduction (%)</u> *High level reduction* UC: 34 I-HE: 35 [1] I-HEC: 44 [10] I-HECE: 46 [12] *Moderate level reduction* UC: 40 I-HE: 41 [1] I-HEC: 51 [11] I-HECE: 56 [16] *Levels of significance unknown*	Costs per 1% of 3 CVD risk factors reduced or prevented *High level reduction* I-HE: 23 I-HEC: 5 I-HECE: 4 *Moderate level reduction* I-HE: 23 I-HEC: 4 I-HECE: 4
Erfurt et al. 2(49)	UC: 27 I-FC: 61 [33] I-HEC: 48 [21] I-HECE: 60 [33]	<u>CVD risk reduction (%)</u> *High level reduction* UC: 35 I-FC: 32 [-3] I-HEC: 44 [9] I-HECE: 45 [10] *Moderate level reduction* UC: 39 I-F: 36 [-3] I-HEC: 48 [9] I-HECE: 51 [12] *Levels of significance unknown*	Costs per 1% of 4 CVD risk factors that reduced or prevented *High level reduction* I-FC: -11 I-HEC: 2 I-HECE: 3 *Moderate level reduction* I-F: -11 I-HEC: 2 I-HECE: 3

Study	Costs Mean [incremental]	Outcomes Mean [incremental]	Incremental cost-effectiveness ratio
Proper et al.(53)	UC: 2591[1] I: 2979 [387][1]	Increase in participants meeting PA recommendations (%) UC: -6; I: -6.6 [-0.6]	-1308 per 1% increase in participants meeting PA recommendations from the employer's perspective
		Increase in energy expenditure (kilocalorie/day) UC: -129; I: 64.2 [193.2]*	7 per extra kilocalorie/day from the employer's perspective
		Decrease in sub-maximal heart rate (beats/minute) UC: -2.5; I: 2.2 [4.7]*	299 per beat/minute decrease in sub-maximal heart rate from the employer's perspective
		Decrease in participants with upper-extremity symptoms (%) UC: 6.2; I: 17.9 [-11.7]	68 per 1% decrease in participants with upper-extremity symptoms from the employer's perspective

UC: Usual care
I: Intervention group
NS: Not stated
kg: Kilogram
CVD: Cardiovascular disease
cm: Centimeter
SF-36: Short Form-36
WPAI questionnaire: Work Productivity and Activity Impairment questionnaire
WEL subscales: Weight Efficacy Lifestyle subscales
QALY: Quality Adjusted Life Year
PA: Physical activity
#: Incremental cost-effectiveness ratio calculated based on the information provided in the article and other related materials
* Significant at $p < 0.05$
[1] Average costs of all participants with complete cost and effect data. Costs included in the cost-effectiveness analyses were variable and depended on the number of participants with complete follow-up data in terms of that outcome measure
Costs are expressed in USD 2010 and rounded to the nearest dollar

Table 4: Main characteristics and outcomes of studies evaluating both costs and reductions in body weight

Study	Focus intervention	Intervention duration	Follow-up duration	Incremental intervention costs (per participant)	Incremental total costs (per participant)	Weight reduction (kilograms)	Incremental cost-effectiveness ratio
Only intervention costs were considered							
Katcher et al.(51)	Diet	22 weeks	22 weeks	226	226	5.3*	43#
Siggaard et al.(55)	Diet	12 weeks	12 weeks	NS	NS	3.4*	20
Rasu et al.(54)	PA & Diet	6 months	6 months	49	49	1.9*	26
Employer's perspective							
Gussenhoven et al.(57)	PA & Diet	6 months	12 months	256 (I-Phone) 227 (I-Internet)	369 (I-Phone) 61 (I-Internet)	0.2 (I-Phone) 0.8 (I-Internet)	1534 (I-Phone) 75 (I-Internet)
Societal perspective							
van Wier et al.(58)	PA & Diet	6 months	24 months	256 (I-Phone) 227 (I-Internet)	447 (I-Phone) 18 (I-Internet)	0.3 (I-Phone) 0.9 (I-Internet)	1282 (I-Phone) 20 (I-Internet)
Groeneveld et al.(50)	PA & Diet	6 months	12 months	730	351	2.0*	174

DISCUSSION

The present review critically appraised and summarized the current evidence on the cost-effectiveness of worksite physical activity and/or nutrition programs. Ten studies (published in 12 articles), evaluating 18 programs, were included in the review. None of the studies evaluated WHP programs aimed solely at increasing physical activity. From various perspectives, all worksite nutrition as well as worksite physical activity and nutrition programs (N=6) were more costly and more effective in reducing body weight compared to usual care during the first years after implementation. If only intervention costs were considered, most worksite nutrition (N=4/5) and worksite physical activity and nutrition programs (N=5/6) were more costly and more effective in reducing cholesterol level and CVD risks, respectively. Currently, however, there are no set levels for how much different stakeholders are willing to pay for reductions in body weight, cholesterol level, and CVD risks. It is therefore unknown whether the costs associated with achieving these results are acceptable, ie, whether these programs are cost-effective. Therefore, it is up to individual decision-makers to judge whether or not these programs offer good value for money. CEA were also conducted in terms of various other outcome measures (eg, dietary habits, quality of life, physical activity-related outcome measures, and work-related outcome measures). However, ICER in terms of these outcome measures were only calculated for one intervention. Furthermore, only one study evaluated the cost-utility of worksite physical activity and nutrition programs and provided mixed results. When compared to the National Institute for Health and Clinical Excellence (NICE) threshold of GBP20,000 (±USD30 500) to GBP30 000 (±USD45 700) per QALY gained as well as the frequently cited US threshold of USD50 000–100 000 per QALY gained (59), the internet-based intervention of the study can be regarded as cost-effective ($1698 per QALY gained), whereas the phone-based intervention ($311 523 per QALY gained) cannot. All in all, these findings do not necessarily support the conclusion of a previous review (36) that the literature provides “guarded cautious optimism” about the cost-effectiveness of WHP programs.

When only intervention costs were considered, the additional costs per kilogram body weight loss ranged from $20–43, independent of the program focus (ie, nutrition or physical activity and nutrition). From a broader perspective, intervention costs were partially offset by a reduction in medical and/or absenteeism costs. Strikingly, this did not result in lower ICER as the three programs evaluated by only including intervention costs were equally or more effective in reducing body weight compared to those evaluated from a broader perspective, whereas their intervention costs were similar or lower. This superior effectiveness might be explained by the fact that these studies conducted follow-up measurements immediately after the intervention period (<6 months) as opposed to several months after the completion of the program (≥6 months) in the studies performed from a broader perspective. Systematic reviews show that (partial) weight rebound after the intervention period is common (60, 61). Another explanation may be the non-randomized design of two of these three studies (ie, results may be confounded by selection bias) (40). Nevertheless, it would also be insightful to investigate the relationship between intervention costs, which are strongly related to intervention composition and intensity, and effect size in more detail. If it is established that more costly programs do not necessarily produce better health outcomes or cost-savings, cost containment strategies during the design phase of a program may be a useful strategy to optimize cost-effectiveness.

A risk of bias assessment revealed that most of the included studies had several methodological shortcomings. For example, few studies specifically stated their perspective and an incremental analysis of costs and effects were not performed in all studies. Furthermore, many studies applied a rather restrictive perspective by only including intervention costs in their cost estimate. However, as WHP programs are thought to be associated with other cost categories (eg, medical and productivity-related costs) (27), the adoption of a broader perspective is recommended. Costs were only measured in physical units in four studies, and of these, only one valued them appropriately by calculating them based on depleted sources and stating their reference year. Furthermore, although research has indicated that presenteeism accounts for a larger proportion of productivity-related losses compared to absenteeism, none of the studies conducted from the societal and/or employer's perspective included presenteeism costs in their cost estimates. This likely resulted from the fact that a "gold standard" for measuring and valuing presenteeism does not exist. Nevertheless, up until now, various instruments have been developed to measure presenteeism, of which several capture lost productivity suitable for direct translation into a monetary unit (62–64). In addition, although economic analyses require that assumptions are made (30, 65), few studies conducted a sensitivity analysis and hardly any of the studies reported on the uncertainty around their ICER. Sensitivity analyses are useful to test the robustness of the study results, but do not give insight into the uncertainty due to sampling variation (30, 66, 67). To quantify precision, non-parametric bootstrapping can be used as a statistical technique for dealing with the highly skewed nature of cost data (30, 65) and the uncertainty around an ICER can be illustrated graphically using cost-effectiveness planes (30). It is also important to mention that three studies did not even report on the uncertainty around their effects. Economic evaluations rely heavily on the assessment of the clinical effectiveness (30). Not reporting on the uncertainty around the effects strongly hampers the interpretation of the reported ICER. Using results of economic evaluations with a high risk of bias for deciding how resources should be optimally allocated, may lead to inappropriate decisions (40, 65). Therefore, strong conclusions about the cost-effectiveness of worksite physical activity and/or nutrition programs cannot be made due to the methodological shortcomings of the included studies. This should be addressed in future studies. In particular, future studies should include presenteeism costs and emphasis should be placed on the handling of uncertainty.

One of the main strengths of this review was that it incorporated a risk of bias assessment using a standardized quality checklist based on consensus among experts in the field of economic evaluations. Furthermore, four additional studies were identified compared to previous reviews on the cost-effectiveness of WHP programs, all of which evaluated costs and kilogram body weight loss. As a result, the present review was the first to compare ICER in terms of costs per kilogram body weight loss from different perspectives. However, due to heterogeneity of outcome measures, follow-up (long- versus short-term), and perspectives, results could not be pooled. As a result of the relatively limited number of included studies, it was also not possible to conduct subgroup analyses to investigate the impact of program format [ie, (self-) assessment, educational/informational, behavioral, exercise, environmental, and incentive components] or participant characteristics (eg, age, gender, and white-versus blue-collar workers) on the cost-effectiveness of the interventions. Therefore, the present review cannot indicate which program formats are important for attaining cost-effectiveness or how worksite physical activity and/or nutrition programs should optimally be designed. Furthermore, a program's cost-effectiveness may depend on the characteristics of its participants. Blue-collar workers, for example, may respond differently compared to white-collar workers as a result of their difference in underlying health risks (68). It is important to address these issues in future reviews when additional research on the cost-effectiveness of worksite physical activity and/or nutrition programs has been completed. Another limitation of the present review was the possible effect of publication bias. That is, economic evaluations may be more likely to be conducted of interventions that had previously been found to be effective, and studies with favorable results may be more likely to be published. It is also important to bear in mind that all CEA conducted from the employer's perspective were performed in The Netherlands. These results are not necessarily generalizable to other countries, as their health and social security systems may differ. US employers, for example, bear a large part of the medical costs of their employees, whereas in Europe these accrue to the government and/or insurance companies (28). Furthermore, only trial-based economic evaluations with relatively short follow-ups (≤3 years) were identified and included. As cost-savings due to improved health might only occur after a longer period, this may have resulted in an underestimation of a possible absenteeism and/or medical cost-offset effect. Due to their relatively short follow-ups, studies were also only able to assess the programs' cost-effectiveness in terms of intermediate outcome measures relating to aspects of diet and physical activity (eg, CVD risk, body weight, and cholesterol level reduction), whereas disease prevention (eg, CVD, diabetes type 2) can be regarded as the primary endpoint of worksite physical activity and/or nutrition programs (69). To bridge the gap between what has been observed in the trial-based economic evaluations and what the cost-effectiveness of worksite physical activity and/or nutrition programs would be over a longer time horizon, decision analytic modelling could be used (30). However, currently little is known about the longevity of the intermediate outcomes of WHP programs and the relationship of these outcomes with changes in long-term medical and productivity-related costs. More research should therefore be done in this field to allow for the development of credible decision analytic models.

Conclusion

Current evidence indicates that worksite physical activity and/or nutrition programs can result in reductions in body weight, cholesterol level, and CVD risks, but at a higher cost than usual care. Because it is unknown how much decision-makers are willing to pay for these health outcomes, conclusions about their cost-effectiveness cannot be made. Most of the included studies had several methodological shortcomings, which hinders the validity of their results. Therefore, there is substantial need for improvement of the methodological quality of studies evaluating the cost-effectiveness of worksite physical activity and/or nutrition programs and particular emphasis should be placed on the handling of uncertainty.

Acknowledgements

We thank Marijke Mol for her contribution in developing the search strategy.

REFERENCES

1. Hankinson AL, Daviglus ML, Bouchard C, Carnethon M, Lewis CE, Schreiner PJ, et al. Maintaining a high physical activity level over 20 years and weight gain. *JAMA*. 2010;304(23):2603–10.
2. Jakicic JM, Otto AD. Physical activity considerations for the treatment and prevention of obesity. *Am J Clin Nutr*. 2005;82(1):226S–9S.
3. Kumanyika SK, Obarzanek E, Stettler N, Bell R, Field AE, Fortmann SP, et al. Population-based prevention of obesity. The need for comprehensive promotion of healthful eating, physical activity, and energy balance. A scientific statement from American heart association council on epidemiology and prevention, interdisciplinary committee for prevention (formerly the expert panel on population and prevention science). *Circulation*. 2008;118(4):428–64.
4. Mozaffarian D, Hao T, Rimm EB, Willett WC, Hu FB. Changes in diet and lifestyle and long-term weight gain in women and men. *N Engl J Med*. 2011;364(25):2392–404.
5. World health organization. *Diet, nutrition and the prevention of chronic diseases: Report of a Joint WHO/FAO Expert Consultation*. Geneva: World health organization; 2003.
6. Bennett GG, Wolin KY, Puleo EM, Masse LC, Atienza AA. Awareness of national physical activity recommendations for health promotion among US adults. *Med Sci Sports Exerc*. 2009;41(10):1849–55.
7. Krebs-Smith SM, Cleveland LE, Ballard-Barbash R, Cook DA, Kahle LL. Characterizing food intake patterns of American adults. *Am J Clin Nutr*. 1997;65(4):1264S–8S.
8. Martínez-González MA, Varo JJ, Santos JL, De Irala J, Gibney M, Kearney J, et al. Prevalence of physical activity during leisure time in the European Union. *Med Sci Sports Exerc*. 2001;33(7):1142–6.
9. Pronk NP, Anderson LH, Crain AL, Martinson BC, O'Connor PJ, Sherwood NE, et al. Meeting recommendations for multiple healthy lifestyle factors: Prevalence, clustering, and predictors among adolescent, adult, and senior health plan members. *Am J Prev Med*. 2004;27(2, Supplement 1):25–33.
10. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilbert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc*. 2008;40(1):181–8.
11. Troiano RP, Macera CA, Ballard-Barbash R. Be physically active each day. How can we know? *J Nutr*. 2001;131(2):451S–60S.
12. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999-2000. *JAMA*. 2002;288(14):1723–7.
13. Adult overweight and obesity in the European Union (EU27) based on measured height and weight [Internet]. International organisation for the study of obesity. May 2011 [Cited May 2011]. Available from: http://www.oaso.org/site_media/uploads/Adult_EU_27_May_2011.pdf
14. Trogdon JG, Finkelstein EA, Hylands T, Dellea PS, Kamal-Bahl SJ. Indirect costs of obesity: a review of the current literature. *Obes Rev*. 2008;9(5):489–500.
15. Yach D, Stuckler D, Brownell KD. Epidemiologic and economic consequences of the global epidemics of obesity and diabetes. *Nat Med.* 2006;12(1):62–6.
16. Thompson D, Edelsberg J, Colditz GA, Bird AP, Oster G. Lifetime health and economic consequences of obesity. *Arch Intern Med*. 1999;159(18):2177–83.

17. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Annu Rev Public Health*. 2008;29:303–23.

18. National Institute for Health and Clinical Excellence (NICE). *Workplace health promotion: How to encourage employees to be physically active*. London: NICE; 2008. NICE Public Health Guidance 13.

19. NiMhurchu C, Aston L, Jebb S. Effects of worksite health promotion interventions on employee diets: a systematic review. *BMC Public Health*. 2010;10:62.

20. Anderson LM, Quinn TA, Glanz K, Ramirez G, Kahwati LC, Johnson DB, et al. The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: A systematic review. *Am J Prev Med.* 2009;37(4):340–57.

21. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Lifestyle-focused interventions at the workplace to reduce the risk of cardiovascular disease--a systematic review. *Scand J Work Environ Health*. 2010;36(3):202–15.

22. Verweij LM, Coffeng J, van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. *Obes Rev*. 2011;(12):406–29.

23. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. *N Engl J Med*. 1977;296(13):716–21.

24. Burdorf A. Economic evaluation in occupational health—its goals, challenges, and opportunities. Scand J Work Environ Health. 2007;33(3):161–4.

25. Leigh JP. Expanding research on the economics of occupational health. *Scand J Work Environ Health*. 2006;32(1):1–4.

26. Miller P, Haslam C. Why employers spend money on employee health: Interviews with occupational health and safety professionals from British Industry. *Safety Science*. 2009;47(2):163–9.

27. Nicholson S, Pauly MV, Polsky D, Baase CM, Billotti GM, Ozminkowski RJ, et al. How to present the business case for healthcare quality to employers. *Appl Health Econ Health Policy* 2005;4(4):209–18.

28. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev*. 2011;12(12):1031–49.

29. Oster G, Thompson D, Edelsberg J, Bird AP, Colditz GA. Lifetime health and economic benefits of weight loss among obese persons. *Am J Public Health* 1999;89(10):1536–42.

30. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes. 3rd edition.* New York: Oxford University Press; 2005.

31. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1991-1993 update. *Am J Health Promot*. 1993;8(1):50–62.

32. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1993-1995 update. *Am J Health Promot*. 1996;10(5):380–8.

33. Pelletier KR. Clinical and cost outcomes of multifactorial, cardiovascular risk management interventions in worksites: a comprehensive review and analysis. *J Occup Environ Med*. 1997;39(12):1154–69.

34. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1995-1998 update (IV). *Am J Health Promot*. 1999;13(6):333–45, iii.

35. Pelletier KR. A review and analysis of the clinical- and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1998-2000 update. *Am J Health Promot.* 2001;16(2):107–16.

36. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VI 2000-2004. *J Occup Environ Med*. 2005;47(10):1051–8.

37. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: Update VII 2004-2008. *J Occup Environ Med*. 2009;51(7):822–37.

38. Proper KI, Van Mechelen W. *Effectiveness and economic impact of worksite interventions to promote physical activity and healthy diet Background paper prepared for the WHO/WEF Joint Event on Preventing Non-communicable Diseases in the Workplace* (Dalian/China, September 2007). Geneva: WHO press; 2008.

39. Soler RE, Leeks KD, Razi S, Hopkins DP, Griffith M, Aten A, et al. A systematic review of selected interventions for worksite health promotion: The assessment of health risks with feedback. *Am J Prev Med*. 2010;38(2, Supplement 1):S237–S262.

40. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from: http://cochrane-handbook.org.

41. Jefferson T, Demicheli V, Vale L. Quality of systematic reviews of economic evaluations in health care. *JAMA*. 2002;287(21):2809–12.

42. Chapman LS. Meta-evaluation of worksite health promotion economic return studies: 2005 update. *Am J Health Promot*.2005;19(6):1–11.

43. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria. *Int J Technol Assess Health Care*. 2005;21(02):240–5.

44. Gerkens S, Crott R, Cleemput I, Thissen JP, Closon MC, Horsmans Y, et al. Comparison of three instruments assessing the quality of economic evaluations: A practical exercise on economic evaluations of the surgical treatment of obesity. *Int J Technol Assess Health Care*. 2008;24(03):318–25.

45. U.S. Bureau of Labor Statistics [Internet]. [Cited 2010 December 1]. Available from: http://www.bls.gov.

46. Organisation for Economic Co-operation and Development (OECD). StatExtracts. OECD Selected Data [Internet]. [Cited December 1 2010]. Available from: http://stats.oecd.org/Index.aspx.

47. Byers T, Mullis R, Anderson J, Dusenbury L, Gorsky R, Kimber C, et al. The costs and effects of a nutritional education program following work-site cholesterol screening. *Am J Public Health*. 1995;85(5):650–5.

48. Erfurt JC, Foote A, Heirich MA. The cost-effectiveness of worksite wellness programs for hypertension control, weight loss, and smoking cessation. *J Occup Med*. 1991;33(9):962–70.

49. Erfurt J, Foote A, Heinrich M. The cost-effectiveness of worksite wellness programs for hypertension control, weight loss, smoking cessation, and exercise. *Pers Psychol*. 1992;45(1):5–27.
50. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med*. 2011;53(6):610–17.
51. Katcher HI, Ferdowsian HR, Hoover VJ, Cohen JL, Barnard ND. A Worksite vegan nutrition program is well-accepted and improves health-related quality of life and work productivity. *Ann Nutr Metab*. 2010;56(4):245–52.
52. Oldenburg B, Owen N, Parle M, Gomel M. An economic evaluation of four work site based cardiovascular risk factor interventions. *Health Educ Behav*. 1995;22(1):9–19.
53. Proper KI, de Bruyne MC, Hildebrandt VH, van der Beek AJ, Meerding WJ, van MW. Costs, benefits and effectiveness of worksite physical activity counseling from the employer's perspective. *Scand J Work Environ Health*. 2004;30(1):36–46.
54. Rasu RS, Hunter CM, Peterson AL, Maruska HM, Foreyt JP. Economic evaluation of an Internet-based weight management program. Am J Manag Care. 2010;16(4):e98–104.
55. Siggaard R, Raben A, Astrup A. Weight loss during 12 week's ad libitum carbohydrate-rich diet in overweight and normal-weight subjects at a Danish work site. *Obes Res*. 1996;4(4):347–56.
56. Wilson MG, Edmunson J, DeJoy DM. Cost Effectiveness of work-site cholesterol screening and intervention programs. *J Occup Environ Med*. 1992;34(6):642–9.
57. Gussenhoven AHM, van Wier MF, Bosmans JE, Dekkers JC, van Mechelen W. Economic evaluation of a distance counselling lifestyle programme among overweight employees from a company perspective, Alife@Work: a randomized controlled trial. *Work*. In press 2012
58. van Wier MF, Dekkers JC, Bosmans JE, Heymans MW, Hendriksen IJM, Pronk NP et al. Economic evaluation of a weight control program with e-mail and telephone counseling among overweight employees: a randomized controlled trial. *Int J Behav Nutr Phys Act.* In press 2012.
59. Ubel PA, Hirth RA, Chernew ME, Fendrick AM. What is the -price of life and why doesn't it increase at the rate of inflation? *Arch Intern Med*. 2003;163(14):1637–41.
60. Barte JCM, Ter Bogt NCW, Bogers RP, Teixeira PJ, Blissmer B, Mori TA, et al. Maintenance of weight loss after lifestyle interventions for overweight and obesity, a systematic review. *Obes Rev.* 2010;11(12):899–906.
61. Loveman E, Frampton GK, Shepherd J, Picot J, Cooper K, Bryant J, et al. The clinical effectiveness and cost-effectiveness of long-term weight management schemes for adults: a systematic review. *Health Technol Assess.* 2011;15(2):1–182.
62. Mattke S, Balakrishnan A, Bergamo G, Newberry SJ. A review of methods to measure health-related productivity loss*. Am J Manag Care.* 2007;13(4):211–7.
63. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res.* 2005;5(1):23–8.
64. Koopmanschap M, Meeding WJ, Evers S, Severens J, Burdorf A, Brouwer W. *Handleiding voor het gebruik van PRODISQ versie 2.1.* [Handbook on use of PRODISQ.] Rotterdam/Maastricht, Erasmus MC - Instituut voor Medical Technology Assessment, Instituut Maatschappelijke Gezondheidszorg, Universiteit van Maastricht - Beleid Economie en Organisatie van de Zorg; 2004.

65. Uegaki K, de Bruijne MC, Lambeek L, Anema JR, van der Beek AJ, van MW, et al. Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. *Scand J Work Environ Health*. 2010;36(4):273–88.

66. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technol Assess*. 1999;3(2):1–134.

67. Sendi P, Gafni A, Birch S. Opportunity costs and uncertainty in the economic evaluation of health care interventions. *Health Econ.* 2002;11:23–31.

68. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Aff (Millwood).* 2010;29(2):304–11.

69. Prentice RL, Willett WC, Greenwald P, Alberts D, Bernstein L, Boyd NF, et al. Nutrition and physical activity and chronic disease prevention: research strategies and recommendations. *J Natl Cancer Inst.* 2004;96(17):1276–87.

70. Levin SM, Ferdowsian HR, Hoover VJ, Green AA, Barnard ND. A worksite programme significantly alters nutrient intakes. *Public Health Nutr.* 2010;13(10):1629–35.

Appendix 1: EMBASE search strategy

Combined search	(#1 AND #2 AND #3 AND #4) NOT #5
#1 Intervention type	'health promotion'/exp OR 'harm reduction'/exp OR 'high risk behavior'/exp OR 'risk reduction'/exp OR 'health behavior'/de OR 'primary prevention'/exp OR 'secondary prevention'/exp OR 'occupational health'/exp OR health:ab,ti OR intervention:ab,ti OR 'life style':ab,ti OR lifestyle:ab,ti OR prevention:ab,ti OR preventive:ab,ti OR 'risk factor':ab,ti OR 'risk factors':ab,ti NOT 'rehabilitation'/exp
#2 Intervention aim	'fitness'/exp OR 'exercise'/exp OR 'physical activity'/exp OR 'sport'/exp OR fitness:ab,ti OR exercis*:ab,ti OR sport*:ab,ti OR 'physical activity':ab,ti OR 'diet'/exp OR 'nutrition'/exp OR diet*:ab,ti OR nutrition*:ab,ti OR food:ab,ti OR vegetable*:ab,ti OR fruit*:ab,ti OR 'weight reduction'/exp OR 'cholesterol'/exp OR 'hypertension'/exp OR cholesterol:ab,ti OR hypertensi*:ab,ti
#3 Participant/setting type	'manpower'/exp OR 'workplace'/exp OR employ*:ab,ti OR worker*:ab,ti OR workplace*:ab,ti OR 'work site':ab,ti OR personnel*:ab,ti OR workforce:ab,ti OR staff:ab,ti
#4 Study design	'economic evaluation'/exp OR 'economic evaluation':ab,ti OR 'economic analysis':ab,ti OR (cost:ab,ti OR costs:ab,ti AND (benefit*:ab,ti OR utilit*:ab,ti OR effective*:ab,ti OR minimi?ation:ab,ti)) OR ROI:ab,ti OR "return on investment":ab,ti
#5 Limits	'newborn'/exp OR 'child'/exp OR 'adolescent'/exp NOT 'adult'/exp

3

Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity

Johanna M van Dongen
Karin I Proper
Marieke F van Wier
Allard J van der Beek
Paulien M Bongers
Willem van Mechelen
Maurits W van Tulder

Obes Rev 2011, 12: 1031-1048

ABSTRACT

This systematic review summarizes the current evidence on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. Data on study characteristics and results were extracted from 18 studies published up to 14 January 2011. Two reviewers independently assessed the risk of bias of included studies. Three metrics were (re-)calculated per study: the net benefits, benefit cost ratio (BCR) and return on investment (ROI). Metrics were averaged, and a *post hoc* subgroup analysis was performed to compare financial return estimates between study designs. Four randomized controlled trials (RCTs), 13 non-randomized studies (NRSs) and one modelling study were included. Average financial return estimates in terms of absenteeism benefits (NRS: ROI 325%, BCR 4.25; RCT: ROI -49%, BCR 0.51), medical benefits (NRS: ROI 95%, BCR 1.95; RCT: ROI -112%, BCR -0.12) or both (NRS: ROI 387%, BCR 4.87; RCT: ROI -92%, BCR 0.08) were positive in NRSs, but negative in RCTs. Worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity generate financial savings in terms of reduced absenteeism costs, medical costs or both according to NRSs, whereas they do not according to RCTs. Since these programmes are associated with additional types of benefits, conclusions about their overall profitability cannot be made.

INTRODUCTION

An imbalance between energy intake (nutrition) and output (physical activity) among the population has led to an increased prevalence of overweight, obesity, and their attributable diseases (e.g. type 2 diabetes, and cardiovascular disease) (1). Nowadays, 33.8% of US adults are obese (body mass index ≥ 30) and the combined prevalence of overweight and obesity is 68.0% (body mass index ≥ 25) (2). In the UK, the combined prevalence of overweight and obesity is 57% in adult women and 65% in adult men (3).

Next to the toll that overweight and obesity take on the health and well-being of individuals, they impose a substantial economic burden in terms of healthcare costs and lost productivity (1,4–7). For example, obesity-related medical payments are estimated to account for 5% of health insurance expenditures among US businesses with employer-provided health insurance (5). Moreover, the estimated US national costs of obesity attributable absenteeism range from $3.38 billion to $6.38 billion per year (6).

Employers bear the financial consequences of reduced productivity. In countries with employer-provided health insurance (e.g. the US), they also bear a large part of the financial consequences of increased medical spending. Therefore, employers may financially benefit from implementing worksite health promotion programmes (WHP programmes) aimed at weight gain prevention among their workforce by improving nutrition and/or increasing physical activity (8). In addition, the worksite provides a useful setting for implementing these programmes since employees spend the majority of their waking hours at the worksite (9), large enterprises often have the infrastructure available to offer such programmes at relatively low costs (10), and organizational and social support can be made available when behaviour change efforts are attempted (11).

Worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity were previously found to be effective in reducing body fat and body weight (12–14). Employers, however, may like to know whether these programmes generate a positive financial return. A useful way for communicating the financial ramifications of a given programme is a 'return on investment' analysis

(ROI analysis), a form of investment analysis often used in business administration in which programme costs are compared to its resulting financial benefits (15).
Several efforts have been undertaken to summarize the literature on the financial return of WHP programmes (8,9,16,17). Estimated financial returns, as defined by averted medical costs, productivity-related costs or both, ranged from \$1.4 to \$4.6 per dollar spent (8,17). Furthermore, medical costs were found to decrease by \$3.3, and absenteeism costs by \$2.7 per dollar spent (9). Most of these reviews, however, did not adjust for the different methodologies used in the included studies to estimate the financial return and a risk of bias assessment was often missing. Furthermore, these reviews focused on WHP programmes in general, instead of programmes aimed at improving nutrition and/or increasing physical activity in particular. Therefore, the present study aimed to critically appraise and summarize the current evidence on the financial return of WHP programmes aimed at improving nutrition and/or increasing physical activity, compared to usual care (including no intervention) or a cut-down version of the programme.

METHODS

Inclusion criteria

English, Dutch, German and French-written studies evaluating the financial return of WHP programmes aimed at improving nutrition and/or increasing physical activity in the working population were eligible for inclusion. The WHP programme should be compared to usual care (including no intervention) or a cut-down version of the programme. Studies should contain a ROI analysis, assessing and presenting both programme costs and its resulting benefits. Benefits, defined as programme outcomes converted to monetary values, should be directly measured or modelled based on primary data. Benefits related to WHP programmes are mostly defined in terms of averted medical and productivity-related costs (18). Examples of productivity-related costs are costs associated with absenteeism and reduced productivity at work (presenteeism) (18). No limitations were set as to the perspective of the ROI analysis (e.g. employer’s and societal perspective), programme format (e.g. assessment, counselling and exercise programme), worksite characteristics (e.g. age, gender, occupation, proportion of full-time employees and number of employees) and follow-up duration. Studies targeting employees with chronic conditions (e.g. diabetes and cardiovascular diseases), long-term sick-listed employees, retirees or children were excluded.

Search strategy

To identify relevant studies, eight electronic databases (EMBASE, MEDLINE, SPORTDiscus, PsycINFO, NIOSHTIC-2, NHSEED, HTA and Econlit) were searched for studies published from inception to 14 January 2011. An information specialist of the VU University Medical Center was consulted to develop and run the search strategy. Databases were searched on participant/setting type (e.g. 'Workplace', 'Employee' and 'Workforce'), intervention type (e.g. 'Health Promotion', 'Lifestyle'), intervention aim (e.g. 'Exercise', 'Physical Activity', 'Nutrition' and 'Diet') and study design (e.g. 'Return on Investment', 'Cost Effectiveness'). A broad search strategy was used so that the results could be used for both the present study and a review on the cost-effectiveness of WHP programmes aimed at improving nutrition and/or increasing physical activity (van Dongen et al., unpublished data). An example of the EMBASE search can be found in Table 1. The electronic search was supplemented by searching references of relevant review articles (9–12,16,17,19–26) and those of the retrieved full texts. Articles were also identified from the authors' own literature databases. To identify unpublished studies, authors of included studies which were published during the last decade, were contacted. During the search, a 'search diary' was maintained consisting of keywords used, searched databases and search results. Titles and abstracts of the retrieved studies were stored in an electronic database using Reference Manager 11.0 (ISI Research Soft Inc., Berkeley, CA, USA).

Study selection

On the basis of abstracts and titles, two reviewers (J. v. D. and K. P.) independently determined the eligibility of the retrieved studies. If studies met the inclusion criteria or uncertainty remained about inclusion, full texts were retrieved. All full texts were read and checked for eligibility. To resolve disagreements between the two reviewers regarding inclusion of a study, a consensus procedure was used. A third reviewer (M. v. W.) was consulted when disagreements persisted; this was necessary in two occasions.

Table 1: EMBASE search strategy

Combined search	(#1 AND #2 AND #3 AND #4) NOT #5
#1 Intervention type	‘health promotion’/exp OR ‘harm reduction’/exp OR ‘high risk behavior’/exp OR ‘risk reduction’/exp OR ‘health behavior’/de OR ‘primary prevention’/exp OR ‘secondary prevention’/exp OR ‘occupational health’/exp OR health:ab,ti OR intervention:ab,ti OR ‘life style’:ab,ti OR lifestyle:ab,ti OR prevention:ab,ti OR preventive:ab,ti OR ‘risk factor’:ab,ti OR ‘risk factors’:ab,ti NOT ‘rehabilitation’/exp
#2 Intervention aim	‘fitness’/exp OR ‘exercise’/exp OR ‘physical activity’/exp OR ‘sport’/exp OR fitness:ab,ti OR exercis*:ab,ti OR sport*:ab,ti OR ‘physical activity’:ab,ti OR ‘diet’/exp OR ‘nutrition’/exp OR diet*:ab,ti OR nutrition*:ab,ti OR food:ab,ti OR vegetable*:ab,ti OR fruit*:ab,ti OR ‘weight reduction’/exp OR ‘cholesterol’/exp OR ‘hypertension’/exp OR cholesterol:ab,ti OR hypertensi*:ab,ti
#3 Participant/setting type	‘manpower’/exp OR ‘workplace’/exp OR employ*:ab,ti OR worker*:ab,ti OR workplace*:ab,ti OR ‘work site’:ab,ti OR personnel*:ab,ti OR workforce:ab,ti OR staff:ab,ti
#4 Study design	‘economic evaluation’/exp OR ‘economic evaluation’:ab,ti OR ‘economic analysis’:ab,ti OR (cost:ab,ti OR costs:ab,ti AND (benefit*:ab,ti OR utilit*:ab,ti OR effective*:ab,ti OR minimi?ation:ab,ti)) OR ROI:ab,ti OR “return on investment”:ab,ti
#5 Limits	‘newborn’/exp OR ‘child’/exp OR ‘adolescent’/exp NOT ‘adult’/exp

Data extraction

Data were extracted on study design (e.g. perspective, research design, setting and follow-up duration), characteristics of the study population (e.g. participants and job characteristics), programme focus (e.g. improving nutrition, increasing physical activity or both), programme format (e.g. assessment, educational/informational, behavioural, exercise, environmental and incentive components), measurement and valuation methods of costs and benefits and study results (e.g. reported costs, benefits and ROI outcomes). One reviewer (J. v. D.) extracted data using a pre-designed data extraction form. Ten percent of the extracted data was checked by a second reviewer (K. P.). No disagreements were identified between reviewers. If articles did not contain sufficient information on study results, authors were contacted for additional information. Research designs were classified into three categories (i) randomized controlled trials (RCTs); (ii) non-randomized studies (NRSs) comparing data between an intervention and a self-selected or matched control group and (iii) modelling studies.

Risk of bias assessment

An instrument assessing the risk of bias of ROI analyses does not exist. Therefore, the Consensus Health Economic Criteria list (CHEC list) was used, representing a minimum set of methodological criteria addressing internal and external validity aspects of economic evaluations (27,28). If a CHEC list item was not adequately performed, or if insufficient information about the performance regarding that item was available in the article or in related materials, the item was scored as negative (27). The CHEC list includes six items related to costs and benefits. Costs were defined as programme costs and outcomes as benefits. The CHEC list does not include items for assessing modelling studies. Therefore, two items of the BMJ checklist were added (‘Details of any model used are given’ and ‘The choice of model used and the key parameters on which it is based are justified’) (29). Two reviewers (J. v. D. and K. P.) independently assessed the risk of bias of included studies. If one of the reviewers was a (co-)author of a study, M. v. W. or M. v. T. acted as the second reviewer. A third reviewer (M. v. W. or M. v. T.) was consulted when disagreements remained, which happened three times.

Data synthesis

To provide a complete picture of the financial return, three ROI metrics were (re-) calculated for each intervention evaluated in the included studies: net benefits (NB), benefit cost ratio (BCR) and ROI (30,31).

$$NB = Benefits - Costs$$

$$BCR = \frac{Benefits}{Costs}$$

$$ROI(\%) = \frac{Benefits - Costs}{Costs}[\times 100]$$

Costs were calculated as the difference in programme costs between the intervention and control groups (incremental costs). Benefits were calculated as the difference in monetized outcome measures (e.g. absenteeism and medical costs) between the intervention and control groups during follow-up and, if available, subtracted by their difference before the intervention (incremental benefits). All monetized outcome measures presented in the article and other related materials were included. If a study did not provide incremental costs and benefits, they were calculated based on figures and tables. Consumer price indices (32) and purchasing power parities (33) were used to standardize costs and benefits to annual costs per participant in 2010 US dollars.

Costs and benefits beyond 1 year have to be discounted to correct for the fact that people place greater value on something that they have today than on something that they will have in the future (29,31). However, cost and benefits are usually reported as a total and not per year, making it impossible to apply a discount rate (34). Therefore, discounting was not standardized in this study. For those studies that reported discounted costs and/or benefits as their main results, these were the costs and benefits that were presented and used for the recalculations. For those studies that did not discount costs beyond 1 year, no additional discounting was performed. Since ROI metrics are highly dependable on the number and type of included benefits, benefit-standardized financial return estimates were calculated per intervention. If, e.g. both medical and absenteeism benefits were included in a ROI analysis, three types of benefit-standardized financial return estimates were calculated: including medical benefits, including absenteeism benefits and including both.

Standard deviations of financial return estimates are often lacking (28,34), which makes statistically pooling impossible. To summarize the results of the included studies and to compare the results of the present review with those of previous reviews, BCRs and ROIs were averaged. One reviewer (J. v. D.) carried out the data analyses, which were all checked by a second reviewer (M. v. W.).

Subgroup analysis

A *post hoc* subgroup analysis was performed comparing the average BCRs and ROIs between study designs. In addition, the differences in ROI between study designs were depicted graphically using scatter plots.

RESULTS

Literature search and study selection

The electronic search yielded 3,835 results. After removing 605 duplicates, 3,230 titles and abstracts were screened for inclusion and 47 full texts were retrieved. Thirty-one additional full texts were retrieved after screening references of relevant review articles and the retrieved full texts. After reading those 78 full texts, 16 articles were identified that met the inclusion criteria. Additionally, two unpublished articles were identified from the authors' own databases. Contacting authors of included studies did not yield any results. Eventually, 18 studies were included in the review (Figure 1).

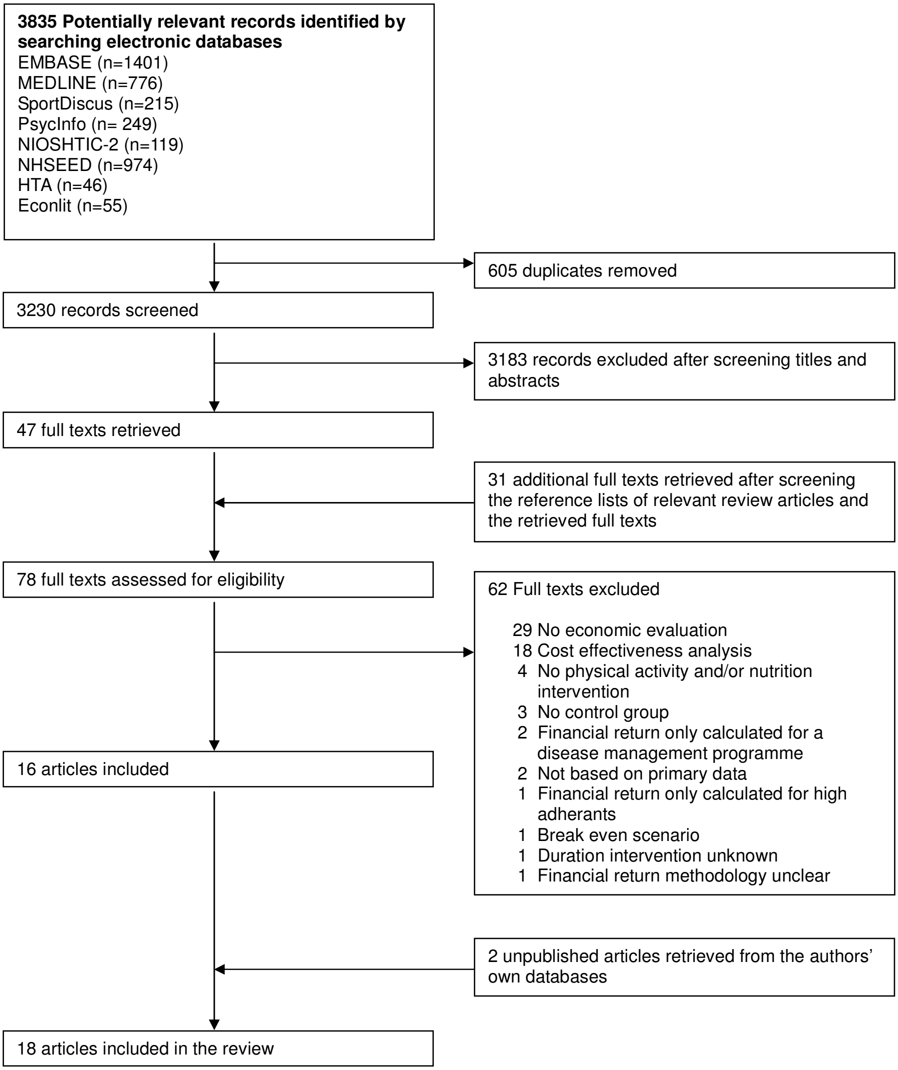

Figure 1: Flow chart for inclusion of studies

Study characteristics

Thirteen NRSs (15 interventions) (35–47), four RCTs (five interventions; (48–50); Gussenhoven et al., unpublished data) and one modelling study (one intervention) (51) were included in the review (Table 2). Ten studies ((40–42;45–50); Gussenhoven et al., unpublished data) were performed from the employer's perspective, indicating that only costs and benefits to the employer were included in the ROI analysis (52). Eight studies (35–39,43,44,51) did not state their perspective. Fourteen studies (35–39,41–47,49,51) were carried out in the USA, three ((48,50); Gussenhoven et al., unpublished data) in the Netherlands and one (40) in the UK. Two studies (38,45) evaluated the financial return of a physical activity intervention and 16 ((35,37,39–44,46–51,53); Gussenhoven et al., unpublished data) that of a comprehensive WHP programme aimed at improving nutrition and increasing physical activity as well as other unhealthy lifestyle behaviours, such as smoking and alcohol consumption. In general, interventions consisted of a (self-)assessment, educational/informational, behavioural, exercise, environmental and/or an incentive component. In the majority of the studies, the control group received no intervention (35–40,42,45–47,51). The length of the interventions varied from 6 months to 5 years (median: 23.7 months, mean: 21.1 months). Financial returns were estimated during the first years after implementation and over a somewhat longer period than the interventions lasted (follow-up: 6 months to 5 years, median: 24 months, mean: 25.1), because four studies ((39,48,50); Gussenhoven et al., unpublished data) had a follow-up beyond the intervention period. Absenteeism benefits were provided by 13 studies (15 interventions; (37,38,40,43–50,53); Gussenhoven et al., unpublished data), medical benefits by 11 studies (13 interventions; (35,38,39,41,42,44,46,48,49,51); Gussenhoven et al., unpublished data), and absenteeism as well as medical benefits by 6 studies (9 interventions; (38,44,46,48,49); Gussenhoven et al., unpublished data). Three of them (three interventions) also provided presenteeism benefits (40,49,51).

Table 2: Characteristics of the included studies (n=18)

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Non-Randomised Studies						
Wood et al.(47)	Perspective: Employer Setting: U.S. 1984 (baseline) 1985-1986 (follow-up) Length intervention: 2 years Follow-up: 2 years	1075 field sales employees I: 688 C: 387	I: <u>Comprehensive health promotion intervention:</u> HRA + feedback, every three months participants are required to complete one of three optional lifestyle activities (fitness, nutrition and weight control, safety, stress management, recreation, relaxation and entertainment, smoking, chemical use, interpersonal relations, cancer prevention, positive thinking, goal setting), incentive programmes, quarterly newsletter C: None	Focus: PA & Diet I: a, b, c, f C: -	*In USD (reference year not stated)* <u>Intervention costs</u> Valued using budget expenditures	<u>Absenteeism benefits</u> Number of days missed because of a health condition based on disability absence data multiplied by an average wage rate
Shore et al.(45)	Perspective: Employer Setting: U.S. 1985 (baseline) 1985-1986 (follow-up) Length intervention: 6 months Follow-up: 6 months	Ambulance and management employees I: 134 C: Not stated	I: <u>Physical activity intervention:</u> Fitness assessment + feedback, individual exercise programme, incentive programme C: None	Focus: PA I: a, d, f C: -	*In USD (reference year not stated)* <u>Intervention costs</u> Valued using budget expenditures	<u>Absenteeism benefits</u> Not stated
Schultz et al.(43)	Perspective: Not stated Setting: U.S. 1995 (baseline), 1996-2000 (follow-up) Length intervention: 5 years Follow-up: 5 years	4189 employees of a manufacturing company I: 2596 C: 1593	I: <u>Comprehensive health promotion intervention:</u> HRA, on-site health screening, on-site and telephonic wellness programmes, medical office visit vouchers, a telephonic nurse counselling line, quarterly newsletters C: Quarterly newsletters	Focus: PA & Diet I: a, b, c, f C: b	*In USD (reference year not stated)* <u>Intervention costs</u> Valued using budget expenditures	<u>Absenteeism benefits</u> Number of days missed because of a health condition based on disability absence data, multiplied by an average wage rate

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Aldana et al.(53)	Perspective: Not stated Setting: U.S. 2001-2002 Length intervention: 2 years Follow-up: 2 years	4710 school district employees I: 2401 (1224 (1 year), 1177 (2 years)) C: 2309	I: Comprehensive health promotion intervention: Programme website, health challenges (11 types),incentive programmes C: None	Focus: PA & Diet I: b, c, f C: -	*In USD (reference year not stated)* Intervention costs Valuation method not stated	Absenteeism benefits Actual wages paid to the participants on days missed because of a health condition
Bertera et al.(37)	Perspective: Not stated Setting: U.S. 1984 (baseline), 1985-1986 (follow-up) Length intervention: 2 years Follow-up: 2 years	43888 employees of a manufacturing company I: 29315 C: 14573	I: Comprehensive health promotion intervention: Environmental interventions (healthy food choices (cafeteria/vending machines) and blood pressure and weight machines/scales), health promotion activity committees, HRA + feedback, group sessions and/or individual consultation, 4- to 10- week health educational class cycles, bi-monthly health and fitness magazine, health challenges, incentive programmes C: None	Focus: PA & Diet I: a, b, c, e, f C: -	*In 1986 USD* Intervention costs Micro-costed, valued using tariffs	Absenteeism benefits Actual wages paid to the participants on days missed because of a health condition

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Mills et al.(40)	Perspective: Employer Setting: U.K. 2004 (baseline) 2005 (follow-up) Length intervention: 1 year Follow-up: 1 year	1508 employees of a manufacturing company I: 266 C: 1242	I: <u>Comprehensive health promotion intervention:</u> HRA + feedback, personalised web-portal (articles/ assessments/interactive behaviour change programmes), biweekly tailored emails, newsletters, health promotional literature, 4 on-site seminars C: None	Focus: PA & Diet I: a, b, c C: -	*In USD (reference year not stated)* <u>Intervention costs</u> Valued using budget expenditures	<u>Absenteeism benefits</u> Self-reported days missed because of a health condition, measured by the WHO-HPQ, multiplied by an average wage rate <u>Presenteeism benefits</u> Self-reported on-the-job productivity, measured by the WHO-HPQ, multiplied by an average wage rate
Gettman(38)	Perspective: Not stated Setting: U.S. 1982-1983 Length intervention: 2 years Follow-up: 2 years	778 (1982) & 707 (1983) employees of an oil and gas exploitation and production company <u>1982</u> <u>1983</u> I: 453 I: 442 C: 325 C: 265	I: <u>Physical activity intervention:</u> Exercise facility (headquarters), health club membership, or home exercise programmes (field offices) C: None	Focus: PA I: d C: -	*In USD (reference year not stated)* <u>Intervention costs</u> Valued using budget expenditures	<u>Absenteeism benefits</u> Actual wages paid to the participants on days missed because of a health condition <u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Shi(44)	Perspective: Not stated Setting: U.S. 1988 (baseline), 1989 (follow-up) Length intervention: 1 year Follow-up: 1 year	1188 utility company employees I-Low: 301 I-Medium: 295 I-High: 180 C: 412	1. I-Low: <u>Comprehensive health promotion intervention:</u> HRA, bimonthly newsletter, health resource centre, self-care books 2. I-Medium: <u>Comprehensive health promotion intervention:</u> HRA, bimonthly newsletter, health resource centre, self-care books, behaviour change workshops/ classes, Division Health Wise team 3. I-High: <u>Comprehensive health promotion intervention:</u> HRA, bimonthly newsletter, health resource centre, self-care books, behaviour change workshops/ classes, Division Health Wise team, case management, environmental policy C: HRA, bimonthly newsletter	Focus: PA & Diet I-Low: a, b I-Medium: a, b, c I-High: a, b, c, e C: a, b	*In 1988 USD* <u>Intervention costs</u> Micro-costed, valued using tariffs	<u>Absenteeism benefits</u> Self-reported days missed because of a health condition, multiplied by an average wage rate <u>Medical benefits</u> Self-reported health care utilization, multiplied by cost prices
Stave et al.(46)	Perspective: Employer Setting: U.S. 1996 (baseline), 1998-2000 (follow-up) Length intervention: 3 years Follow-up: 3 years	3962 employees of a pharmaceutical company I: 1275 C: 2687	I: <u>Comprehensive health promotion intervention:</u> Contracts for good health and programme attendance, self-assessment to measure "readiness for change", on-site health education seminars, marketing strategies C: None	Focus: PA & Diet I: a, b, f C: -	*In 2000 USD* <u>Intervention costs</u> Valuation method not stated	<u>Absenteeism benefits</u> Actual wages paid to the participants on days missed because of a health condition. <u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Ozminkowski et al.(42)	Perspective: Employer Setting: U.S. 1994-1997 (baseline & follow-up) Length intervention: 23,33 months Follow-up: 23,33 months	22838 bank employees I: 11194 C: 11644	I: <u>Comprehensive health promotion intervention:</u> HRA - Low risk HRA participants: Feedback, general health education and self-care materials, programme telephone line - High risk HRA participants: Feedback, 3 additional HRA questionnaires + feedback, recommendations for action, health education materials, books, and videos, telephone counselling, audio library and health tapes C: None	Focus: PA & Diet I: a, b, c C: -	*In 1996 USD* <u>Intervention costs</u> Valued using budget expenditures	 <u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records
Naydeck et al.(41)	Perspective: Employer Setting: U.S. 2001 (baseline), 2002-2005 (follow-up) Length intervention: 4 years Follow-up: 4 years	3784 insurance company employees I: 1892 C: 1892	I: <u>Comprehensive health promotion intervention:</u> HRA, biometric screening, on-line health education programmes, health education classes, telephone counselling, individual coaching, biometric screening, various 6- to 12-week campaigns to increase fitness participation and awareness of disease preventions strategies, fitness centre, health promotion campaigns, incentive programmes C: None	Focus: PA & Diet I: a, b, c, d, e, f C: -	*In 2005 USD* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	 <u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Aldana et al.2(35)	Perspective: Not stated Setting: U.S. 1987-1988 (baseline), 1989-1990 (follow-up) Length intervention: 2 years Follow-up: 2 years	680 city employees I: 340 C: 340	I: <u>Comprehensive health promotion intervention:</u> HRA, counselling, contracts, incentive programmes, additional medical examinations for high risk subjects, one hour exercise instruction class, monthly health seminars, bimonthly newsletter C: None	Focus: PA & Diet I: a, b, c, d, f C: -	*In 1990 USD* <u>Intervention costs</u> Valuation method not stated	<u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records
Gibbs et al.(39)	Perspective: Not stated Setting: U.S. 1978-1982 Length intervention: 6 months, optional to participate more than once Follow-up: 57 months	1559 employees of a health insurance company I: 667 C: 892	I: <u>Comprehensive health promotion intervention:</u> Publicity and health risk education, HRA + feedback & referral, telephone follow-up, group programmes (nutrition, weight reduction, smoking cessation & fitness),individual therapy (alcohol and drug abuse) C: Publicity and health risk education	Focus: PA & Diet I: a, b, c C: b	*In USD (reference year not stated)* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	<u>Medical benefits</u> Health care costs, paid by the employer, determined from claim records
Randomized Controlled Trials						
Proper et al.(50)	Perspective: Employer Setting: NL 2000-2002 Length intervention: 9 months Follow-up: 18 months	299 civil servants I: 131 C: 168	I: <u>Comprehensive health promotion intervention</u>: 9 month counselling programme (7 face-to-face sessions), health education materials C: Health education materials	Focus: PA & Diet I: b, c C: b	*In Euros (reference year not stated)* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	<u>Absenteeism benefits</u> Number of days missed because of a health condition based on disability absence data, multiplied by an average wage rate

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
(Gussenhoven et al - unpublished data)	Perspective: Employer Setting: NL 2004-2005 Length intervention: 6 months Follow-up: 1 year	1386 *overweight* employees from 7 variable types of companies I-Phone: 462 I-Internet: 464 C: 460	I- phone: <u>Comprehensive health promotion intervention</u>: Ring binder containing 10 health education modules (lifestyle information and techniques for changing behaviour), telephone counselling, pedometer, health education materials + oral instructions I-internet: <u>Comprehensive health promotion intervention</u>: Website containing 10 health education modules (lifestyle information and techniques for changing behaviour), email counselling, pedometer, health education materials + oral instructions C: Health education materials + oral instructions	Focus: PA & Diet I-Phone: b, c I-Internet: b, c C: b	*In 2004 Euros* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	 <u>Absenteeism benefits</u> Number of days missed because of a health condition based on disability absence data, multiplied by an average wage rate <u>Medical benefits</u> [a] Self-reported health care utilization, valued using Dutch standard costs
Groeneveld et al.(48)	Perspective: Employer Setting: NL 2007-2009 Length intervention: 6 months Follow-up: 1 year	573 construction workers with an *elevated CVD risk* I: 293 C: 280	I: <u>Comprehensive health promotion intervention:</u> 6 month counselling programme (3 face-to-face sessions / 4 telephone contacts), periodical health screening, health education materials C: Periodical health screening, health education materials	Focus: PA & Diet I: a, b, c C: a, b	*In 2008 Euros* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	 <u>Absenteeism benefits</u> Self-reported days missed because of a health condition, multiplied by an average wage rate <u>Medical benefits</u> [b] Self-reported health care utilization, valued using Dutch standard costs

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
Meenan et al.(49)	Perspective: Employer Setting: U.S. 2005-2006 (baseline), 2006-2008 (follow-up) Length intervention: 2 years Follow-up: 2 years	6958 hotel employees I: 3346 C: 3612	I: <u>Comprehensive health promotion intervention:</u> HRA + feedback/advice, environmental interventions (electronic sign messages, healthier food choices, management support, other negotiated environmental changes), monthly newsletter, contests, weekly on-site health promotion groups and more intense weekly of-site health promotion groups for obese employees C: HRA + feedback/advice	Focus: PA & Diet I: a, b, c, e, f C: a, b	*In 2008 USD* <u>Intervention costs</u> Micro-costed, valued using tariffs and depleted sources	<u>Absenteeism benefits</u> Self-reported days missed because of a health condition, measured by the WHO-HPQ, multiplied by an average wage rate <u>Medical benefits</u> Medical costs were calculated as a function of: 1) The proportion of the participants with a chronic condition, 2) monthly rate of employee turnover, 3) proxies of current medical costs, and 4) proxies of the annual dollar value of a unit change in BMI. <u>Presenteeism benefits</u> Self-reported on-the-job productivity, measured by the WHO-HPQ, multiplied by an average wage rate

Study	Study Details	Population	Intervention and control conditions	Programme focus & format(s)	Included costs and benefits	
			Modelling studies			
Baker et al.(51)	Perspective: Not stated Setting: U.S. 2006-2007 Length intervention: 1 year Follow-up: 1 year	850 employees from 119 variable types of companies I: 850 C: Artificial controls	I: <u>Comprehensive health promotion intervention</u>: HRA, up to 48 telephonic counselling sessions + supportive materials, personal health improvement plan, exercise planning support, nutrition education, web-based health tracker, health improvement website including educational materials C: None	Focus: PA & Diet I: a, b, c, d C: -	*In 2007 USD* <u>Intervention costs</u> Valuation method not stated	<u>Medical benefits</u> Medical expenditures for employees with a certain health risk, based on the Thomas Reuters MarketScan database <u>Presenteeism benefits</u> On-the-job productivity losses linked to having a certain health risk as found in the literature

U.S.: United States
U.K.: United Kingdom
NL: The Netherlands
C: Control group
I: Intervention group
HRA: Health Risk Assessment
CVD: Cardiovascular Disease
PA: Physical Activity
Programme format(s): a: (self-)assessment, b: Education/Information, c: Behavioural, d: Exercise programme, e: Environment, f: Incentives
USD: U.S. Dollar
WHO-HPQ: Health and Work Performance Questionnaire of the World Health Organization
[a] Personal communication with the authors
[b] Presented in the article as part of a cost-effectiveness analysis from the societal perspective
BMI: Body Mass Index

Risk of bias assessment

Reviewers disagreed on 58 of the 344 items (17%). Disagreements were mainly due to misreading and different interpretations of the CHEC-list items. Nine out of 19 CHEC list items (47%) were fulfilled by more than 50% of the studies and seven items (37%) by more than 75%, indicating that the risk of bias of the included studies was high. RCTs, however, had a lower risk of bias compared to NRSs. On average, they fulfilled almost 13 out of 19 CHEC-list items (68%), whereas NRSs fulfilled almost 9 (47%) (Table 3). In five studies ((41,48–50); Gussenhoven et al., unpublished data) costs were measured appropriately in physical units, and of these two, (41,49) valued them appropriately by calculating them based on depleted sources and stating their reference year. One study (49) appropriately collected benefits to the chosen perspective (employer's perspective). At a minimum, these comprise medical, absenteeism and presenteeism benefits in countries with employer-provided health insurance (e.g. US). In countries with nationalized health insurance or health service programmes (e.g. the Netherlands and the UK), the last two apply (54). Seven studies ((39,41,42,48,49,51); Gussenhoven et al., unpublished data) appropriately discounted costs and benefits by converting them to a single year based on a motivated discount rate. Sensitivity analyses were performed in six studies ((41,42,44,48,49); Gussenhoven et al., unpublished data).

Costs and benefits

Average annual programme costs per participant ranged from \$11 to \$1,075 (median: \$155, n = 21). Average annual absenteeism and medical benefits per participant ranged from -\$113 to \$1,384 (median: \$324, n = 15) and -\$82 to \$554 (median: \$187, n = 13), respectively. One study (46) included absenteeism and medical benefits in the total benefits and could therefore not be presented separately. Average annual presenteeism benefits per participant ranged from \$2 to \$1,528 (median: \$158, n = 3) (Table 4, columns 2–5).

Table 3: Risk of bias assessment of included studies using the CHEC-list and BMJ-checklist.

Items	Studies scoring "Yes" [No. (%)]		
	RCTs (n = 4)	NRSs (n = 14)	Overall (n = 18)
CHEC-list			
1) Study population	3	3	6 (33)
2) Competing alternatives	4	2	6 (33)
3) Research question	1	9	10 (56)
4) Study design	4	13	17 (94)
5) Time horizon	4	14	18 (100)
6) Perspective	4	6	10 (56)
7) Costs identified	4	12	16 (89)
8) Costs measured	4	1	5 (28)
9) Costs valued	1	1	2 (11)
10) outcomes identified	1	0	1 (6)
11) Outcomes measured	3	13	16 (89)
12) Outcomes valued	3	12	15 (83)
13) Incremental analysis	3	12	15 (83)
14) Discounted	3	4	7 (39)
15) Sensitivity analysis	3	3	6 (33)
16) Conclusions	4	13	17 (94)
17) Generalizability	1	2	3 (17)
18) Conflict of interest	1	2	3 (17)
19) Ethical and distributional issues	0	0	0 (0)
BMJ-checklist			
20) Model details	N.A.	1	1 (100)
21) Model and key parameters	N.A.	1	1 (100)

RCT: Randomized Controlled Trial
NRS: Non-Randomized Study
N.A.: Not Applicable

Financial return

The NB ranged from -\$451 to \$2,757 (median; \$91, n = 21), indicating the amount of money gained after costs were recovered. The BCR ranged from -0.76 to 18.84 (median: 1.42, mean: 3.76, SD: 5.36), indicating the amount of money returned per dollar invested. The ROI ranged from -176% to 1,784% (median: 42%, mean: 276%, SD: 536%), indicating the percentage of profit per dollar invested (30). The financial return was positive in 14 out of 21 interventions (NB > 0, BCR > 1 and ROI > 0) (Table 4, column 7).

Table 4: Costs, benefits and financial return of the included studies (2010 U.S. Dollars)

Study	Costs ($)	Benefits ($)				Financial Return NB($) [ROI(%);BCR]				
		Absenteeism	Presenteeism	Medical	Total	Total	Absenteeism	Medical	Presenteeism	Absenteeism & Medical
Non Randomized Studies										
Wood et al.(47)	123	348	-	-	348	225 [184; 2.84]	225 [184; 2.84]	-	-	-
Shore et al.(45)	664	937	-	-	937	273 [41; 1.41]	273 [41; 1.41]	-	-	-
Schultz et al.(43)	103	305	-	-	305	201 [195; 2.95]	201 [195; 2.95]	-	-	-
Aldana et al.*(53)	66	1033	-	-	1033	967 [1460; 15.6]	967 [1460; 15.6]	-	-	-
Bertera et al.(37)	73	104	-	-	104	31 [42; 1.42]	31[42; 1.42]	-	-	-
Mills et al.(40)	155	1384	1528	-	2912	2757 [1784; 18.84]	1129 [795; 8.95]	-	1373 [888; 9.88]	-
Getmann*(38)	1075	504	-	477	980	-94 [-9; 0.91]	-571 [-53; 0.47]	-598 [-56; 0.44]	-	-94 [-9; 0.91]
Shi(44)			-						-	
I-Low	11	-2		13	11	-0.1 [-1; 0.99]	-13 [-117; -0.17]	2 [16; 1.16]		-0.1 [-1; 0.99]
I-Medium	21	152		99	251	230 [1129; 12.29]	131 [646; 7.46]	78 [383; 4.83]		230 [1129; 2.29]
I-High	213	324		187	511	298 [140; 2.40]	111 [52; 1.52]	-26 [-12; 0.88]		298 [140; 2.40]
Stave et al.(46)	127	See Total	-	See Total	987	860 [677; 7.77]	-	-	-	860 [677; 7.77]

Study	Costs ($)	Benefits ($)				Financial Return NB($) [ROI(%);BCR]				
		Absenteeism	Presenteeism	Medical	Total	Total	Absenteeism	Medical	Presenteeism	Absenteeism & Medical
Ozminkowski et al.(42)	119	-	-	554	554	435 [364; 4.64]	-	435 [364; 4.64]	-	-
Naydeck et al.(41)	155	-	-	198	198	42 [27; 1.27]	-	42 [27; 1.27]	-	-
Aldana et al.2(35)	107	-	-	197	197	91 [85; 1.85]	-	91 [85; 1.85]	-	-
Gibbs et al.*(39)	469	-	-	248	248	-221 [-47; 0.53]	-	-221 [-47; 0.53]	-	-
MEAN(SD)							*ROI: 325 (497)* *BCR: 4.25 (4.97)*	*ROI: 95 (178)* *BCR: 1.95 (1.78)*	-	*ROI: 387 (501)* *BCR: 4.87 (5.01)*
Randomized Controlled Trials										
Proper et al.(50)	385	680	-	-	680	295 [77; 1.77]	295 [77; 1.77]	-	-	-
Gussenhoven et al, unpublished data)										
I-Phone	256	-113	-	-82	-195	-451 [-176; -0.76]	-369 [-144; -0.44]	-338 [-132; -0.32]	-	-451[-176; -0.76]
I-Internet	227	166		-66	100	-127 [-56; 0.44]	-61 [-27; 0.73]	-293 [-129; -0.29]		-127[-56; 0.44]
Groeneveld et al.(48)	730	364	-	81	445	-285 [-39; 0.61]	-365 [-50; 0.50]	-649 [-89; 0.11]	-	-285 [-39; 0.61]
Meenan et al.(49)	139	-0.2	2	5	6	-132 [-95; 0.05]	-139 [-100; -0.001]	-134 [-97; 0.03]	-136 [-98; 0.02]	-140 [-97; 0.03]

Study	Costs ($)	Benefits ($)				Financial Return NB($) [ROI(%);BCR]				
		Absenteeism	Presenteeism	Medical	Total	Total	Absenteeism	Medical	Presenteeism	Absenteeism & Medical
MEAN(SD)							*ROI: -49 (84)* *BCR: 0.51 (0.84)*	*ROI: -112 (22)* *BCR: -0.12 (0.22)*		*ROI: -92 (61)* *BCR: 0.08 (0.61)*
Modelling studies										
Baker et al.(51)	331		158	229	387	55 [17; 1.17]	-	-102 [-31; 0.69]	-174 [-52; 0.48]	-
All studies										
MEAN(SD)						*ROI: 276 (536)* *BCR: 3.76 (5.36)*	*ROI: 200 (440)* *BCR: 3.00 (4.40)*	*ROI: 22 (168)* *BCR: 1.22 (1.68)*	*ROI: 246 (557)* *BCR: 3.46 (5.57)*	*ROI: 174 (438)* *BCR: 2.74 (4.38)*

Empty cells indicate that a specific type of benefit was not presented by an article
*Benefits are not corrected for baseline, but calculated as the difference in costs between the intervention and control group during follow-up
NB: Net Benefits
ROI: Return On Investment
BCR: Benefit Cost Ratio
NRS: Non-Randomized Study
RCT: Randomized Controlled Trial

Benefit-standardized financial return

On average, benefit-standardized ROIs and BCRs were positive, indicating that WHP programmes aimed at improving nutrition and/or increasing physical activity generate financial savings during the first years after implementation. For example, the average ROI in terms of absenteeism benefits was 200% (SD: 440%), in terms of medical benefits 22% (SD: 168%), in terms of presenteeism benefits 246% (SD: 557%), and in terms of both absenteeism and medical benefits 174% (SD: 438%) (Table 4, columns: 8–11).

Subgroup analysis

Average benefit-standardized ROIs and BCRs were positive in NRSs, but negative in RCTs (Table 4, columns: 8–11). For example, the average ROI in terms of absenteeism benefits was 325% (SD: 497%) in NRSs, but -49% (SD: 84%) in RCTs. This indicates that WHP programmes aimed at improving nutrition and/or increasing physical activity generate financial savings during the first years after implementation according to NRSs, whereas they do not pay for themselves in terms of absenteeism benefits, medical benefits or both according to RCTs. The average ROI and BCR in terms of presenteeism benefits could not be compared between study designs, since presenteeism benefits were only provided by three studies. The differences in ROI between NRSs and RCTs are depicted graphically in Figure 2.

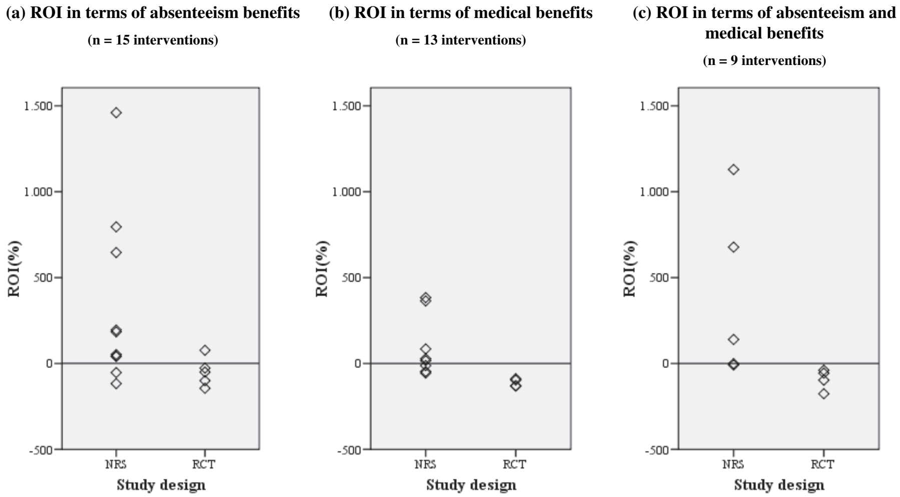

Figure 2: Distribution of Return On Investments (ROIs) in terms of (a) absenteeism, (b) medical, and (c) both absenteeism and medical benefits of Non-Randomized Studies (NRSs) and Randomized Controlled Trials (RCTs)
ROI: Return On Investment
NRS: Non-Randomized Study
RCT: Randomized Controlled Trial
A ROI of more than 0 indicates that the financial profitability is positive
Note that the number of interventions is higher than the number of studies, because some studies included more than one intervention.

DISCUSSION

This review critically appraised and summarized the current evidence on the financial return of WHP programmes aimed at improving nutrition and/or increasing physical activity. On average, financial returns in terms of absenteeism benefits, medical benefits or both were positive during the first years after implementation. This is in accordance with previous reviews (9,16,17,53) concluding that WHP programmes should be considered as an effective method for reducing employee-related expenses (16,17,53) and producing positive financial returns in terms of absenteeism and medical benefits (9). A subgroup analysis, however, revealed that the average financial return estimates were positive due to the inclusion of NRSs; they were positive in NRSs, but negative in RCTs. This is in line with previous findings indicating that NRSs of healthcare interventions tend to result in larger estimates of effect compared to

RCTs (55). These findings also support researchers arguing that the cost savings and high ROI estimates found in WHP studies are likely the result of selection bias (11). Selection bias arises when allocation methods other than randomization are used, meaning that the intervention and control group are unlikely to be comparable (56). Consequently, it is difficult to attribute any differences found in outcomes between both groups to the intervention and to rule out the possibility that they were biased by baseline differences in group characteristics or confounders (e.g. motivation to change health) (57). It has been argued that results of RCTs may not reflect ‘real-life’ effectiveness, since they evaluate the efficacy of programmes in well-controlled experimental circumstances. However, although other research designs can add to the existing knowledge on WHP programmes, RCTs are the ‘gold standard’ for investigating their effectiveness untainted by bias (58,59).
The overall risk of bias of the included studies was high. Few studies explicitly stated the perspective of their ROI analysis and properly measured and valued costs and benefits. More than half of the studies did not state the reference year of their monetary outcomes, which limits their interpretation. In addition, an incremental analysis of costs and benefits was not performed in all studies. One study (35), for example, included the decrease in medical costs of both the intervention and control group in their benefit estimate, resulting in an overestimation of the financial return. Furthermore, although economic analyses require that assumptions are made (28), few studies conducted a sensitivity analysis and hardly any of the studies reported on the uncertainty around their financial return estimates. To quantify the precision, non-parametric bootstrapping can be used as a statistical technique for dealing with the highly skewed nature of cost data (28,52). These findings are not unique to the present review. A systematic review appraising the methodological quality of economic evaluations of occupational health and safety interventions also concluded that most of them had a high risk of bias (28). Using the results of ROI analyses with a high risk of bias to advise companies, however, may lead to inappropriate business decisions (28). Therefore, the methodological quality of ROI analyses in WHP programme research should be improved. This can be achieved by developing a methodological guideline for ROI analyses. Furthermore, since NRSs had a higher risk of bias compared to RCTs, the discrepancies found between their financial

return estimates may also be explained by types of bias other than selection (e.g. performance, detection, attrition and reporting bias) (56).

The results of the present review indicate that financial return estimates derived from NRSs should be interpreted with caution. RCTs with a low risk of bias indicate that WHP programmes aimed at improving nutrition and/or increasing physical activity do not pay for themselves in terms of reduced absenteeism costs, medical costs or both during the first years after implementation. This is in contrast with the conclusions of previous reviews (9,16,17,53). An explanation for this discrepancy may be that the previous reviews were mainly based on NRSs, which might have confounded their results as well.

Several strengths of the present review are noteworthy. First, to improve comparability among the included studies, costs and benefits were standardized to annual costs per participant in 2010 dollars and ROI metrics were (re-)calculated per study using the same methodology. Second, when reporting the financial return of WHP programmes, economists and policy makers prefer the NB, whereas the BCR and ROI are more familiar to business managers (60). By providing all three of them, the results of the present review can be easily interpreted by all stakeholders. In addition, this makes the results easily comparable with those of other studies, since different ROI metrics are used in the literature to estimate the financial return of WHP programmes. Third, the present study was the first review on the financial return of WHP programmes in which subgroup analyses were performed to compare financial return estimates of RCTs and NRSs, yielding substantial differences.

A first limitation concerns the fact that none of the interventions were solely aimed at improving nutrition and only two of them were solely aimed at increasing physical activity. Therefore, the present review examined the financial return of WHP programmes aimed at improving nutrition and/or increasing physical activity in general. Further research is needed to investigate whether financial returns vary between interventions with a different focus (i.e. improving nutrition, increasing physical activity or both). Additionally, only the financial return in terms of absenteeism and/or medical benefits were compared between RCTs and NRSs. WHP programmes, however, are suggested to provide additional types of financial benefits, such as reduced presenteeism, turnover, disability management and workers' compensation costs (16,54). Presenteeism benefits were only presented in three studies, which likely resulted from the fact that a 'gold standard' for measuring and valuing presenteeism does currently not exists. The other three types of financial benefits were not presented at all (61). Consequently, conclusions about the overall profitability of WHP programmes aimed at improving nutrition and/or increasing physical activity cannot be made. Furthermore, WHP programmes may yield intangible benefits (e.g. improved reputation or increased worker satisfaction) (34), which were not reported by any of the studies. Since intangible benefits may also be important drivers of business decisions (34), it is advisable to report them alongside ROI analyses or to conduct a cost-effectiveness analysis in which the total incremental costs are compared to the incremental intangible benefits. Furthermore, the varying number and type of benefits included in the studies indicate that consensus should be reached about a minimum set of benefits to be included in ROI analyses of WHP programmes. Another limitation may be that no requirements were set as to programme format, subject and worksite characteristics, intervention length and follow-up duration. Consequently, NRSs and RCTs may differ with respect to these characteristics contributing to the discrepancies found in financial return estimates between both study designs. For example, the follow-up duration of NRSs was, on average, longer than that of RCTs. Since WHP programme costs are more costly at the start while health benefits accumulate gradually (9), this may have resulted in lower financial return estimates in the RCTs. Therefore, conclusions about the extent to which financial return estimates were overestimated in NRSs cannot be made. It is also important to mention that US employers bear a large part of the medical costs of their employees, whereas in Europe these accrue to the government or insurance companies. As a result, ROI analyses from the employer's perspective conducted in the USA and Europe are limited in their comparability. To provide information that would be useful to both sides of the Atlantic, benefit-standardized financial return estimates were calculated, including financial returns in terms of absenteeism benefits, medical benefits and both. Benefit-standardized financial returns in terms of medical benefits assume that no benefits accrue in terms of reduced absenteeism costs and vice versa for financial returns in terms of medical benefits. Thus, US employers are informed by the total benefits, whereas

European employers are informed by the productivity-related benefits and European governments and insurance companies by the medical benefits. An advantage of this approach is that RCTs and NRSs could be compared, without distortion resulting from differences in the jurisdictions in which they were conducted. It should also be noted that no corrections were made for transatlantic differences in healthcare costs. Per capita spending on health care in the USA is double that of most European countries, leaving more room for reductions in medical costs in the USA than in Europe (62). This may have influenced the differences found between RCTs and NRSs as all but one of the NRSs were performed in the USA, whereas all but one of the RCTs were performed in Europe. Nevertheless, in accordance with the overall results, financial returns were negative in the RCT conducted in the USA, whereas those of the NRSs conducted in the USA were on average positive.

Conclusion

During the first years after implementation, WHP programmes aimed at improving nutrition and/or increasing physical activity generate financial savings in terms of reduced absenteeism costs, medical costs or both according to NRSs, whereas they do not according to RCTs. However, since these programmes are associated with additional types of benefits, conclusions about their overall profitability cannot be made. Therefore, more ROI analyses should be performed that are based on RCTs and include a consensus-based set of financial benefits.

REFERENCES

1. Yach D, Stuckler D, Brownell KD. Epidemiologic and economic consequences of the global epidemics of obesity and diabetes. *Nat Med* 2006; 12: 62–66.
2. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA* 2002; 288: 1723–1727.
3. Donaldson L. On the state of the public health: annual report of the chief medical officer for England 2002. Department of Health, 2003.
4. Finkelstein EA, Trogdon JG, Cohen JW, Dietz W. Annual medical spending attributable to obesity: payer-and service-specific estimates. *Health Aff (Millwood)* 2009; 28: w822–w831.
5. Thompson D, Edelsberg J, Colditz GA, Bird AP, Oster G. Lifetime health and economic consequences of obesity. *Arch Intern Med* 1999; 159: 2177–2183.
6. Trogdon JG, Finkelstein EA, Hylands T, Dellea PS, Kamal-Bahl SJ. Indirect costs of obesity: a review of the current literature. *Obes Rev* 2008; 9: 489–500.
7. van Duijvenbode DC, Hoozemans MJM, van Poppel MNM, Proper KI. The relationship between overweight and obesity, and sick leave: a systematic review. *Int J Obes* 2009; 33: 807–816.
8. Goetzel RZ, Juday TR, Ozminkowski RJ. What's the ROI? A systematic review on return of investment (ROI) studies of corporate health and productivity management initiatives. *AWPH's Worksite Health* 1999; 6: 12–21.
9. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Aff (Millwood)* 2010; 29: 304–311.
10. National Institute for Health and Clinical Excellence (NICE). *Workplace Health Promotion: How to Encourage Employees to Be Physically Active*. NICE: London, 2007.
11. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Annu Rev Public Health* 2008; 29: 303–323.
12. Anderson LM, Quinn TA, Glanz K, Ramirez G, Kahwati LC, Johnson DB *et al*. The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: a systematic review. *Am J Prev Med* 2009; 37: 340–357.
13. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Lifestyle-focused interventions at the workplace to reduce the risk of cardiovascular disease – a systematic review. *Scand J Work Environ Health* 2010; 36: 202–215.
14. Verweij LM, Coffeng J, Van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. *Obes Rev* 2011; 12: 406–429. doi: 10.1111/j.1467-789X.2010.00765.x.
15. Cavallo D. Using return on investment analysis to evaluate health promotion programs: challenges and opportunities. *Health Promot Econ Issue Briefs* 2006; 1: 1–4. RTI-UNC Center of Excellence. [WWW document]. URL http://www.rti.org/pubs/issuebrief_3.pdf (accessed 12 October 2010).
16. Chapman LS. Meta-evaluation of worksite health promotion economic return studies: 2005 update. *Am J Health Promot* 2005; 19: 1–11.
17. Soler RE, Leeks KD, Razi S, Hopkins DP, Griffith M, Aten A *et al*. A systematic review of selected interventions for worksite health promotion: the assessment of health risks with feedback. *Am J Prev Med* 2010; 38(Suppl. 1): S237–S262.

18. Sexner S, Gold DB. Calculating the economic return of health and productivity programs. In: Pronk NP (ed.). *ACSM's Workplace Health Handbook*. Human Kinetics: Minneapolis, MN, 2009, pp. 165–174.
19. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1991–1993 update. *Am J Health Promot* 1993; 8: 50–62.
20. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1993–1995 update. *Am J Health Promot* 1996; 10: 380–388.
21. Pelletier KR. Clinical and cost outcomes of multifactorial, cardiovascular risk management interventions in worksites: a comprehensive review and analysis. *J Occup Environ Med* 1997; 39: 1154–1169.
22. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1995–1998 update (IV). *Am J Health Promot* 1999; 13: 333–345, iii.
23. Pelletier KR. A review and analysis of the clinical- and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1998–2000 update. *Am J Health Promot* 2001; 16: 107–116.
24. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VI 2000–2004. *J Occup Environ Med* 2005; 47: 1051–1058.
25. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VII 2004–2008. *J Occup Environ Med* 2009; 51: 822–837.
26. Proper KI, Van Mechelen W. *Effectiveness and Economic Impact of Worksite Interventions to Promote Physical Activity and Healthy Diet*. WHO Press: Geneva, 2008.
27. Evers S, Goossens M, de Vet H, van Tulder M, Ament A. Criteria list for assessment of methodological quality of economic evaluations: consensus on health economic criteria. *Int J Technol Assess Health Care* 2005; 21: 240–245.
28. Uegaki K, de Bruijne MC, Lambeek L, Anema JR, van der Beek AJ, van Mechelen W *et al*. Economic evaluations of occupational health interventions from a corporate perspective – a systematic review of methodological quality. *Scand J Work Environ Health* 2010; 36: 273–288.
29. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996; 313: 275–283.
30. Phillips JJ. *Return on Investment in Training and Performance Improvement Programs*, 2nd edn. Elsevier: Burlington, VT, 2003.
31. Stone PW. Return-on-investment models. *Appl Nurs Res* 2005; 18: 186–189.
32. U.S. Bureau of Labor Statistics. [WWW document]. URL http://www.bls.gov (accessed 1 December 2010).
33. Organisation for Economic Co-operation and Development (OECD). StatExtracts. OECD Selected Data. [WWW document]. URL http://stats.oecd.org/Index.aspx (accessed 1 December 2010).
34. Verbeek J, Pulliainen M, Kankaanpaa E. A systematic review of occupational safety and health business cases. *Scand J Work Environ Health* 2009; 35: 403–412.
35. Aldana SG, Jacobson BH, Harris CJ, Kelley PL, Stone WJ. Influence of a mobile worksite health promotion program on health care costs. *Am J Prev Med* 1993; 9: 378–383.

36. Aldana SG, Merrill RM, Price K, Hardy A, Hager R. Financial impact of a comprehensive multisite workplace health promotion program. *Prev Med* 2005; 40: 131–137.

37. Bertera RL. The effects of workplace health promotion on absenteeism and employment costs in a large industrial population. *Am J Public Health* 1990; 80: 1101–1105.

38. Gettman LR. Cost/benefit analysis of a corporate fitness program. *Fitness Bus* 1986; 1: 11–17.

39. Gibbs JO, Mulvaney D, Henes C, Reed RW. Work-site health promotion: five-year trend in employee health care costs. *J Occup Environ Med* 1985; 27: 826–830.

40. Mills PR, Kessler RC, Cooper J, Sullivan S. Impact of a health promotion program on employee health risks and work productivity. *Am J Health Promot* 2007; 22: 45–53.

41. Naydeck BL, Pearson JA, Ozminkowski RJ, Day BT, Goetzel RZ. The impact of the Highmark employee wellness programs on 4-year healthcare costs. *J Occup Environ Med* 2008; 50: 146–156.

42. Ozminkowski RJ, Dunn RL, Goetzel RZ, Cantor RI, Murnane J, Harrison M. A return on investment evaluation of the Citibank, N.A., health management program. *Am J Health Promot* 1999; 14: 31–43.

43. Schultz AB, Lu C, Barnett TE, Yen LT, McDonald T, Hirschland D *et al*. Influence of participation in a worksite health promotion program on disability days. *J Occup Environ Med* 2002; 44: 776–780.

44. Shi L. Health promotion, medical care use, and costs in a sample of worksite employees. *Eval Rev* 1993; 17: 475–487.

45. Shore G, Prasad P, Zroback M. Metrofit: a cost-effective fitness program. *Fitness Bus* 1989; 4: 147–153.

46. Stave GM, Muchmore L, Gardner H. Quantifiable impact of the contract for health and wellness: health behaviors, health care costs, disability, and workers' compensation. *J Occup Environ Med* 2003; 45: 109–117.

47. Wood EA, Olmstead GW, Craig JL. An evaluation of lifestyle risk factors and absenteeism after two years in a worksite health promotion program. *Am J Health Promot* 1989; 4: 128–133.

48. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, Van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med* 2011; 56: 610–617.

49. Meenan RT, Vogt TM, Williams AE, Stevens VJ, Albright CL, Nigg C. Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii. *J Occup Environ Med* 2010; 52(Suppl. 1): S8–S13.

50. Proper KI, de Bruyne MC, Hildebrandt VH, van der Beek AJ, Meerding WJ, van Mechelen W. Costs, benefits and effectiveness of worksite physical activity counseling from the employer's perspective. *Scand J Work Environ Health* 2004; 30: 36–46.

51. Baker KM, Goetzel RZ, Pei X, Weiss AJ, Bowen J, Tabrizi MJ *et al*. Using a return-on-investment estimation model to evaluate outcomes from an obesity management worksite health promotion program. *J Occup Environ Med* 2008; 50: 981–990.

52. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*, 3rd edn. Oxford University Press: New York, 2005.

53. Aldana SG. Financial impact of health promotion programs: a comprehensive review of the literature. *Am J Health Promot* 2001; 15: 296–320.
54. Nicholson S, Pauly MV, Polsky D, Baase CM, Billotti GM, Ozminkowski RJ *et al*. How to present the business case for healthcare quality to employers. *Appl Health Econ Health Policy* 2005; 4: 209–218.
55. Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev* 2007; Issue 2. Art. No. MR000012.
56. Higgins JPT, Altman DG. Chapter 8. Assessing risk of bias in included studies. In: Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.0.1 (updated September 2008). The Cochrane Collaboration; 2008 [WWW document]. URL http://www.igh.org/Cochrane/tools/Ch08_Bias.pdf (accessed 15 December 2010).
57. Linden A. Identifying spin in health management evaluations. *J Eval Clin Pract* 2011. doi: 10.1111/j.1365-2753.2010.01611.x/pdf (accessed 10 February 2011).
58. Driessen MT, Anema JR, Proper KI, van der Beek AJ. Authors' response: RCTs of ergonomic interventions. *Occup Environ Med* 2010; 67: 218–219.
59. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I *et al*. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996; 276: 637–639.
60. Ozminkowski RJ, Goetzel RZ. Getting closer to the truth: overcoming research challenges when estimating the financial impact of worksite health promotion programs. *Am J Health Promot* 2001; 15: 289–295.
61. Lofland JH, Pizzi L, Frick KD. A review of health-related workplace productivity loss instruments. *PharmacoEconomics* 2004; 22: 165–184.
62. Reinhardt UE, Hussey PS, Anderson GF. U.S. health care spending in an international context. *Health Aff (Millwood)* 2004; 23: 10–25.

4

A cost-effectiveness and return-on-investment analysis of a worksite vitality intervention among older hospital workers: results of a randomized controlled trial

Johanna M van Dongen
Jorien E Strijk
Karin I Proper
Marieke F van Wier
Willem van Mechelen
Maurits W van Tulder
Allard J van der Beek

J Occup Environ Med 2013, 55(3): 337-346

ABSTRACT

Objective: To conduct a cost-effectiveness and return-on-investment analysis comparing a worksite vitality intervention with usual care.

Methods: A total of 730 older hospital workers were randomized to the intervention or control group. The 6-month intervention consisted of yoga and aerobic exercising, coaching, and fruit. At baseline, and 6 and 12 months, general vitality, work-related vitality, and need for recovery were determined. Cost data were collected on a 3-monthly basis. The cost-effectiveness analysis was performed from the societal perspective and the return-on-investment analysis from the employer's perspective using bootstrapping techniques.

Results: No significant differences in costs and effects were observed. Incremental cost-effectiveness ratios in terms of general vitality (range, 0 to 100), work-related vitality (range, 0 to 6), and need for recovery (range, 0 to 100) were, respectively, €280, €7506, and €258 per point improvement. Per euro invested, €2.21 was lost.

Conclusions: The intervention was neither cost-effective nor cost-saving.

INTRODUCTION

In various European countries, people aged 60 years and older will comprise up to one third of the population during the next decades. Because a shrinking labor force will have to support a growing number of retired people (1), there is a need for workers who are able to prolong their working life in good health (2). In the Vital@ Work study, a worksite vitality intervention was developed that aimed to improve physical activity, nutrition, and relaxation, as a potentially effective tool to keep older workers vital (ie, at a perceived high energy level, lowlevels of fatigue, and feeling fit) and healthy, thereby contributing to prolonged employability (2).

An evaluation of the Vital@Work intervention's effectiveness has been reported elsewhere (3,4). Nevertheless, budgets for occupational health care are restricted. Therefore, decisions about investments in worksite interventions may be guided not only by the evidence on their effectiveness, but also by considerations of their costs in relation to these effects (5). In occupational health care research, cost-effectiveness analyses (CEAs) are conducted to gain insight into the (additional) costs of an intervention per additional unit of effect gained. These results can be used by decision makers to decide how resources should optimally be allocated to maximize health or welfare (6,7). Within business administration, the primary interest may not be in maximizing health or welfare but in maximizing the financial return of an intervention (8). This is often determined using a return-on-investment (ROI) analysis, in which intervention costs are compared with their resulting financial benefits (ie, program outcomes converted to monetary values) (9–11). As CEAs and ROI analysis are based on the same data, both can be conducted simultaneously and doing so provides information that can be used by business managers and experts in occupational health care research.

The aim of the present study was to conduct a CEA and ROI analysis in which the Vital@Work intervention was compared with usual care. The CEA was performed from the societal perspective, which is generally advocated for when various stakeholders may be affected by an intervention (7,12). This is clearly the case for worksite health promotion interventions, as employers invest in the program and may benefit from it through reduced productivity-related spending, whereas (in the

Dutch situation) the government and health insurance companies may benefit from it through reduced medical costs. Because employers are the ones deciding whether or not to implement such intervention, and in doing so may have an explicit interest in its financial return, the ROI analysis was performed from the employer's perspective.

METHODS

Study population and design

The present study was conducted alongside a randomized controlled trial (RCT) (2). The follow-up was 12 months and data collection took place during 2009 and 2010. Older workers (45 years or older) from two Dutch academic hospitals were invited to participate: VU University Medical Center Amsterdam (VUMC, Amsterdam, the Netherlands) and Leiden University Medical Center (LUMC, Leiden, the Netherlands). The criteria for inclusion were: (1) working at least 16 hours a week, and (2) no risk for developing adverse health effects when becoming physically active. At enrollment, workers provided written informed consent. After baseline measurements, they were individually randomized to the intervention or control group by a research assistant using Random Allocation Software (version 1.0, May 2004, Isfahan University of Medical Sciences, Iran). The research assistant had no information about the workers to ensure concealment of treatment allocation. The study protocol was approved by the medical ethics committee of the VUMC Amsterdam (2). The sample size was based on detecting a 10% difference in work-related vitality, measured by the Utrecht Work Engagement Scale (UWES) (13). Assuming a mean baseline UWES Vitality Score (range, 0 to 6) of 3.99 (standard deviation [SD], 1.11) (14), a power of 0.90, and a confidence interval (CI) of 95% (α = 0.05), 189 workers were needed per group at follow-up (2). Taking into account a loss to follow-up of 15%, at least 446 workers (223 per group) needed to be included at baseline.

Control and intervention condition

After randomization, all workers received written information about a healthy lifestyle regarding physical activity, nutrition, and relaxation. Subsequently, workers in the intervention group received the Vital@Work intervention.

A full description of the Vital@Work intervention has been given elsewhere (2). Briefly, the intervention consisted of a Vitality Exercise Program (VEP), three Personal Vitality Coach (PVC) visits, and free fruit (2).

The VEP lasted 24 weeks. Once a week, workers were invited to participate in a guided group yoga session, a guided group workout session, and 45 minutes of unsupervised vigorous physical activity (eg, fitness and spinning). Guided group sessions were provided in small groups (16 participants or fewer) and lasted 45 minutes as well. During working days (Monday to Friday), group sessions were provided in two time blocks: (1) during lunchtime, and (2) directly after working hours (after 4 PM). Yoga sessions were guided by qualified yoga instructors and took place at the worksite. Workout sessions were guided by certified fitness instructors and took place at a fitness center near the worksite (2).

PVC visits took place at the worksite. The first visit was scheduled at the start of the intervention and was followed by two visits at 4 to 6 and 10 to 12 weeks. Before the start of the intervention, the PVC protocol and accompanying materials (eg, coaching registration forms) were explained to the coaches during 4-hour training sessions (2). Free fruit was provided during the guided group sessions of the VEP (2).

Effect measures

Vitality and need for recovery (NFR) from work-induced efforts, which is thought to increase with age (15), were assessed at baseline and 6 and 12 months.

Vitality was measured using two questionnaires. The RAND-36 Vitality Scale was used to measure general vitality and included four items assessing a worker's general vitality during the previous 4 weeks. Items were scored on a 6-point scale ranging from "all of the time" (1) to "none of the time" (6) (16). The RAND-36 Vitality Score ranged from 0 to 100 (higher scores indicate a better general vitality). The RAND-36 Vitality Scale has shown to be sufficiently reliable; internal consistency was 0.82 (Cronbach α), and the 6-month test–retest stability coefficient was 0.63 (16). Work-related vitality was measured using a subscale of the UWES (ie, UWES Vitality Scale). This scale included six items, scored on a 7-point scale ranging from "never" (0) to "always" (6). The UWES Vitality Score ranged from 0 to 6 (higher scores indicate a better work-related vitality) (13). The UWES Vitality Scale has shown sufficient internal consistency (Cronbach α =0.83). Also, two longitudinal studies carried out in Australia and Norway showed 1-year test–retest stability coefficients ranging between 0.64 and 0.71 (13).

The NFR was assessed using a subscale of the "Dutch Questionnaire on the Experience and Evaluation of Work" (ie, NFR scale). The NFR scale contains 11 statements, answered with "Yes" or "No", and has shown sufficient internal consistency (Cronbach α =0.88) (17). Also, a 2-year test–retest intra class coefficient of 0.80 was found among Dutch hospital nurses (18). The NFR score ranged from 0 to 100 (lower scores indicate a better NFR) (17).

Resource use and valuation

Intervention costs were estimated using a bottom-up microcosting approach (ie, detailed data were collected regarding the quantity and unit prices of resources consumed). During the study period, data on other resource use (ie, health care, absenteeism, presenteeism, and sports activities) were collected on a 3-monthly basis using retrospective questionnaires. All costs were converted to 2010 Euros using consumer price indices (19). As the follow-up of the trial was 1 year, discounting of costs and effects was not necessary (7).

Intervention costs were those related to implementing and operating the Vital@Work intervention (ie, costs for VEP, PVC visits, fruit, and printed materials). The number of guided group sessions was monitored using attendance registration forms. The number of PVC visits per worker and their average duration were recorded by the coaches. Labor costs were valued using the total time investments of the intervention staff and their gross salaries including holiday allowances and premiums. Capital costs were valued using cost data collected from project and finance department staff. Costs of printed materials and the provision of fruit were estimated using invoices.

Health care utilization was assessed using 3-monthly retrospective questionnaires and included cost categories relevant to the study outcomes and intervention; primary health care (ie, general practitioner, allied health professionals, and complementary medicine) and secondary health care (ie, medical specialist and hospitalization). Dutch standard costs were used to value health care utilization (20). If these were unavailable, prices according to professional organizations were used.

Absenteeism was assessed using an item of the "Productivity and Disease Questionnaire" (PRODISQ) asking workers to report their total number of sick leave days during the past 3 months (21). The absenteeism module of the PRODISQ showed satisfactory responsiveness and construct validity (22). In accordance with the Dutch Manual of Costing, costs associated with one sick leave day were calculated per worker by dividing their gross annual salary including holiday allowances and premiums by their total number of workable days per year (20). Gross annual salaries including holiday allowances and premiums were calculated using a worker's self-reported net salary. Therefore, Dutch total tax on income rates (23) and the percentage of holiday allowances and premiums according to the Dutch Manual of Costing were used (20). Using the Friction Cost Approach (FCA), absenteeism costs were estimated by multiplying the total number of sick leave days during follow-up by their associated costs. The FCA assumes that costs are limited to the friction period (ie, period needed to replace a sick worker). A friction period of 23 weeks and an elasticity of 0.8 were used (20,24).

Presenteeism (ie, reduced productivity while at work) (25) was assessed using an item of The World Health Organization Health and Work Performance Questionnaire (WHO-HPQ) (26). Workers were asked to rate their overall work performance during the previous 4 weeks on an 11-point scale, ranging from "worst performance" (0) to "best performance" (10). The WHO-HPQ Work Performance Scale has been validated against objective measures of performance (ie, archival performance data) and good concordance was found between both measures (27). Assuming linearity, their average work performance during follow-up (Wown) was calculated. Because presenteeism is conceptualized in the WHO-HPQ as a measure of actual performance in relation to "best performance" (10) (26,28), a worker's average level of presenteeism during follow-up (presenteeism score) was calculated using the following formula:

presenteeism score = (10 - Wown)/10

Using the Human Capital Approach (HCA), presenteeism costs were calculated by multiplying a worker's presenteeism score by their gross annual salary including holiday allowances and premiums.

Costs related to the sports activities of the workers (eg, membership fees and sports equipment costs) were collected using two items with a 3-month recall period.

Potential confounders and effect modifiers

At baseline, data about potential confounders and effect modifiers were assessed by questionnaire, including age (years), sex (female/male), education level (low = elementary school or less, medium = secondary education, and high = college/ university), chronic disease status (yes/no), smoking (yes/no), intervention location (VUMC/LUMC), and marital status (having a partner: yes/no).

Statistical analyses

Analyses were performed according to the intention-to-treat principle. All missing data about general vitality, work-related vitality, NFR, and costs were imputed using Fully Conditional Specification and Predictive Mean Matching (29,30). Forty different data sets were created and pooled estimates were calculated according to Rubin's rules (31). Baseline characteristics were compared between completers and non-completers using descriptive statistics. Missing data were imputed on the cost level and not on the level of resource use. Therefore, a descriptive analysis on resource use was performed based on the complete cases using t tests for normally distributed data and Mann-Whitney U tests for non-normally distributed data. Unless otherwise stated, data were analyzed in PASW (v18.0, SPSS Inc, Chicago, IL). Statistical significance was set at $P < 0.05$.

Societal perspective: CEA

The CEA was conducted from the societal perspective (ie, all costs related to the intervention were taken into account irrespective of who pays for them). The intervention effect on both vitality measures and NFR was analyzed using linear regression. Because the addition of potential confounders did not change the intervention effects by more than 10% and no effect modifiers were found, outcome measures were only adjusted for their baseline values. Mean cost differences between the intervention and control group were calculated for total and disaggregated costs. Using R (Version 2.13.1., Free Software Foundation Inc., Boston, MA), their

95% CIs were estimated by means of approximate bootstrap confidence intervals (32). Incremental cost-effectiveness ratios (ICERs) were calculated by dividing the difference in total costs between both groups by the difference in effects. Bootstrapped incremental cost-effect pairs, using 5000 replications, were plotted on cost-effectiveness planes to graphically illustrate the uncertainty around the ICERs (33). A summary measure of the joint uncertainty of costs and effects was presented using cost-effectiveness acceptability curves. These indicate the probability of cost-effectiveness at different ceiling ratios (ie, the maximum amount of societal costs decision makers are willing to pay per unit of effect) (34).

Employer's perspective: ROI analysis

The ROI analysis was performed from the employer's perspective (ie, only the costs relevant to the employer were considered, including intervention, absenteeism, and presenteeism costs). Three ROI metrics were calculated; (1) net benefits (NBs), (2) benefit:cost ratio (BCR), and (3) ROI (10).

NB = benefits − costs
BCR = benefits/costs
*ROI = (benefits − costs)/costs [*100]*

Costs were defined as intervention costs. Benefits were defined as the difference in monetized outcome measures (ie, absenteeism, and presenteeism costs) between the intervention and control groups during follow-up, with positive benefits indicating reduced spending. To quantify precision, 95% CIs around the benefit estimates and NB were estimated by means of approximate bootstrap confidence intervals (32). Financial returns are positive if the following criteria are met: NB > 0, BCR > 1, and ROI > 0%.

Sensitivity analyses

To test the robustness of the results, four sensitivity analyses (SAs) were conducted. First, analyses were performed using the complete cases only (SA1). Second, analyses were performed in which intervention costs were based on prices paid (ie, intervention costs were solely valued using invoices) (SA2). Third, analyses were performed in which absenteeism costs were estimated using the HCA instead of the FCA (SA3). In the HCA, total sick leave days are neither "truncated" as in the FCA nor is elasticity considered (24). Fourth, because of the lack of overall consensus regarding the inclusion of presenteeism costs in economic evaluations, analyses were performed in which presenteeism costs were excluded (SA4) (10).

RESULTS

Participants

A total of 730 workers were randomized to the intervention (n = 367) or control group (n = 363). At baseline, no meaningful differences were found between both groups (Table 1). Complete follow-up data were obtained from 68.5% of the workers on the effect measures (n = 500; 250 intervention group workers and 250 control group workers) and from 53.4% of the workers on the cost measures (n = 390; 199 intervention group workers and 191 control group workers) (Figure 1). Data about VEP and PVC visits were complete for all intervention group workers. No significant differences in baseline characteristics were found between workers with complete and incomplete follow-up data.

Effectiveness

During follow-up, intervention group workers increased their general vitality by 2.5 points (range, 0 to 100) and their work-related vitality by 0.12 points (range, 0 to 6), whereas both remained about the same in the control group (general vitality, 0.0 points; work-related vitality, 0.03 points). Furthermore, the intervention group decreased their NFR by 1.8 points (range, 0 to 100), whereas that of the control group increased by 0.8 points. None of these between-group differences were statistically significant.

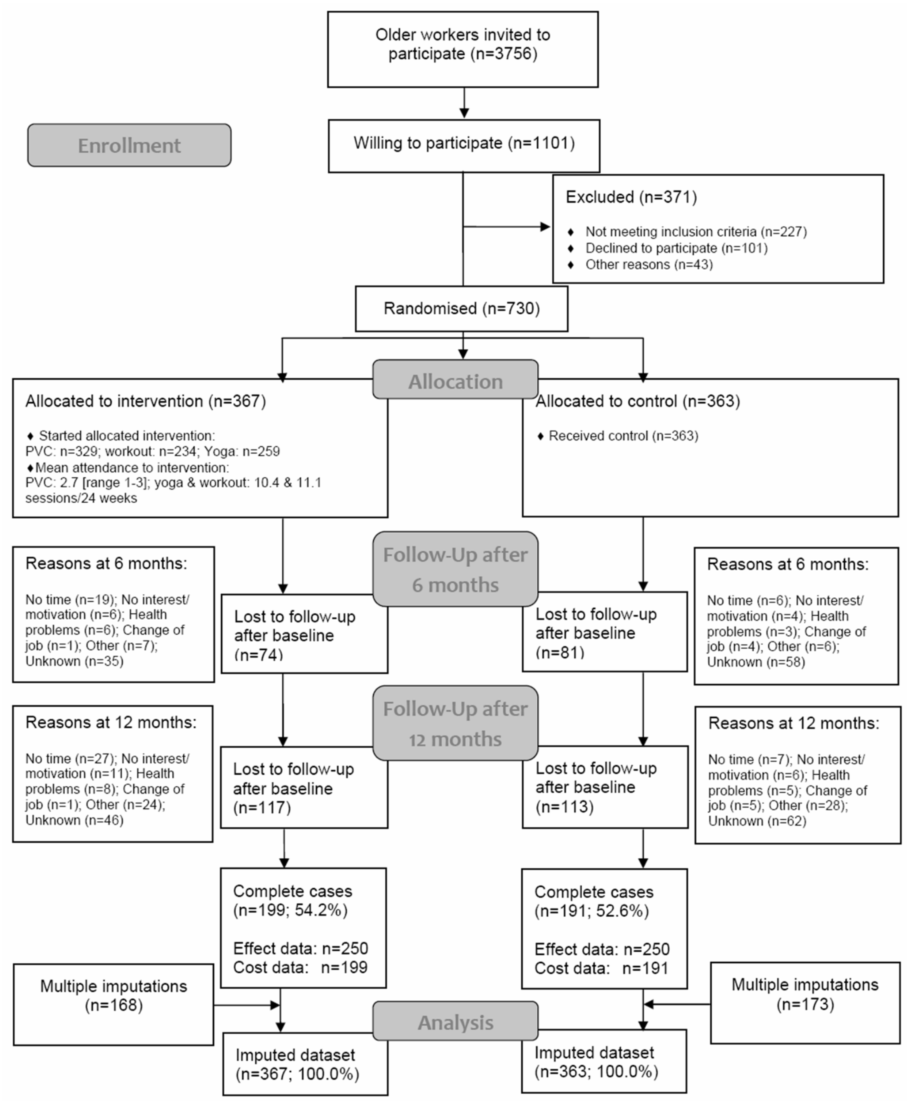

Figure 1: Flow diagram of older workers in the Vital@Work study

4

Table 1: Baseline characteristics of the study population

Baseline characteristics	Intervention group All (n=367)	Control group All (n=363)
Female [n. (%)]	274 (74.7)	277 (76.3)
Age (years) [mean (SD)]	52.5 (4.8)	52.3 (4.9)
Education level [n. (%)][1]		
Low	42 (11.4)	32 (8.8)
Intermediate	100 (27.3)	110 (30.3)
High	225 (61.3)	221 (60.9)
Working hours per week [mean (SD)]	30.4 (7.3)	29.8 (7.0)
Irregular working hours [n. (%)]		
Yes	44 (12.0)	52 (14.3)
No	323 (88.0)	311 (85.7)
General vitality (Range 0-100) [mean (SD)]	66.7 (16.9)	68.1 (16.0)
Work-related vitality (Range 0-6) [mean (SD)]	4.9 (0.9)	4.9 (0.9)
Need for recovery (Range 0-100) [mean (SD)]	29.6 (27.7)	27.8 (28.1)

Abbreviations: n: number, SD: standard deviation
[1] Education level was classified according to the definition of Statistics Netherlands (http://www.cbs.nl)

Resource use

During the intervention period, 894 PVC visits, 459 workout sessions, and 392 yoga sessions were provided. On the basis of the complete cases, workers in the intervention and control groups did not differ in terms of their median number of visits to a care provider (2.0 vs 2.0; P = 0.96), median number of days of hospitalization (0.0 vs 0.0; P = 0.74), median number of sick leave days (2.0 vs 1.0; P = 0.127), and average presenteeism scores (0.2 vs 0.2; 95% CI, −0.01 to 0.02) during follow-up.

Costs

On average, intervention costs were €149 per worker (Table 2). Medical, absenteeism, presenteeism, and total costs were higher in the intervention than in the control group during follow-up. Sports costs, however, were lowest in the intervention group. None of these between-group differences were statistically significant (Table 3).

Table 2: Cost of the Vital@Work intervention per worker

Intervention components	Cost categories	Staff	Units	Unit prices	Total costs (Euros 2010) (n=367)	Costs per worker (Euros 2010)
Vitality Exercise Program (VEP)						
Workout sessions	Labor costs	Fitness instructors	459 sessions	€ 24.35/hour	€ 10,244.14	€ 27.91
	Capital costs		Total time investment: 420,75 hours	€ 25.33/hour	€ 10,655.49	€ 29.03
Yoga sessions	Labor costs	Yoga instructors	392 sessions	€ 24.35/hour	€ 8,748.81	€ 23.84
	Capital costs		Total time investment: 359.33 hours	€ 5.87/hour	€ 2,110.72	€ 5,75
					Total	**€ 86.54**
Personal Vitality Coaching (PVC)						
Training sessions	Labor costs	Principal investigator	2 sessions	€ 26.65/hour	€ 213.19	€ 0.58
		Research assistant	Total time investment: 8 hours	€ 25.45/hour	€ 203.57	€ 0.55
		PVC coaches		€ 29.79/hour	€ 714.99	€ 1.95
	Capital costs			€ 5.87/hour	€ 46.99	€ 0.13
PVC visits	Labor costs	PVC coaches	894 visits	€ 29.79/hour	€ 13,316.67	€ 36.29
	Capital costs		Total time investment: 447 hours	€ 2.94/hour	€ 1,312.84	€ 3.58
					Total	**€ 43.08**
Provision of free fruit			288 fruit-boxes (12 per week)	€ 20.52/box	€ 5,905.23	€ 16.09
					Total	**€ 16.09**
Printed materials						
General information folder			1900 folders	€ 0.33/folder	€ 623.13	€ 1.70
Intervention logbook			367 logbooks	€ 1.53/logbook	€ 561.99	€ 1.53
PVC protocol and accompanied materials			10 ring binders	€ 14.22/binder	€ 142.20	€ 0.39
					Total	**€ 3.62**
Total intervention costs						**€ 149.33**

Abbreviations: n: number.
Note: Costs are expressed in 2010 Euros

Table 3: Mean costs per worker in the intervention and control group, and mean cost differences between both groups during the 12-month follow-up

	Imputed dataset		
Cost category	**Intervention group n=367; mean (SEM)**	**Control group n=363; mean (SEM)**	**Mean cost difference (95% CI)**
Medical costs	847 (73)	593 (53)	254 (-246 – 670)
Absenteeism costs	2793 (250)	2570 (249)	223 (-1284 – 1637)
Presenteeism costs	11580 (408)	11475 (396)	106 (-1454 – 1650)
Sports costs	553 (37)	714 (38)	-162 (-466 – 228)
Intervention costs	149 (NA)	0 (NA)	149 (NA)
Total costs	15922 (624)	15353 (574)	570 (-1968 – 2905)
	Complete dataset		
Cost category	**Intervention group n=199; mean (SD)**	**Control group n=191; mean (SD)**	**Mean cost difference (95% CI)**
Medical costs	295 (587)	277 (562)	19 (-94 - 132)
Absenteeism costs	793 (1764)	686 (1779)	107 (-259 – 446)
Presenteeism costs	9466 (4963)	9782 (6745)	-315 (-1549 – 855)
Sports costs	449 (502)	505 (608)	-56 (-170 – 45)
Intervention costs	149 (NA)	0 (NA)	149 (NA)
Total costs	11153 (5828)	11249 (7671)	-96 (-1578 – 1237)

Abbreviations: n: number; SEM: standard error of the mean, CI: confidence interval, NA: not applicable, SD: standard deviation
Note: Costs are expressed in 2010 Euros

Societal perspective: Cost-effectiveness

For general vitality, an ICER of 280 was found. This indicates that the additional societal costs per 1-point improvement in general vitality were €280. ICERs in similar directions were found for work-related vitality (ICER, 7506) and NFR (ICER, −258) (Table 4). Note that the ICER for NFR was negative because lower scores indicate a better NFR. In all cost-effectiveness planes, the majority of incremental cost-effect pairs were located in the northeast quadrant (Figure 2 [1A–1C]), indicating that the intervention was more expensive than usual care in obtaining an additional unit of effect. The uncertainty surrounding the cost-effectiveness was large, as is reflected by the wide distribution of incremental cost-effect pairs (Table 4). Cost-effectiveness acceptability curves are presented in Figure 2 (2A–2C). To illustrate, if society is not willing to pay anything to obtain a 1-point improvement in general vitality, there is a probability of 0.3 that the intervention is cost-effective. If society is willing to pay ±€3500, there is a probability of 0.9.

Table 4: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes

Analysis	Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
	I	C		€	points		NE[1]	SE[2]	SW[3]	NW[4]
Main analysis	367	363	General vitality	570 (-1968 – 2905)	2.0 (-0.2 - 4.3)	280	66.4	29.8	1.4	2.4
Imputed dataset	367	363	Work-related vitality	570 (-1968 – 2905)	0.08 (-0.02 - 0.18)	7506	63.8	29.5	1.7	5.1
	367	363	Need for recovery	570 (-1968 – 2905)	-2.2 (-5.5 – 1.1)	-258	62.4	28.5	2.8	6.6
SA1	199	191	General vitality	-96 (-1578 – 1237)	2.1 (-0.6 - 4.9)	-44	41.4	53.6	2.2	2.8
Complete dataset	199	191	Work-related vitality	-96 (-1578 – 1237)	0.08 (-0.04 - 0.2)	-1167	40.4	51.6	4.0	4.0
	199	191	Need for recovery	-96 (-1578 – 1237)	-3.7 (-7.5 - 0.1)	26	42.5	54.6	1.3	1.6
SA2	367	363	General vitality	557 (-1979 - 2892)	2.0 (-0.2 - 4.3)	274	66.2	30.2	1.4	2.2
Prices paid	367	363	Work-related vitality	557 (-1979 - 2892)	0.08 (-0.02 - 0.18)	7338	63.5	29.7	1.6	5.2
	367	363	Need for recovery	557 (-1979 - 2892)	-2.2 (-5.5 – 1.1)	-253	61.9	28.8	2.8	6.5
SA3	367	363	General vitality	604 (-2251 - 3184)	2.0 (-0.2 - 4.3)	297	66.0	30.2	1.4	2.3
HCA, elasticity: 1.0	367	363	Work-related vitality	604 (-2251 - 3184)	0.08 (-0.02 - 0.18)	7956	63.3	29.8	1.8	5.1
	367	363	Need for recovery	604 (-2251 - 3184)	-2.2 (-5.5 – 1.1)	-274	62.2	28.5	2.8	6.4
SA4	367	363	General vitality	464 (-1261 – 1911)	2.0 (-0.2 - 4.3)	228	71.2	25.2	1.2	2.4
Excluding presenteeism	367	363	Work-related vitality	464 (-1261 – 1911)	0.08 (-0.02 - 0.18)	6137	68.4	24.8	1.4	5.4
	367	363	Need for recovery	464 (-1261 – 1911)	-2.2 (-5.5 – 1.1)	-211	66.8	23.7	2.5	7.0

Abbreviations: CI: confidence interval, C: costs, E: effects, ICER: incremental cost-effectiveness ratio, CE-plane: cost-effectiveness plane, I: Intervention, C: Control, SA: sensitivity analysis, HCA: human capital approach

Note: Costs are expressed in 2010 Euros

[1] Refers to the northeast quadrant of the CE plane, indicating that the Vital@Work intervention is more effective and more costly compared to the control group

[2] Refers to the southeast quadrant of the CE plane, indicating that the Vital@Work intervention is more effective and less costly compared to the control group

[3] Refers to the northwest quadrant of the CE plane, indicating that the Vital@Work intervention is less effective and more costly compared to the control group

[4] Refers to the southwest quadrant of the CE plane, indicating that the Vital@Work intervention is less effective and less costly compared to the control group

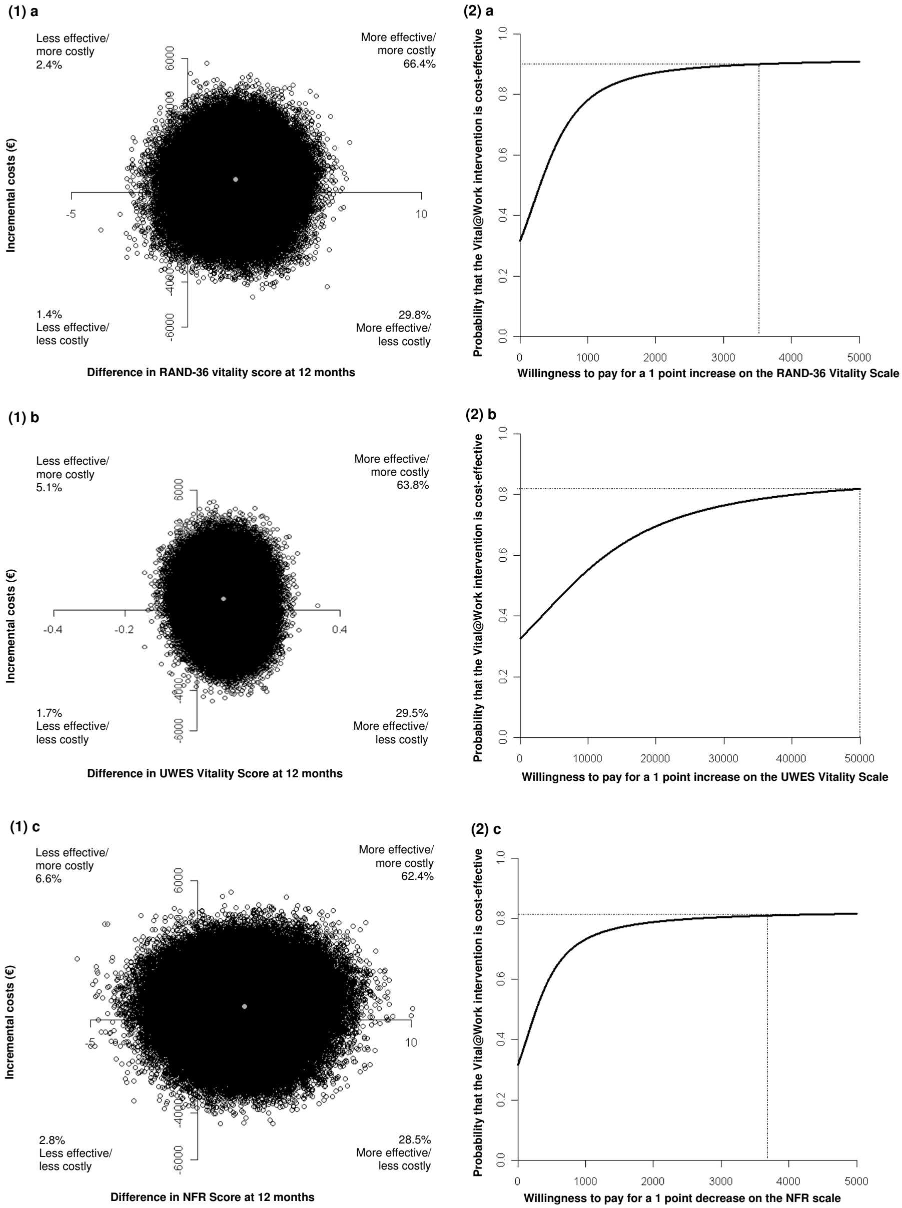

Figure 2: Cost-effectiveness planes indicating the uncertainty around the incremental cost-effectiveness ratios (1) and cost-effectiveness acceptability curves indicating the probability of cost-effectiveness for different values (€) of willingness to pay per unit of effect gained (2) for general vitality (a), work-related vitality (b), and need for recovery (c) (based on the imputed dataset).

Employer's perspective: Financial return

During follow-up, average absenteeism (−€223; 95% CI, −1636 to 1284) and presenteeism (−€106; 95% CI, −1650 to 1454) benefits per worker were negative, suggesting that the intervention increased productivity-related spending (Table 5). The NB was on average −€478 (95% CI, −2663 to 1816) per worker, suggesting a net loss to the employer of €478. Nevertheless, as indicated by the 95% CIs, the uncertainty surrounding the benefit estimates and NB was large and they cannot be regarded as statistically significant. The BCR (ie, amount of money returned per euro invested) and ROI (ie, percentage of profit per euro invested) were −2.21 and −321%, respectively (11). Overall, these findings suggest that the intervention was not cost saving to the employer during the 12-month follow-up.

Sensitivity analyses

The overall conclusions would not change when using the results from SA2 (using prices paid), SA3 (using HCA), and SA4 (excluding presenteeism) (Tables 4 and 5). When solely analyzing the complete cases (SA1), however, total societal costs were lower in the intervention than in the control group, whereas they were highest in the intervention group according to the main analysis. This difference is mostly explained by differences in presenteeism costs, which were lowest in the intervention group among the complete cases, whereas they were lowest in the control group after multiple imputation (Table 3). Effect sizes, on the contrary, were about the same in both analyses. In the complete-case analysis, the majority of the incremental cost-effect pairs were located in the southeast quadrant of the cost-effectiveness plane, indicating that the intervention was less expensive than usual care to obtain an additional unit of effect. Nevertheless, the uncertainty surrounding this cost-effectiveness was large. For the employer, the complete-case analysis resulted in an NB of €59 (95% CI, −1137 to 1471), a BCR of 1.40, and an ROI of 40%, indicating that the intervention produced a positive financial return. Again, however, the range of uncertainty was large.

Table 5: Intervention costs, benefits, net benefit (NB), benefit cost ratio (BCR), and return-on-investment (ROI) per worker

Analysis	Sample size		Costs	Benefits			Financial return		
	I	C	I (95% CI)	Absenteeism (95% CI)	Presenteeism (95% CI)	Total (95% CI)	NB[1] (95% CI)	BCR[2]	ROI (%)[3]
Main analysis *Imputed dataset*	367	363	149 (NA)	-223 (-1636 – 1284)	-106 (-1650 – 1454)	-329 (-2514 – 1964)	-478 (-2663 – 1816)	-2.21	-321
SA1 *Complete dataset*	199	191	149 (NA)	-107 (-466 – 259)	315 (-855 – 1549)	208 (-993 – 1574)	59 (-1137 – 1471)	1.40	40
SA2 *Prices paid*	367	363	137 (NA)	-223 (-1636 – 1284)	-106 (-1650 – 1454)	-329 (-2514 – 1964)	-466 (-2651 – 1829)	-2.40	-340
SA3 *HCA, elasticity: 1.0*	367	363	149 (NA)	-257 (-2042 – 1674)	-106 (-1650 – 1454)	-363 (-2806 - 2261)	-512 (-2952 – 2115)	-2.44	-344
SA4 *Excluding presenteeism*	367	363	149 (NA)	-223 (-1636 – 1284)	NA	-223 (-1636 – 1284)	-372 (-1785 – 1135)	-1.50	-250

Abbreviations: CI: confidence interval, NB: net benefit, BCR: benefit cost ratio, ROI: return-on-investment, I: Intervention, C: Control, SA: sensitivity analysis, HCA: human capital approach, NA: not applicable
Note: Costs are expressed in 2010 Euros
[1] Indicates the amount of money returned after intervention costs are recovered
[2] Indicates the amount of money returned per Euro invested in the intervention
[3] Indicates the percentage of profit per Euro invested in the intervention

DISCUSSION

The present study aimed to assess the cost-effectiveness and financial return of a worksite vitality intervention among older workers versus usual care. No significant differences in effects and costs were found. The intervention can neither be regarded as cost-effective from the societal perspective nor cost saving from that of the employer.

Effects and Costs

The lack of effect on the study outcomes might be due to their baseline values already being in the upper limit range of those measures, leaving less room for improvement. This might indicate a "healthy worker effect" (ie, healthier workers are more likely to stay in the workforce than those who are sick or physically unfit). Another explanation might be that attendance and compliance were lower than expected among intervention group workers. The attendance rates, defined as the mean percentage of attended group sessions in relation to the number of provided group sessions (n = 24), for the yoga and workout sessions were 51.7% and 44.8%, respectively (35). Furthermore, 108 (29.4%) intervention group workers did not attend any of the yoga sessions and 133 (36.2%) did not attend any of the workout sessions (35).

Until now, few studies evaluated the effectiveness of worksite health promotion intervention in terms of vitality or NFR. One study (36) found a worksite intervention consisting of vegan nutrition education sessions to increase general vitality by 11.0 points (range, 0 to 100) at 22-week follow-up. Their results, however, were based on a nonrandomized study, making it difficult to attribute the effect to the intervention and to rule out the possibility that the study was biased by confounders or baseline differences in group characteristics (ie, selection bias) (10,37). Furthermore, the content of the intervention was different from that of the Vital@Work intervention, the intervention was not specifically aimed at older workers, and it is unknown whether the effect was sustained over the long term.

As for the lack of significant cost differences, it is known that cost data are highly skewed and therefore require large sample sizes to detect relevant differences (38).

In our study, the sample size calculation was based on work-related vitality (2), which may have underpowered it to detect significant cost differences. Although not significant, it is noteworthy that despite the fact that intervention group workers reported a larger increase in weekly sports activities compared with their control group counterparts (3), sports costs were lowest in the intervention group. Further examination of the data revealed that this was mainly due to the fact that intervention group workers purchased fewer sports memberships than those of the control group (data not shown). Therefore, a possible explanation for this finding may be that workers regarded the Vital@Work intervention as a substitute for a membership of a sports club.

Societal perspective: Cost-effectiveness

Joint comparison of costs and effects revealed that a substantial amount of money has to be paid by society to reach a reasonable probability of cost-effectiveness. For example, for a 0.9 probability of cost-effectiveness, society should be willing to pay ±€3500 per 1-point improvement in general vitality (range, 0 to 100). Although it is unknown what relevant improvements on the main study outcomes are, and this will depend on their baseline values, it may be in the 10% to 20% range. Therefore, although it is currently unknown how much decision makers are willing to pay for a 1-point improvement on both vitality measures and NFR, the present study provides no evidence to support the implementation of the Vital@Work intervention on cost-effectiveness grounds. One might argue that this was expected because the intervention did not have a significant effect on the main study outcomes. Nevertheless, CEAs are about the joint distribution of differences in costs and effects, which could even show clear cost-effectiveness when neither cost nor effect differences are individually significant (39).

Comparing these results with previous studies is hampered by the lack of studies evaluating the societal cost-effectiveness of similar interventions in terms of vitality or NFR. Nevertheless, the previously mentioned study did report the intervention costs of their worksite vegan nutrition intervention ($3614/16 participants; $226/participant) (36), but the authors did not measure any other cost and did not perform a full economic evaluation.

Employer's perspective: Financial return

The ROI analysis indicated that the Vital@Work intervention cannot be regarded as cost saving to the employer. So far, only one other study (40) evaluated the financial return of a similar intervention in terms of both absenteeism and presenteeism benefits. On average, this worksite physical activity and nutrition program, consisting of a health risk assessment, a Web portal, and lifestyle seminars, resulted in a reduction of 4.3 absenteeism days (absenteeism benefits: $1236) and a 0.79-point (range, 0 to 10) increase in work performance (presenteeism benefits, $1364). Combining these findings with the reported intervention costs ($138/participant) results in a BCR of 18.84 and an ROI of 1784% (10). These findings differ enormously from those of our study, which might be explained by differences in intervention content, intervention participants (older workers vs general working population), study design (RCT vs nonrandomized study) or a combination of these. The latter is underscored by a recent systematic review, which indicated that worksite physical activity, nutrition programs, or both generate positive financial returns through reduced absenteeism, medical costs, or both according to nonrandomized studies, whereas they do not according to RCTs (10).

Robustness of study results

Sensitivity analyses revealed that the present findings were reasonably robust with respect to the valuation of intervention and absenteeism costs. Excluding presenteeism costs did not change the conclusions either. Nevertheless, differences were found between the main analysis, for which data were imputed, and the complete-case analysis. These differences were mainly caused by differences in presenteeism costs. This may be due to the complete cases being unrepresentative of the whole study population in terms of (presenteeism) costs and, therefore, not satisfying the missing completely at-random assumption (ie, the "missingness" of data does not depend on the unobserved or the observed data) required for a complete-case analysis to provide valid and unbiased results (32).

Strengths and limitations

Several strengths of the present study are noteworthy. First, analyses were performed alongside a pragmatic RCT, which is generally acknowledged as the best vehicle for economic evaluations as it enables the evaluation of an intervention's economic consequences under "real life" conditions and allows prospective collection of relevant cost and effect data (39,41). So far, few studies have used this design to evaluate the financial return of worksite physical activity or nutrition programs, although their results seem to differ from those of nonrandomized studies with a higher risk of bias (10). Second, the CEA was conducted from the societal perspective. Until now, many studies evaluating the cost-effectiveness of similar interventions have applied a rather restrictive perspective by including only intervention costs (42). Worksite interventions, however, are also thought to be associated with medical and productivity-related costs. Both were included in the present study as a result of the adoption of the societal perspective. Third, the present study was one of the first CEAs and ROI analyses of worksite physical activity or nutrition programs to incorporate presenteeism costs (10), which can represent a considerable proportion of total productivity-related costs (43). Nevertheless, it is important to mention that a "gold standard" for estimating presenteeism costs does not exist currently (25). Further research is needed to develop more sophisticated instruments for measuring and valuing presenteeism and to reach consensus about the best way to do so. Until then, the method used in the present study provides at least a crude estimate of the presenteeism costs associated with a worksite vitality intervention.

A first limitation concerns the amount of incomplete data. For 360 workers (48%), complete follow-up data were missing. This is comparable with the amount of missing data in other CEAs of worksite interventions that were conducted alongside RCTs with a follow-up of 1 year or more (44,45). Multiple imputation was used to deal with the missing data, which is acknowledged as a more appropriate way to deal with missing data than complete-case analyses (46). Complete-case analysis will always be inefficient, to some degree, as the sample size is reduced and it will ignore observed cost data, effect data, or both in the excluded participants (32). Multiple imputation, however, relies on the assumption that data are missing at random (ie, the "missingness" depends only on the observed data and not on unobserved data), an assumption that may not necessarily hold true. Therefore, the results of the present study should be interpreted with caution. In future studies, every endeavor should be made to minimize the amount of missing data (32). Another limitation may be that cost and effect data were obtained through self-reported retrospective questionnaires, which may have caused “social desirability bias,” “recall bias,” or both. For example, participants’ health insurance claim data could not be used for calculating medical costs, as these are often practically inaccessible in the Netherlands. As a consequence, self-report of medical resource utilization is the most commonly used method in Dutch economic evaluations and was therefore used in the present study as well. Furthermore, it is unknown whether the results may be generalized to other working populations (ie, “external validity”), as the intervention was specifically tailored to older hospital workers.

Conclusion

The Vital@Work intervention was neither cost-effective from the societal perspective nor cost saving from that of the employer. Therefore, the present study provides no evidence to support its implementation on cost-related grounds.

REFERENCES

1. Siegrist J, Wahrendorf M, von dem Knesebeck O, Jurges H, Borsch-Supan A. Quality of work,well-being, and intended early retirement of older employees: baseline results from the SHARE Study. *Eur J Public Health*. 2007;17:62–68.
2. Strijk J, Proper K, van der Beek A, van Mechelen W. The Vital@Work Study. The systematic development of a lifestyle intervention to improve older workers' vitality and the design of a randomised controlled trial evaluating this intervention. *BMC Public Health.* 2009;9:408.
3. Strijk JE, Proper KI, van der Beek AJ, van Mechelen W. A worksite vitality intervention to improve older workers' lifestyle and vitality-related outcomes: results of a randomised controlled trial. *J Epidemiol Community Health.*2012.
4. Strijk JE, Proper KI, van der Beek AJ, van Mechelen W. A worksite vitality intervention for older hospital workers to improve vitality, work engagement, productivity, and sick leave: results of a randomised controlled trial. *Scand J Work Environ Health.* 2012; (pii):3311.
5. Burdorf A. Economic evaluation in occupational health—its goals, challenges, and opportunities. *Scand J Work Environ Health.* 2007;33:161–164.
6. Centers for Disease Control and Prevention, US Department of Health and Human Services. *Introduction to economic evaluation*. Available at: http://www.cdc.gov/owcd/EET/SeriesIntroduction/1.html. Accessed September 18, 2011.
7. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes.* New York, NY: Oxford University Press Inc; 2005.
8. Miller P, Haslam C. Why employers spend money on employee health: interviews with occupational health and safety professionals from British industry. *Safety Sci.* 2009;47:163–169.
9. Cavallo D. Using return on investment analysis to evaluate health promotion programs: challenges and opportunities. Health promotion economics issue briefs. 2006;1:1–4. RTI-UNC Center of Excellence. Available at: http://www.rti.org/pubs/IssueBrief 3.pdf. Accessed October 12, 2010.
10. van Dongen JM, Proper KI, van Wier MF, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev.* 2011;12:1031–1049.
11. Phillips JJ. *Return on Investment in Training and Performance Improvement Programs.* Burlington, MA: Elsevier; 2003.
12. Tompa E, Dolinschi R, de Oliveira C. Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil.* 2006;16:367–392.
13. Schaufeli WB, Bakker AB. *Utrecht Work Engagement Scale*. Utrecht, the Netherlands: Occupational Health Psychology, Unit Utrecht University; 2003.
14. Schaufeli WB, Bakker AB. Bevlogenheid: een begrip gemeten. *Gedrag en Organisatie.* 2004;17:90–112.
15. Crawford JO, Graveling RA, Cowie HA, Dixon K. The health safety and health promotion needs of older workers. *Occup Med (Lond).* 2010;60:184–192.

16. van der Zee KI, Sanderman R. *Het Meten van Gezondheidstoestand Met de RAND-36: een Handleiding*. Groningen, the Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 1993.
17. van Veldhoven M, Broersen S.Measurement quality and validity of the "need for recovery scale." *Occup Environ Med.* 2003;60:i3–i9.
18. de Croon EM, Sluiter JK, Frings-Dresen MHW. Psychometric properties of the need for Recovery after work scale: test-retest reliability and sensitivity to detect change. *Occup Environ Med.* 2006;63:202–206.
19. Statistics Netherlands. *Consumer Prices*. Den Haag, the Netherlands: Centraal Bureau voor de Statistiek. 2011. Available at: http://www.cbs nl/nl-NL/menu/themas/inkomen-bestedingen/cijfers/default.htm. Accessed September 10, 2011.
20. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. *Handleiding Voor Kostenonderzoek*. Methoden en Standaardkostprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen, the Netherlands: College Voor Zorgverzekeringen; 2010.
21. Koopmanschap M, Meeding WJ, Evers S, Severens J, Burdorf A, Brouwer W. *Handleiding Voor het Gebruik van PRODISQ Versie 2.1*. Maastricht, the Netherlands: Rotterdam/Maastricht, Erasmus MC—Instituut voor Medical Technology Assessment, Instituut Maatschappelijke Gezondheidszorg, Universiteit vanMaastricht—Beleid Economie en Organisatie van de Zorg; 2004.
22. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev PharmacoeconOutcomes Res.* 2005;5:23–28.
23. Belastingdienst. *Tax Income Rates*. Available at: http://belastingdienst.nl/variabel/buitenland/loonheffingen/loonheffingen-78html#P3232_117003. Accessed September 10, 2011.
24. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ.* 1995;14:171–189.
25. Prasad M, Wahlqvist P, Shikiar R, Shih YCT. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. *Pharmacoeconomics.* 2004;22:225–244.
26. Kessler R, Ames M, Hymel P, Loeppke R, McKenas D, Richling D. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med.* 2004;46:S23–S37.
27. Kessler R, Barber C, Beck A, Berglund P, Cleary P, McKenas D. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med.* 2003;45:156–174.
28. Kessler RC, Petukhova M, McInnes K, Üstün TB. *Content and scoringrules for the WHO HPQ absenteeism and presenteeism questions*. 2007. Available at: http://www.hcp.med.harvard.edu/hpq/ftpdir/absenteeism%20presenteeism%20scoring%20050107.pdf. Accessed July 25, 2011.
29. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res.* 2011;20:40–49.
30. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res.* 2007;16:219–242.
31. Rubin DB. *Multiple Imputation for Nonresponse in Surveys.* New York, NY: John Wiley & Sons; 1987.

32. Burton A, Billingham LJ, Bryan S. Cost-effectiveness in clinical trials: using multiple imputation to deal with incomplete cost data. *Clin Trials.*2007;4:154–161.
33. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making.* 1990;10:212–214.
34. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves—facts, fallacies, and frequently asked questions. *Health Econ.* 2004;13:405–415.
35. Strijk JE, Proper KI, van der Beek AJ, van Mechelen W. A process evaluation of a worksite vitality intervention among ageing hospital workers. *Int J Behav Nutr Phys Act.* 2011;8:58.
36. Katcher HI, Ferdowsian HR, Hoover VJ, Cohen JL, Barnard ND. A worksite vegan nutrition program is well-accepted and improves health-related quality of life and work productivity. *Ann Nutr Metab.* 2010;56:245–252.
37. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Oxford, Oxfordshire, UK: The Cochrane Collaboration. Available at: http://www.cochrane-handbook.org. Accessed September 15, 2011.
38. Briggs A. Economic evaluation and clinical trials: size matters. *BMJ.* 2000;321:1362–1363.
39. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ.* 2011;342: d1548.
40. Mills PR, Kessler RC, Cooper J, Sullivan S. Impact of a health promotion program on employee health risks andwork productivity. *Am J Health Promot.*2007;22:45–53.
41. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. *Value Health.* 2005;8:521–533.
42. van Dongen J, Proper K, van Wier M, et al. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health.* 2012;38:393–408.
43. Burton WN, Conti DJ, Chen CY, Schultz AB, Edington DW. The economic burden of lost productivity due to migraine headache: a specific worksite analysis. *J Occup Environ Med.* 2002;44:523–529.
44. Robroek SJW, Polinder S, Bredt FJ, Burdorf A. Cost-effectiveness of a long-term Internet-delivered worksite health promotion programme on physical activity and nutrition: a cluster randomized controlled trial. *Health Educ Res.*2012;27:399–410.
45. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med.* 2011;53:610–617.
46. Marshall A, Billingham L, Bryan S. Can we afford to ignore missing data in cost-effectiveness analyses? *Eur J Health Econ.* 2009;10:1–3.

5

Cost-effectiveness and financial return of a mindfulness-based worksite intervention aimed at improving work engagement: results of a randomized controlled trial

Johanna M van Dongen
Jantien van Berkel
Cécile RL Boot
Judith E Bosmans
Karin I Proper
Paulien M Bongers
Allard J van der Beek
Maurits W van Tulder
Marieke F van Wier

Submitted

ABSTRACT

Objectives: To evaluate the cost-effectiveness and financial return of a mindfulness-based worksite intervention compared to usual practice.

Methods: 257 employees of two Dutch governmental research institutes were randomized to the intervention (n=129) or control group (n=128). Intervention group participants received an intervention consisting of mindfulness training, e-coaching, and several supporting elements. Data on work engagement, general vitality, job satisfaction, and work ability were collected at baseline, six, and 12 months. Salary and absence data were collected from company records. Data on healthcare utilization, work performance, the utilization of occupational health services and/or in-company health promotion activities as well as the participants' expenses on sports equipment and membership fees were collected using 3- or 6-monthly retrospective questionnaires. Missing data were imputed using multiple imputation. Cost-effectiveness analyses were conducted from both the societal and employer's perspective. A return on investment analysis was conducted from the employer's perspective. To test the robustness of the study results, various sensitivity analyses were conducted.

Results: Intervention costs were €171 per participant from the societal perspective (bottom-up micro-costed) and €464 from the employer's perspective (market prices). After 12 months, a statistically significant but not clinically relevant adverse effect on work engagement (-0.19; 95%CI -0.38 – -0.01) was found. There were no differences in job satisfaction (-0.02; 95%CI -0.22 – 0.17), general vitality (-3.0; 95%CI -6.1 – 0.1), work ability (-0.34; 95%CI -0.84 – 0.17), and total costs (societal: 1814; 95%CI -800 – 4588, employer: 2038; 95%CI -548 - 4752). Probabilities of cost-effectiveness were low (≤0.25) and the intervention did not result in a positive financial return to the employer.

Conclusion: The intervention was neither cost saving nor cost-effective. Therefore, the present study provided no evidence to support its implementation.

INTRODUCTION

During the last decades, daily working life has become more emotionally and mentally demanding (1). This may contribute to the development of mental health problems (1;2), of which up to 30% of the population worldwide experiences some form each year (3). Next to the human suffering associated with mental health problems, they negatively affect employers through reduced productivity and (in most European countries) the government and/or health insurance companies through increased health care costs. For example, in Europe alone, the annual costs of anxiety disorders and depression have been estimated at €136.3 billion, of which the majority (i.e. €99.3 billion) was due to productivity losses (4). Furthermore, mental health problems are one of the most important reasons for early retirement and withdrawal from the workforce on health-related grounds (5).

In the Mindful "Vitality in Practice" (VIP) study, a mindfulness-based intervention was developed aimed at promoting mental health among workers by improving their work engagement. Work engagement is defined as *"…a positive, fulfilling, work-related state of mind that is characterized by vigor (i.e. vitality), dedication, and absorption"* (6). Work engagement was found to be negatively associated with burnout, depression, and psychosomatic complaints (2;7). In addition, work engagement is considered to be related to increased job satisfaction and reduced turnover intentions (8;9) and plays an important role in the promotion of work ability (10). A mindfulness-based approach was chosen as it is thought to produce "psychological flexibility"; i.e. the propensity to persist with behaviors that are consistent with ones values and to desist from those that are not (11). This in turn may lead to improved work engagement and mental health (12).

Evaluations of the Mindful VIP intervention's effectiveness have been reported elsewhere (13;14). However, numerous occupational health interventions exist, of which only a limited number can be provided with the resources available (15). Therefore, decision-makers may not only be interested in the effectiveness of worksite interventions, but also in their (additional) costs per unit of effect gained (16). This is explored using "cost-effectiveness analyses" (CEAs), of which the results can be used to decide how resources should optimally be allocated to maximize health or welfare (17). Decision-makers, and employers in particular, may also be interested in the financial return of worksite interventions. This can be determined using “return on investment analyses” (ROI analyses), in which the costs of an intervention are compared to its financial benefits (18-20).

The aim of the present study was to conduct a CEA and ROI analysis comparing the Mindful VIP intervention to usual practice. As various stakeholders may be affected by the implementation of this intervention (e.g. employers, government, and health insurance companies), CEAs were performed from both the societal and employer’s perspective. Additionally, a ROI analysis was performed from the employer’s perspective.

METHODS

Study population and design

The present study was conducted alongside a 12-month randomized controlled trial (21), which took place in 2010 and 2011. All employees of two Dutch governmental research institutes were invited to participate. Participants were recruited through available communication channels (e.g. intranet, and employee magazine). Exclusion criteria were: 1) being on sickness absence for more than four weeks, and 2) being pregnant. All participants provided written informed consent. After baseline measurements, they were randomly allocated to the intervention or control group by a research assistant using a computer-generated randomization sequence in SPSS (v15.0, Chicago, IL). The research assistant was blinded to group allocation. As a result of the nature of the intervention, blinding of participants and trainers was not possible. The study protocol was approved by the Medical Ethics Committee of the VU University Medical Center (21).

Control and intervention condition

After randomization, all participants were granted access to an intranet webpage containing links to various health promotion activities of the participating research institutes (e.g. in-company fitness). Additionally, intervention group participants received the 6-month Mindful VIP intervention. The intervention and its development

have been fully described elsewhere (21). Briefly, the intervention consisted of mindfulness training, e-coaching, and several supporting elements (i.e. fruit and vegetables, lunch walking routes, and a buddy system).
The mindfulness training lasted eight weeks. Once a week, employees participated in 90-minute mindfulness group sessions (4-17 participants per group), given by certified trainers. Group compositions were the same during the training period, but participants were allowed to switch between groups if necessary. All sessions took place at the worksite and were scheduled so that participants could attend them outside working hours. Participants received hand-outs containing homework exercises, a mindfulness exercise booklet, and an audio disc with relaxation exercises. Participants formed pairs (i.e. the buddy system) for discussing homework exercises. An intranet webpage ("VIP webpage") was developed where participants could access the intervention materials at any time. The webpage also provided suggestions for lunch walking routes in the vicinity of the worksite. The mindfulness training was followed by an 8-week e-coaching trajectory to continue the implementation of the mindfulness principles, which was provided by the mindfulness trainers as well. If deemed necessary by the trainers, participants were referred to an occupational health service (e.g. occupational physician, occupational social worker, and career coach) and/or an in-company health promotion activity. Throughout the 6-month intervention period, free fruit and vegetables were provided at the worksite (21).

Effect measures

Effect measures were assessed at baseline, six, and 12 months and included work engagement, general vitality, job satisfaction, and work ability.
Work engagement was assessed using the "Utrecht Work Engagement Scale" (UWES). The UWES includes 17 items concerning three aspects of work engagement; 1) vigor (6 items), 2) dedication (5 items), and 3) absorption (6 items). All items were scored on a 7-point scale ranging from "never"(0) to "always"(6). The Work Engagement Score ranges from 0 to 6 (higher scores indicate a better work engagement) (6).
General vitality was assessed using the RAND-36 Vitality Scale, which includes four items assessing a participant's general vitality during the previous four weeks. Items were scored on a 6-point scale ranging from "all of the time"(1) to "never"(6). The

RAND-36 Vitality Score ranges from 0-100 (higher scores indicate a better general vitality) (22).
Job satisfaction was explored using a 1-item question of the “Netherlands Working Conditions Survey” (23). Participants were asked to rate their overall job satisfaction on a 5-point scale ranging from “very dissatisfied”(1) to “very satisfied”(5).
Work ability was explored using the Work Ability Index (WAI). The WAI originally consists of seven concepts (24). However, as sub-items can also be used as a concise indicator of work ability (25;26), only two concepts were used in the present study; 1) current work ability (one item, range: 0-10), and 2) work ability in relation to physical and mental job demands (two items, range: 0-10). The Work Ability Score was calculated by combining the raw scores of both WAI concepts and ranges from 0 to 20 (higher scores indicate a better work ability).

Resource use and valuation
For the societal perspective, bottom-up micro-costing was used to estimate intervention costs (i.e. data were collected regarding the quantity of resources consumed during the implementation of the intervention as well as their unit prices) (27). Intervention costs comprised those related to developing, implementing, and operating the Mindful VIP intervention (i.e. costs of mindfulness training, e-coaching, printed materials, fruit and vegetables, and “VIP webpage” hosting and maintenance). Frequency, duration, and preparation time of group meetings and e-mail contacts were registered by the trainers. Labor costs of intervention staff were valued by multiplying their total time investments by their gross hourly salaries including holiday allowances and premiums. Capital costs were valued using cost data collected from finance department staff. Costs of printed materials, the provision of fruit and vegetables, and website hosting and maintenance were estimated using invoices. Development costs were estimated by dividing the total costs related to the development of the intervention by the expected number of program users during the first five years after implementation. For the analyses conducted from the employer’s perspective, intervention costs were valued using market prices (i.e. the true cost to the employer, namely the amount of money employers have to pay when implementing the Mindful VIP intervention).

Health care utilization was assessed using 3-monthly retrospective questionnaires and included primary health care (i.e. general practitioner, allied health professionals, and complementary medicine), secondary health care (i.e. medical specialists, and hospitalization) and the use of prescribed and over-the-counter medications. Dutch standard costs were used to value health care utilization (28). If these were not available, prices according to professional organizations were used. Medication use was valued using unit prices provided by the Dutch Society of Pharmacy (29).

The use of occupational health services and/or in-company health promotion activities was assessed using 6-monthly retrospective questionnaires. Occupational health costs were valued using both a micro-costing approach (societal perspective) and market prices (employer's perspective). For both costing methods, information was collected from finance department staff.

Participants' expenses on sports membership fees and sports equipment were assessed using 3-monthly retrospective questionnaires.

Sickness absence at baseline (i.e. a one year period prior to baseline) and during follow-up as well as gross annual salaries of participants were collected from company records. Costs associated with one sick leave day were calculated per participant by dividing their gross annual salaries including holiday allowances and premiums by their total number of workable days per year (28). Holiday allowances and premiums were estimated according to the Dutch manual of costing (28). Using the Friction Cost Approach (FCA), absenteeism costs were estimated by multiplying the total number of sickness absence days during follow-up by their associated costs. Therefore, a friction period of 23 weeks was used, which is the time-span organizations need to replace a sick worker. Also, an elasticity of 0.8 was used, which implies that a 100% loss of work time corresponds with an 80% reduction in productivity (30).

Presenteeism was assessed on a 3-monthly basis using an item of "The World Health Organization Health and Work Performance Questionnaire" (WHO-HPQ). Participants were asked to rate their overall work performance during the previous three months on an 11-point scale, ranging from "worst performance"(0) to "best performance"(10). Subsequently, their average work performance during follow-up (*Wown*) was calculated. In the WHO-HPQ, presenteeism is conceptualized as a measure of actual work performance in relation to "best performance"(10), irrespective of the presence or absence of health complaints. Therefore, a participant's level of presenteeism (P_{HPQ}) was calculated using the following formula (31;32):

$$P_{HPQ} = (10 - Wown)/10$$

Presenteeism costs were calculated by multiplying a participant's P_{HPQ} by their gross annual salary including holiday allowances and premiums, corrected for absenteeism costs.

All costs were converted to 2011 Euros using consumer price indices (33). As the follow-up of the trial was one year, discounting of costs and effects was not necessary (34). An overview of the price weights used for valuing resource use can be found in Table 1 (Columns: 2-3).

Statistical analyses

Analyses were performed according to the intention-to-treat principle. Using descriptive statistics, baseline characteristics were compared between intervention and control group participants as well as those with complete and incomplete data. Missing data were imputed using multiple imputation. Imputations were performed separately for the intervention and control group. Amongst others, the imputation model included age, gender, number of working days, baseline sickness absence, baseline effect measure values, and available midpoint and follow-up cost and effect measure values (6- and 12 months). Using Fully Conditional Specification and Predictive Mean Matching, 10 complete data sets were created in PASW (V18, Chicago, IL) (Loss of Efficiency ≤5%) (35-37). Pooled estimates were calculated according to Rubin's rules (38). Furthermore, a descriptive analysis on resource use was performed. This analysis was based on the complete-cases, as missing data were imputed on the cost level rather than the level of resource use. T-tests were performed for normally distributed data. For skewed data, 95% confidence intervals (95%CIs) around the mean resource use differences were calculated using the bias-corrected and accelerated (BCA) bootstrap method (5000 replications). Unless otherwise stated, data were analyzed using Stata (V12, Stata Corp, College Station, TX). Statistical significance was set at $p<0.05$.

Table 1: Price weights used for valuing resource use in the Mindfulness VIP study and resources consumed during follow-up by the intervention and control group participants (based on the complete-cases)

Units [Units of measurement]	Price weight		Resources consumed	
	Societal perspective	Employer's perspective	Intervention group (n=91)	Control group (n=84)
Intervention costs	€ 171.45 [a]	€ 464.94 [b]	91	0
Medical costs				
Visits to a care provider [No. of visits; mean (SD)]			6.9 (9.9)	8.5 (13.8)
General practitioner				
Office consultation	€ 28.96 [c]	N.A.	1.6 (2.0)	1.9 (2.2)
Telephone consultation	€ 14.48 [c]	N.A.	0.3 (0.9)	0.5 (1.0)
House call	€ 44.47 [c]	N.A.	0.0 (0.1)	0.0 (0.2)
Allied health professionals			3.3 (6.2)	4.7 (11.7)
Psychologist	€ 82.47 [c]	N.A.	0.4 (2.0)	0.8 (3.6)
Social worker	€ 67.23 [c]	N.A.	0.3 (1.7)	0.1 (0.6)
Physical therapist	€ 37.23 [c]	N.A.	1.6 (3.6)	3.1 (9.4)
Remedial therapist	€ 36.20 [c]	N.A.	0.2 (1.3)	0.2 (1.4)
Other allied health professionals	Variable [c,d]	N.A.	0.9 (2.4)	0.6 (1.6)
Medical specialists			1.0 (2.4)	1.1 (2.3)
Psychiatrist	€ 106.53 [c]	N.A.	0.2 (1.7)	0.1 (0.3)
Other medical specialists	€ 74.47 [c]	N.A.	0.8 (2.0)	0.9 (1.7)
Complementary medicine	Variable [c,d]	N.A.	0.7 (2.9)	0.3 (1.7)
Hospitalization [No. of days; mean (SD)]				
Ward	€ 472.66 [c]	N.A.	0.3 (2.4)	0.2 (0.7)
Intensive care	€ 2257.82 [c]	N.A.	0.0 (0.0)	0.0 (0.0)
Medications [No. of participants using medication; Number (%)]	Variable [e]	N.A.	61.0 (67.0)	59.0 (70.0)
Absenteeism costs				
Sick leave [days; Mean (SD)]	Variable [f]	Variable [f]	6.6 (15.9)	4.4 (9.4)

Units [Units of measurement]	Price weight		Resources consumed	
	Societal perspective	Employer's perspective	Intervention group (n=91)	Control group (n=84)
Presenteeism costs				
Presenteeism [Presenteeism score; Mean (SD)]	Variable[g]	Variable[g]	0.26 (0.08)*	0.23 (0.07)*
Sports costs [No. of participants with sports costs; Number (%)]	Variable[h]	N.A.	76 (83.5)	68 (81.0)
Occupational health costs				
Occupational health specialists [No. of visits; Mean (SD)]				
Physician	€ 35.00[a]	€ 89.00[i]	0.0 (0.0)	0.0 (0.1)
Social Worker	€ 60.00[a]	€ 121.00[i]	0.0 (0.2)	0.0 (0.2)
Courses [No. of courses]				
Career counseling	€ 683.00[a]	€ 1025.00[j]	0.0 (0.2)	0.0 (0.2)
Pimp your resume	€ 23.00[a]	€ 55.00[j]	0.0 (0.0)	0.0 (0.1)
Other courses	Variable[a]	Variable[j]	0.0 (0.2)	0.0 (0.2)
In-company fitness [No. of months; mean (SD)]	€ 30.00[k]	€ 30.00[k]	2.7 (4.6)	2.8 (4.6)
Workplace massage [No. of massages; Mean (SD)]	€ 19.31[l]	€ 4.31[l]	1.1 (4.1)	1.9 (6.4)
Turnover costs [No. of employees leaving company]	€ 3523.80[m]	€ 3523.80[m]	5 (4)*	2 (2)*

* Significant at $p< 0.05$
Abbreviations: No.: Number; SD: Standard Deviation
Note: Costs are expressed in 2011 Euros
Price weight sources: [a] Bottum-up micro-costed, valued using tariffs and depleted sources (See Appendix 2); [b] Market prices, valued using invoices of contractors; [c] Dutch Manual of Costing; [d] Professional organizations; [e] Dutch Society of Pharmacy; [f] Gross annual salaries including holiday allowances and premiums according to the friction cost approach and an elasticity of 0.8; [g] Gross annual salaries including holiday allowances and premiums (corrected for absenteeism costs); [h] Self-reported expenses on sports memberships and sports equipment; [i] Market prices to the "Health and Safety Executive"; [j] Market prices to the career advice bureau; [k] Market prices to the in-company fitness provider (The total monthly costs were € 45.00, of which € 30.00 were paid by the employer and € 15.00 by the participants. The latter is included in the sports costs of participants); [l] Micro-costed, valued using tariffs and depleted sources (€ 4.31 is paid by the employer, and €15.00 by the participants); [m] Micro-costed according to Smith and Watkins (1978)

Cost-effectiveness analysis

CEAs in terms of work engagement and general vitality were conducted from the societal perspective (i.e. all costs related to the intervention were taken into account irrespective of who pays or benefits). CEAs in terms of work engagement, job satisfaction, and work ability were conducted from the employer's perspective (i.e. only costs relevant to Dutch employers were considered, including intervention, absenteeism, presenteeism, and occupational health costs). Effectiveness at 12-month follow-up was analyzed using linear regression, adjusted for baseline values. Unadjusted mean cost differences between the intervention and control group were calculated for total and disaggregated costs. Their 95%CIs were estimated by means of BCA intervals, with 5000 replications. Seemingly unrelated regression (SUR) analyses were performed in which effect differences were corrected for their baseline values and cost differences for baseline sickness absence (39). Subsequently, incremental cost-effectiveness ratios (ICERs) were calculated by dividing the corrected cost differences by those in effects. The uncertainty surrounding the ICERs was graphically illustrated by plotting bootstrapped incremental cost-effect pairs (CE-pairs) on cost-effectiveness planes (CE-planes) (40). A summary measure of the joint uncertainty of costs and effects was presented using cost-effectiveness acceptability curves (CEACs). CEACs indicate the intervention's probability of cost-effectiveness at different values of willingness-to-pay (41).

Employer's perspective: ROI analysis

The ROI analysis was performed from the employer's perspective. Costs were defined as intervention costs and benefits as the difference in total monetized outcome measures between the intervention and control group during follow-up (i.e. absenteeism, presenteeism, and occupational health costs). Using linear regression, benefits were adjusted for baseline sickness absence. Positive benefits indicate reduced spending. Three ROI-metrics were calculated; 1) Net Benefits (NB), 2) Benefit Cost Ratio (BCR), and 3) Return On Investment (ROI).

NB = Benefits – Costs

BCR = Benefits / Costs

*ROI = ((Benefits – Costs)/Costs)*100*

To quantify precision, 95% bootstrapped confidence intervals were estimated, using 5000 replications. Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0% (18-20).

Sensitivity analyses

To test the robustness of the results, six sensitivity analyses were performed. First, analyses were performed using the complete-cases only (SA1). Second, analyses were performed in which productivity losses were estimated using standard mean labor costs of the Dutch population (i.e. €30.90) (SA2) (28). Third, analyses were performed in which absenteeism costs were estimated using the Human Capital Approach (HCA) instead of the FCA (SA3). In the HCA, total sickness absence days are neither "truncated" to the friction period, nor is an elasticity factor applied. Fourth, analyses were performed in which presenteeism costs were estimated using a slightly modified version of the "PROductivity and DISease Questionnaire" (PRODISQ) (42;43). On a 3-montly basis, participants were asked to report the number of days during the previous two weeks at which they went to work while experiencing health complaints. If applicable, participants were asked to rate the quantity (Q1) and quality (Q2) of their work during these days on an 11-point scale ranging from "Nothing/Very bad quality"(0) to "Same as normal"(10). Assuming linearity, the number of workdays at which participants experienced some level of presenteeism was extrapolated over a 3-month period (W_{pres}). Per 3-month period, total workdays lost due to presenteeism were calculated using the following formula:

$$Pdays_{PRODISQ} = W_{pres} * (1-((Q1*Q2)/100))$$

Subsequently, the total number of workdays lost due to presenteeism during the complete follow-up period was estimated and valued using gross salaries of participants including holiday allowances and premiums. Fifth, due to the lack of overall consensus regarding the inclusion of presenteeism costs in economic evaluations, analyses were performed in which presenteeism costs were excluded (SA5). Sixth, as work engagement was previously found to be related to turnover intentions (10), we set out to include turnover costs in our main analysis. Unfortunately, however, only the number of participants leaving the company during follow-up could be ascertained, instead of the number that was replaced. Therefore, turnover costs were only included in a sensitivity analysis (SA6), which was based on the premise that all participants that left the company were replaced, by new ones. Turnover costs were estimated by multiplying the number of participants that left the company by the organizational costs associated with replacing one employee. Using information from finance and human resource department staff, these costs were estimated according to the costing model of Smith and Watkins (See Appendix 1) (44).

RESULTS

Participants

A total of 257 participants were randomized to the intervention (n=129) or control group (n=128) (Figure 1). At baseline, a relevant difference in sickness absence was found between both groups (Intervention group: Mean=7.1 (SD=25.7); Control group: Mean=3.8 (SD=9.5)) (Table 1). Complete follow-up data were obtained from 88% of participants on the effect measures (n=226; 118 intervention group participants and 108 control group participants) and from 71% on the cost measures (n=181; 91 intervention group participants and 90 control group participants). Data on the total number of provided mindfulness and e-coaching were complete. Relevant differences in terms of sickness absence days were also found between participants with complete and incomplete data in the intervention (Complete: 7.9 (SD=28.0); Incomplete: 3.4 (SD=5.1)) and control group (Complete: 2.5 (SD=5.0); Incomplete: 9.5 (SD=18.9)) (Table 2).

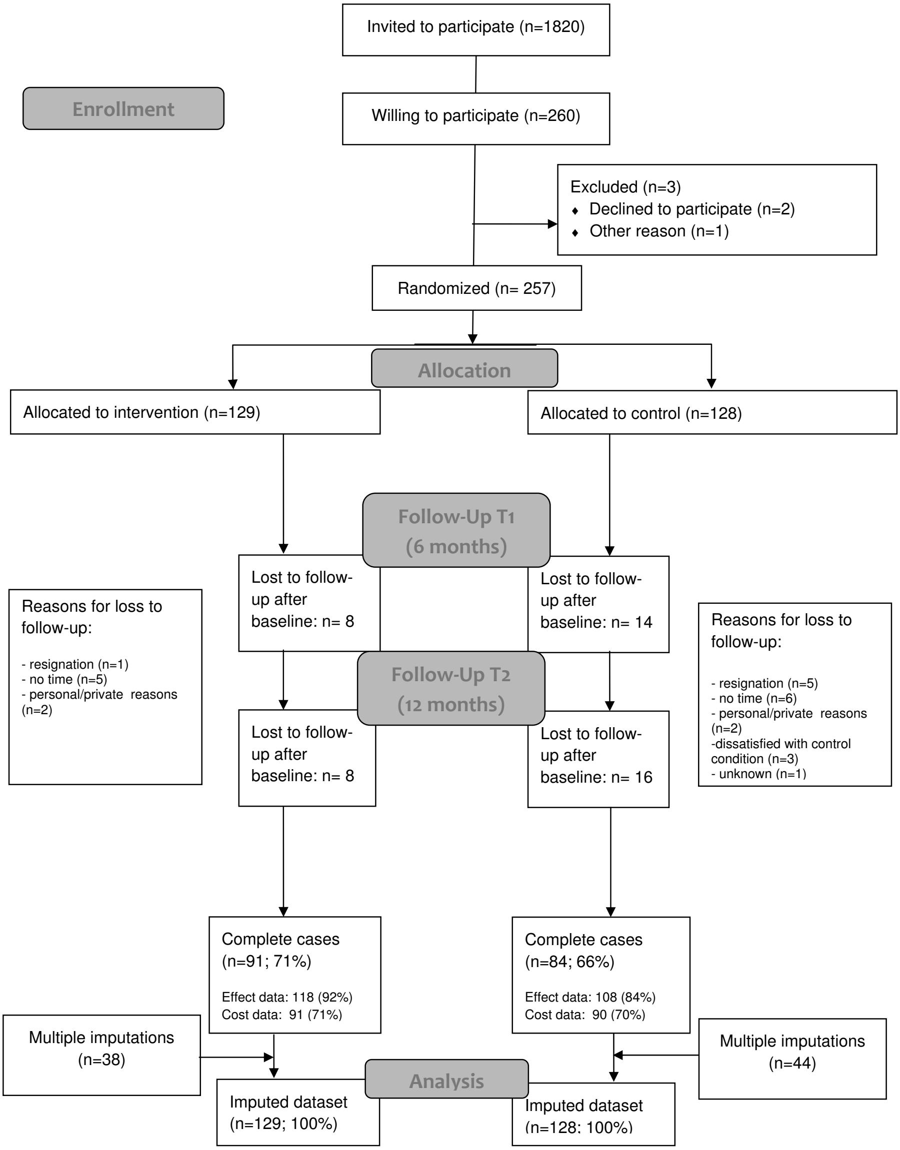

Figure 1: Flow diagram of participants in the Mindful VIP study

Table 2: Baseline characteristics of the study population

Baseline characteristics	Intervention group All (n=129)	Intervention group Complete (n=91)	Intervention group Incomplete (n=38)	Control group All (n=128)	Control group Complete (n=84)	Control group Incomplete (n=44)
Female [n (%)]	82 (63.6)	56 (61.5)	26 (68.4)	91 (71.1)	57 (67.9)	34 (77.3)
Age (years) [mean (SD)]	46.0 (9.5)	46.8 (8.8)	44.1 (10.7)	45.1 (9.6)	45.2 (8.9)	44.7 (10.9)
Marital status: married or significant other (n (%)]	105 (81.3)	78 (85.7)	27 (71.1)	94 (72.9)	68 (91.9)	26 (59.1)
Education level [n (%)][1]						
Low	3 (2.3)	2 (2.2)	1 (2.6)	2 (1.6)	1 (1.2)	1 (2.3)
Intermediate	27 (20.9)	19 (20.9)	8 (21.1)	16 (12.5)	8 (9.5)	8 (18.2)
High	99 (76.7)	70 (76.9)	29 (76.3)	110 (85.9)	75 (89.3)	35 (79.5)
Working hours per week [mean (SD)]	32.6 (5.6)	32.6 (5.6)	32.6 (5.7)	32.0 (5.6)	32.2 (5.0)	31.8 (6.7)
Work engagement (range: 0-6) [mean (SD)]	4.1 (0.8)	4.0 (0.8)	4.3 (0.9)	4.0 (0.9)	4.1 (0.8)	3.8 (1.0)
General Vitality (range: 0-100) [mean (SD)]	67.2 (16.7)	66.9 (16.7)	67.9 (16.9)	67.6 (18.7)	68.2 (18.6)	66.5 (19.2)
Job satisfaction (range: 1-5) [mean (SD)]	4.1 (0.7)	4.1 (0.6)	4.2 (0.8)	3.9 (0.8)	4.0 (0.6)	3.8 (0.8)
Work ability (range: 0-20) [mean (SD)]	16.6 (2.3)	16.5 (2.3)	16.9 (2.5)	16.3 (2.6)	16.6 (2.4)	15.7 (2.8)
Sickness absence: number of sickness absence days during the year prior to baseline [mean (SD)][2]	7.1 (25.7)	7.9 (28.0)	3.4 (5.1)	3.8 (9.5)	2.4 (5.0)	9.5 (18.9)
Work performance: WHO-HPQ work performance score during a 4-week period before baseline [mean (SD)]	7.9 (1.3)	7.8 (1.3)	8.0 (1.1)	7.9 (1.1)	8.0 (1.1)	7.7 (1.2)

Abbreviations: n: number, SD: standard deviation

[1] Education level was classified according to the definition of Statistics Netherlands (http://www.cbs.nl)

[2] Complete data were obtained from 109 intervention group participants and 103 control group participants

Effectiveness

A statistically significant difference in work engagement was found in favor of the control group. Compared to the control group, intervention group participants decreased their work performance (Range: 0-6) by 0.19 points (95%CI -0.38 - -0.01). No statistically significant differences were found for job satisfaction (-0.02; 95%CI –0.22 – 0.17), general vitality (-3.0; 95%CI -6.1 – 0.1), and work ability (-0.34; 95%CI -0.84 – 0.17).

Resource use

During the intervention period, 112 mindfulness sessions and 194 e-coaching sessions were provided. Based on the complete-cases, participants in the intervention and control group did not differ in terms of their average number of visits to a care provider (6.9 versus 8.5), average number of days of hospitalization (0.3 versus 0.3), and their average number of months of in-company fitness (2.7 versus 2.8). However, a statistically significant between-group difference was found in terms of their average presenteeism score (0.26 versus 0.23; p=0.01) (Table 1).

Costs

Average intervention costs per participant were €171 from the societal perspective and €464 from the employer's perspective (Appendix 2). Medical, absenteeism, and presenteeism costs were highest among intervention group participants. Sports and occupational health costs were highest among control group participants. However, none of these between-group differences were statistically significant (Table 3).

Table 3: Mean costs per participant in the intervention and control group, and unadjusted mean cost differences between both groups during the 12-month follow-up

Cost category	Intervention group n=129; mean (SEM)	Control group n=128; mean (SEM)	Mean cost difference (95%CI)
	Societal perspective		
Medical costs	588 (126)	495 (68)	94 (-116 – 472)
Sports costs	449 (44)	491 (52)	-42 (-180 – 86)
Occupational health costs	113 (15)	137 (19)	-24 (-75 – 23)
Absenteeism costs	2160 (423)	1413 (214)	746 (-14 – 1885)
Presenteeism costs	17293 (957)	16424 (904)	869 (-325 - 3930)
Intervention costs	171 (NA)	NA	171 (NA)
Total	20773 (1034)	18960 (963)	1814 (-800 – 4588)
	Employer's perspective		
Absenteeism costs	2160 (423)	1413 (214)	746 (-14 – 1885)
Occupational health costs	113 (17)	155 (28)	-42 (-118 – 15)
Presenteeism costs	17293 (957)	16424 (904)	869 (-325 - 3930)
Intervention costs	464 (NA)	NA	464 (NA)
Total	20029 (1012)	17992 (950)	2038 (-548 – 4752)

Abbreviations: n: number; SEM: Standard Error of the Mean, CI: Confidence Interval, NA: Not Applicable, SD: Standard Deviation
Note: Costs are expressed in 2011 Euros

Societal perspective: cost-effectiveness

For work engagement an ICER of -7321 was found, indicating that a 1-point decrease in work engagement was associated with a societal cost of €7321. An ICER in the similar direction was found for general vitality (ICER:-470). In both cases, the majority of incremental CE-pairs were located in the northwest quadrant of the CE-plane, indicating that the intervention was more costly and less effective than usual practice. This is graphically illustrated for work engagement (Figure 2). The uncertainty surrounding both cost-effectiveness estimates was large, as is reflected in the wide distribution of incremental CE-pairs (Figure 2, Table 4). The CEAC for work engagement presented in Figure 3 shows that the maximum probability of the intervention being cost-effective in comparison with usual practice was 0.17 regardless of the willingness to pay. For general vitality, the maximum probability was 0.17 as well (Figure not shown).

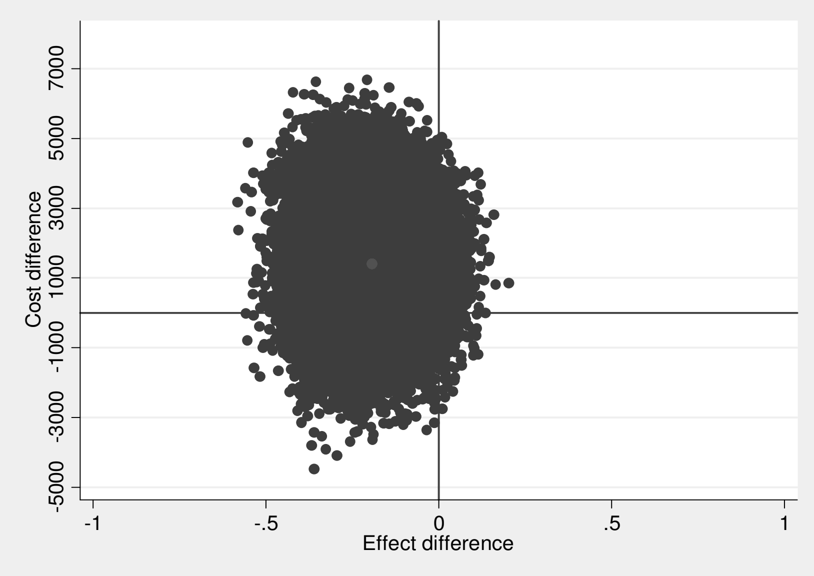

Figure 2: Costs-effectiveness plane for the difference in work engagement at 12-months (societal perspective)

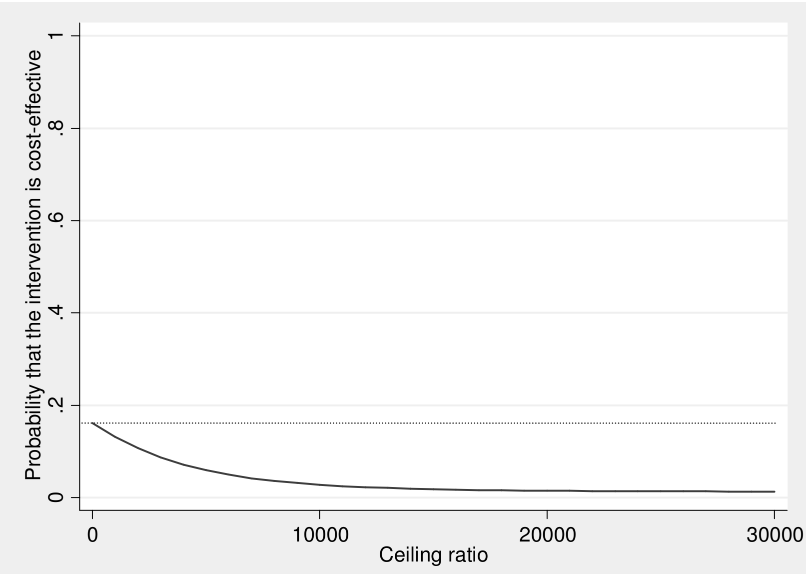

Figure 3: Cost-effectiveness acceptability curve for the difference in work engagement at 12-months (societal perspective)

Table 4: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes

Analysis		Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
				Societal perspective							
		Intervention	**Control**		**€**	**points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
Main analysis	- *Imputed dataset*	129	128	Work engagement (Range: 0-6)	1403 (-1198 – 4155)	-0.19 (-0.38 – -0.01)	-7321	1.5	0.3	14.9	83.3
		129	128	General vitality (Range 0-100)	1397 (-1201 – 4148)	-3.0 (-6.0 – 0.1)	-470	2.3	0.8	14.5	82.5
SA1	- *Complete dataset*	91	84	Work engagement (Range: 0-6)	2027 (-73 – 4070)	-0.15 (-0.34 – 0.05)	-13513	6.9	0.0	2.2	89.3
		91	84	General vitality (Range 0-100)	2049 (-12 – 4110)	-3.5 (-7.4 – 0.3)	-585	3.5	0.2	2.5	93.7
SA2	- *Mean labor costs*	129	128	Work engagement (Range: 0-6)	607 (-113 – 2470)	-0.19 (-0.38 – -0.01)	-3180	1.0	0.5	25.4	73.1
		129	128	General vitality (Range 0-100)	596 (-951 – 2597)	-3.0 (-6.0 – 0.1)	-200	1.2	0.5	17.9	80.4
SA3	- *HCA, elasticity: 1.0*	129	128	Work engagement (Range: 0-6)	1680 (-1206 – 4568)	-0.19 (-0.38 – -0.01)	-8769	1.6	0.3	11.5	86.6
		129	128	General vitality (Range 0-100)	1673 (-1027 – 4558)	-3.0 (-6.0 – 0.1)	-563	2.5	0.6	11.2	85.7
SA4	- *PRODISQ*	129	128	Work engagement (Range: 0-6)	368 (-1910 – 2539)	-0.19 (-0.38 – -0.01)	-1938	1.6	0.3	14.9	83.3
		129	128	General vitality (Range 0-100)	369 (-1906 – 2530)	-3.0 (-6.0 – 0.1)	-123	1.9	1.2	34.1	62.8
SA5	- *Excluding presenteeism costs*	129	128	Work engagement (Range: 0-6)	618 (-128 – 1569)	-0.19 (-0.38 – -0.01)	-3211	1.6	0.1	6.5	91.7
		129	128	General vitality (Range 0-100)	619 (-127 – 1565)	-3.0 (-6.0 – 0.1)	-209	2.9	0.2	6.3	90.6
SA6	- *Including turnover costs*	129	128	Work engagement (Range: 0-6)	1481 (-1120 – 4261)	-0.19 (-0.38 – -0.01)	-7731	1.6	0.3	13.6	84.5
		129	128	General vitality (Range 0-100)	1476 (-1122 – 4246)	-3.0 (-6.0 -0.1)	-496	2.4	0.7	13.2	83.7
				Employer's perspective							
		Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
Main analysis	- *Imputed dataset*	129	128	Work engagement (Range: 0-6)	1646 (-925 – 4323)	-0.19 (-0.38 – -0.01)	-8593	1.6	0.2	10.7	87.5
		129	128	Job satisfaction (Range: 1-5)	1641 (-924 – 4313)	-0.02 (-0.22 – 0.18)	-81295	36.5	4.7	6.3	52.6
		129	128	Work ability (Range 0-20)	1657 (-912 – 4331)	-0.32 (-0.81 – 0.16)	-5081	7.1	1.5	9.2	82.3
SA1	- *Complete dataset*	91	84	Work engagement (Range: 0-6)	2315 (205 – 4220)	-0.15 (-0.34 – 0.04)	-15433	6.8	0.0	1.1	92.0
		91	84	Job satisfaction (Range: 1-5)	2318 (292 – 4238)	-0.05 (-0.24 – 0.16)	-46360	30.5	0.3	0.9	68.4
		91	84	Work ability (Range 0-20)	2357 (386 – 4345)	-0.49 (-1.06 – 0.09)	-4810	4.9	0.0	1.1	94.0
SA2	- *Mean labor costs*	129	128	Work engagement (Range: 0-6)	848 (-965 – 2599)	-0.19 (-0.38 - -0.01)	-4455	1.5	0.4	16.8	81.2
		129	128	Job satisfaction (Range: 1-5)	845 (-967 – 2592)	-0.02 (-0.22 – 0.18)	-51124	34.3	8.2	9.1	48.4
		129	128	Work ability (Range 0-20)	854 (-959 – 2597)	-0.33 (-0.81 – 0.16)	-2645	6.3	1.4	14.6	76.7

5

Analysis		Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
				Employer's perspective							
		Intervention	Control		€	points	€/point	NE[1]	SE[2]	SW[3]	NW[4]
		129	128	Work ability (Range 0-20)	1933 (-721 – 4753)	-0.33 (-0.81 – 0.16)	-5930	7.4	1.1	7.1	84.4
SA4	- *PRODISQ*	129	128	Work engagement (Range: 0-6)	611 (-1633 – 2707)	-0.19 (-0.38 – -0.01)	-3174	1.3	0.5	26.6	71.6
		129	128	Job satisfaction (Range: 1-5)	616 (-1623 – 2704)	-0.02 (-0.22 – 0.18)	-30952	28.2	13.1	13.8	44.9
		129	128	Work ability (Range 0-20)	616 (-1622 – 2706)	-0.33 (-0.82 – 0.15)	-1851	5.4	2.6	24.3	67.7
SA5	- *Excluding presenteeism*	129	128	Work engagement (Range: 0-6)	861 (216 – 1747)	-0.19 (-0.38 - -0.01)	-4478	1.8	0.0	0.7	97.4
		129	128	Job satisfaction (Range: 1-5)	862 (218 – 1755)	-0.02 (-0.22 – 0.18)	-45055	41.3	0.4	0.3	58.0
		129	128	Work ability (Range 0-20)	861 (217 – 1743)	-0.33 (-0.82 – 0.15)	-2581	8.0	0.0	0.7	91.3
SA6	- *Including turnover costs*	129	128	Work engagement (Range: 0-6)	1725 (-853 – 4413)	-0.19 (-0.37 - -0.01)	-9003	1.7	0.2	9.6	88.5
		129	128	Job satisfaction (Range: 1-5)	1720 (-854 – 4404)	-0.02 (-0.22 – 0.18)	-85297	37.0	4.2	5.7	53.1
		129	128	Work ability (Range 0-20)	1735 (-846 – 4412)	-0.33 (-0.81 – 0.16)	-5318	7.3	1.3	8.3	83.1

Abbreviations: CI: Confidence Interval, C: Costs, E: Effects, ICER: Incremental Cost-Effectiveness Ratio, CE-plane: Cost-Effectiveness plane, SA: Sensitivity Analysis, HCA: Human Capital Approach

Note: Costs are expressed in 2011 Euros

[1] Refers to the northeast quadrant of the CE plane, indicating that the Mindful VIP intervention is more effective and more costly compared to the control group

[2] Refers to the southeast quadrant of the CE plane, indicating that the Mindful VIP intervention is more effective and less costly compared to the control group

[3] Refers to the northwest quadrant of the CE plane, indicating that the Mindful VIP intervention is less effective and more costly compared to the control group

[4] Refers to the southwest quadrant of the CE plane, indicating that the Mindful VIP intervention is less effective and less costly compared to the control group

Employer's perspective: cost-effectiveness

For work engagement an ICER of -8593 was found, indicating that a 1-point decrease in work engagement was associated with an employer's cost of €8593. ICERs in the similar direction were found for job satisfaction (ICER:-81295) and work ability (ICER:-5081). In all cases, the majority of incremental CE-pairs were located in the northwest quadrant of the CE-plane (Table 4), indicating that the intervention was more costly and less effective than usual practice. Irrespective of the willingness to pay, the associated maximum probabilities of cost-effectiveness were 0.13 (work engagement), 0.25 (job satisfaction), and 0.12 (work ability) (Figures not shown).

Employer's perspective: financial return

During follow-up, total benefits in terms of absenteeism, presenteeism, and occupational health costs were on average €-1170 (95%CI: -3760 – 1486) (Table 5). The NB was on average €-1635 (95%CI: -4268 – 973), which suggests that the intervention was associated with a net loss to the employer of €1635 per participant. The BCR (i.e. amount of money returned per Euro invested) and ROI (i.e. percentage of profit per Euro invested) were -2.51 (95%CI: -8.19 – 3.10) and -315% (95%CI: -919 – 210), respectively. None of these estimates was statistically significant.

Sensitivity analyses

Results of SA2, SA3, and SA6 were similar to those of the main analysis, whereas the outcomes of SA1, SA4, and SA5 differed in some aspects from those of the main analysis (Table 4, Table 5). Three differences stand out. First, work engagement significantly decreased among intervention group participants compared to their control group counterparts in the main analysis, whereas this difference was non-significant among the complete-cases (SA1). Second, the total cost differences between the intervention and control group were considerably lower in SA4 (PRODISQ) compared to those of the main analysis. Third, financial return estimates were statistically significant among the complete-cases (SA1) and when presenteeism costs were excluded (SA5), whereas they were not in the main analysis.

Table 5: Intervention costs, benefits, Net Benefits (NB), Benefit Cost Ratio (BCR), and Return-On-Investment (ROI) per participant (employer's perspective)

Analysis		Sample size		Costs	Benefits		Financial return	
		I	C	€	Total (95% CI)	NB[1] (95% CI)	BCR[2] (95% CI)	ROI (%)[3] (95% CI)
Main analysis	- *Imputed dataset*	129	128	464	-1170 (-3760 – 1486)	-1635 (-4268 – 973)	-2.51 (-8.19 – 3.10)	-351 (-919 – 210)
SA1	- *Complete dataset*	91	84	464	-1857 (-3833 – 134)	-2321 (-4341 - -376)	-4.00 (-8.35 – 0.19)	-500 (-935 – -81)
SA2	- *Mean labor costs*	129	128	464	-397 (-2165 – 1395)	-861 (-2618 – 940)	- 0.85 (-4.64 – 3.02)	-185 (-564 – 202)
SA3	- *HCA, elasticity: 1.0*	129	128	464	-1404 (-4154 – 1324)	-1869 (-4648 – 832)	-3.02 (-9.01 – 2.79)	-402 (-1001 – 179)
SA4	- *PRODISQ*	129	128	464	-144 (-2212 – 2111)	-609 (-2689 – 1636)	-0.31 (-4.79 – 4.52)	-131 (-597 – 352)
SA5	- *Excluding presenteeism*	129	128	464	-396 (-1130 – 352)	-860 (-1660 - -166)	-0.85 (-2.57 – 0.64)	-185 (-357 - -36)
SA6	- *Including turnover costs*	129	128	464	-1249 (-3842 – 1414)	-1712 (-4354 – 899)	-2.69 (-8.38 – 2.94)	-369 (-938 – 194)

Abbreviations: CI: Confidence Interval, NB: Net Benefit, BCR: Benefit Cost Ratio, ROI: Return-On-Investment, I: Intervention, C: Control, SA: Sensitivity Analysis, HCA: Human Capital Approach

Note 1: Costs are expressed in 2011 Euros

Note 2: Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0

[1] Indicates the amount of money returned after intervention costs are recovered

[2] Indicates the amount of money returned per Euro invested in the intervention

[3] Indicates the percentage of profit per Euro invested in the intervention

DISCUSSION

This study evaluated the cost-effectiveness and financial return of a mindfulness-based worksite intervention aimed at improving work engagement versus usual practice. The intervention had a statistically significant adverse but non-relevant effect (i.e. ≤10% decrease) on work engagement (21), no significant effect on all other cost and effect measures, and low probabilities of cost-effectiveness from both the societal and employer's perspective. In addition, the intervention did not generate a positive financial return to the employer. As such, the present study does not provide evidence for its implementation for economic reasons.

Effects and costs

A possible explanation for the lack of positive effects could be the low compliance with some of the intervention components. Although more than half of the participants (55%) were highly compliant with the mindfulness sessions (i.e. they attended ≥75% of the 8 provided sessions), only 8% were considered highly compliant with the homework exercises, and only 6% with the e-coaching sessions (45). The latter were provided after the 8-week mindfulness training in order to continue the implementation of the mindfulness principles (21). It is therefore likely that participants did not sufficiently integrate these principles into their daily (work) life to affect their work engagement at 12-month follow-up. This reasoning is supported by the fact that participants reported to feel *'revitalised'*, *'fresh'*, *'energetic'*, and *'peaceful'* after the mindfulness sessions, whereas they indicated that this effect faded away with time (45). This might also explain why previous studies, with relatively short follow-up durations (Range: 8–26 weeks, Mean: 12.4 weeks), found mindfulness-based worksite interventions to be effective in terms of various stress- and/or work-related outcomes (46-50), whereas the present study did not. Other factors, however, might also account for these conflicting results, including differences in study population and study design (e.g. three of the five studies were based on a non-randomized design (46;48;49)). Another explanation for the lack of positive effects might be that the favourable long-term effects found among various clinical populations are not necessarily generalizable to a relatively healthy workforce (46;51). However, more long-term RCTs are needed to establish this.

Regarding the lack of significant cost differences, it is known that cost data are heavily right skewed and therefore require large sample sizes to detect relevant differences (52). As in most RCTs, however, the sample size calculation was based on the primary outcome (i.e. work engagement) (21). This probably underpowered the present study to detect relevant cost differences (52). Although statistically non-significant, it is worth mentioning that absenteeism and presenteeism costs were highest among intervention group participants who also had significantly lower levels of work engagement in comparison to the control group. This is in line with previous research findings that found work engagement to negatively predict sickness absence duration and frequency and to positively predict overall work performance (2;53).

Cost-effectiveness

Joint comparison of costs and effects revealed that the maximum probability of cost-effectiveness was very low for all outcome measures, irrespective of the willingness-to-pay (≤0.25). Therefore, the present study provides no evidence to support implementation of the intervention on cost-effectiveness grounds.

Although various studies evaluated the effectiveness of mindfulness-based worksite interventions, none of them conducted a CEA. One study (54), however, did evaluate the societal cost-effectiveness of a worksite vitality intervention (i.e. work-out, yoga, coaching, and free fruit) versus usual practice among older hospital workers. Per 1-point increase in general vitality, the additional societal costs were found to be €280. Although this ICER is considerably more favourable than that of the present study (i.e. -470), the intervention was not considered to be cost-effective either, as a substantial amount of money had to be paid by society to reach a reasonable probability of cost-effectiveness (i.e. €3500 for a probability of 0.9).

Financial return

The ROI analysis indicated that the Mindful VIP intervention cannot be regarded as cost saving to the employer. Other studies evaluating the financial return of mindfulness-based worksite interventions are lacking. However, the previously mentioned study (54) also estimated the financial return of the worksite vitality intervention. In terms of absenteeism and presenteeism, a BCR of -2.21 and a ROI of -321% were found.

Although these financial return estimates are similar to those of the present study, a higher benefit decrease (€-1170 versus €-329) and higher intervention costs (€464 versus €149) were found in the present study. Also, occupational health costs were included in the present study, but not in the previous one.

Robustness of study results

The outcomes of the sensitivity analyses differed in some aspects from those of the main analysis (i.e. value sensitivity). First, the effect on work engagement was statistically significant in the main analysis (for which data were imputed), but not in the complete-case analysis. As their mean effect differences were comparable (i.e. -0.19 versus -0.15), this is likely explained by the increased power resulting from multiply imputing the missing values. Second, the total cost difference between the intervention and control group was considerably smaller when presenteeism costs were estimated using a slightly modified version of the PRODISQ, instead the WHO-HPQ. This was due to the fact that presenteeism costs were highest among intervention group participants when using the WHO-HPQ, but lowest when using the modified-PRODISQ (data not shown). Both instruments likely produced different results because they conceptualize presenteeism in a slightly different way (WHO-HPQ: reduced overall work performance, modified-PRODISQ: reduced work performance due to health complaints). The WHO-HPQ was used in the main analysis, as work engagement was found to be positively related to overall work performance (53), whereas evidence regarding an association between work engagement and reduced work performance due to health complaints is currently lacking. Third, in the main analysis, financial return estimates were not statistically significant, whereas they were statistically significant among the complete-cases and when presenteeism costs were excluded. The first may be due to the complete-cases being unrepresentative of the whole study population in terms of their total costs and therefore not satisfying the "Missing Completely At Random" assumption (i.e. the "missingness" of data does neither depend on the observed nor the unobserved data) required for a complete-case analysis to provide valid results (35-37). Based on the latter, it can be concluded that the intervention generated a financial loss to the employer in terms of absenteeism and occupational health costs alone.

None of the sensitivity analyses, however, led to changes in the overall conclusion about whether or not to implement the intervention (i.e. decision sensitivity). As CEAs and ROI analyses are conducted to inform implementation decisions, the results of the present study can be considered robust (15).

Strengths and limitations

An important strength of the present study is its pragmatic RCT design, which enabled the evaluation of the intervention's cost-effectiveness and financial return under "real world" circumstances. The randomization aspect is of particular importance, as financial return estimates based on RCTs seem to differ from those of non-randomized studies with a higher risk of bias (55). A second strength concerns the measurement and valuation of productivity-related costs. For estimating absenteeism costs, sickness absence and salary data were retrieved from company records, which eliminated recall bias. Also, the present study incorporated presenteeism costs, which can represent a considerable proportion of total productivity-related spending. Not including this cost category could lead to an underestimation of the cost impact of worksite interventions. It is important to mention, however, that a "gold standard" for measuring and valuing presenteeism does currently not exist. Further research is therefore needed to develop more sophisticated methods for estimating and valuing presenteeism and to reach consensus about the best way to do so. Until then, the methods used in the present study provides at least a crude estimate of the presenteeism costs associated with a mindfulness-based worksite intervention. A third strength concerns the use of state-of-the-art statistical methods that are not or infrequently used in occupational health research. Multiple imputation was used to avoid the problems of lost power and inefficiency associated with complete-case analyses. SUR analyses were used for the cost and effect components of the CEA allowing us to adjust for various confounders that are not required to be the same for costs and effects. Also, this method has the advantage that it allows for the correction for the possible correlations between error terms of regression equations (i.e. cost and effect equations) (39). Finally, until now, BCR and ROI estimates were presented without an indication of their uncertainty. In the present study, bootstrapping techniques, which are frequently used for estimating the uncertainty surrounding skewed data (34;52), were used to estimate their 95%CIs.

Some methodological limitations deserve attention as well. A first limitation concerns the amount of missing data; i.e. complete data were missing from 32% of participants. Although this percentage is relatively low compared to that of similar studies (56;57), and multiple imputation was used for handling missing data, multiple imputation cannot be regarded as a solution to prevent missing data in the first place. Even with the most sophisticated imputation techniques, cost and effect estimates are less reliable than those based on a 100% complete dataset. Every endeavor should therefore be made in future studies to minimize the amount of missing data. Second, the present study was carried out at two Dutch governmental research institutes and it is unknown whether the results are generalizable to other working populations or jurisdictions (i.e. external validity). As for the jurisdictions, it is important to bear in mind that employee medical cost policies differ between countries. In countries with employer-provided health insurance (e.g. the United States), employers bear a large part of the medical costs of their employees, whereas in the Netherlands (which has a dual payer system) they accrue to the government or health insurance companies. Therefore, the results of the analyses performed from the employer's perspective are mainly of interest to decision makers in countries with comparable policies. Third, as work engagement seems to be a predictor for turnover intentions (8;9), we set out to include turnover costs in our cost estimate. Unfortunately, the participating institutes were not able to provide us with all the required information and turnover costs were therefore only included in a sensitivity analysis. Nevertheless, the current methods used for estimating them could be insightful for researchers that intend to include this cost category in a future economic evaluation. Also, some researchers argue that economic evaluations should not be performed of interventions that were previously found to be ineffective. Such interventions, however, may still be considered cost-effectiveness and/or show a positive financial return if there is a non-negligible probability that the intervention is associated with cost-savings (52;58). Even if this would not be the case, not reporting on their cost-effectiveness and/or financial return would lead to biased systematic reviews on the resource implications of interventions.

Conclusion

The intervention was neither cost-effective from the societal perspective nor from that of the employer. Also, the intervention did not produce a positive financial return to the employer. Therefore, although mindfulness-based interventions are increasingly being offered in the workplace, the present study did not provide evidence for its implementation on cost-related grounds.

Acknowledgements

This project is part of a research program called “Vitality In Practice”, which is funded by Fonds Nuts Ohra (Nuts Ohra Foundation). The authors wish to thank Rosan Oostveen and Henrike van der Does for their help with the data collection. The authors would also like to thank all participants and trainers.

REFERENCES

1. World Health Organization. *Mental health and well-being at the workplace - protection and inclusion in challenging times*. Copenhagen; 2010.
2. Schaufeli WB, Bakker AB, Van Rhenen W. How changes in job demands and resources predict burnout, work engagement, and sickness absenteeism. *J Organiz Behav* 2009;30(7):893-917.
3. Chisholm D, Flisher AJ, Lund C, Patel V, Saxena S, Thornicroft G, et al. Scale up services for mental disorders: a call for action. Lancet 2007;370(9594):1241-1252.
4. Andlin-Sobocki P, Jönsson B, Wittchen HU, Olesen J. Cost of disorders of the brain in Europe. *Eur J Neurol* 2005;12:1-27.
5. McDaid D, Park AL. Investing in mental health and well-being: findings from the DataPrev project. *Health Promot Int* 2011;26 Suppl 1:i108-i139.
6. Schaufeli WB, Bakker AB. *Utrecht Work Engagement Scale*. Utrecht, the Netherlands: Occupational Health Psychology, Unit Utrecht University; 2003.
7. Schaufeli WB, Taris TW, Van Rhenen W. Workaholism, Burnout, and Work Engagement: Three of a Kind or Three Different Kinds of Employee Well-being? *Appl Psychol* 2008;57(2):173-203.
8. Alarcon GM, Lyons JB. The Relationship of Engagement and Job Satisfaction in Working Samples. *J Psychol* 2011;145(5):463-480.
9. Alarcon GM, Edwards JM, Menke LE. Student Burnout and Engagement: A Test of the Conservation of Resources Theory. *J Psychol* 2011;145(3):211-227.
10. Airila A, Hakanen J, Punakallio A, Lusa S, Luukkonen R. Is work engagement related to work ability beyond working conditions and lifestyle factors? *Int Arch Occup Environ Health* 2012;85(8):915-925.
11. Baer RA. Mindfulness based treatment approaches: Clinician's guide to evidence base and applications.: Elsevier Academic Press: Burlington, MA, USA, 2006.
12. Bond FW, Flaxman PE. The Ability of Psychological Flexibility and Job Control to Predict Learning, Job Performance, and Mental Health. *Journal of Organizational Behavior Management* 2006;26(1-2):113-30.
13. van Berkel J, Boot CRL, Proper KI, Bongers PM, van der Beek AJ. Effectiveness of a worksite mindfulness-related multi-component health promotion intervention on work engagement and mental health: results of a randomized controlled trial. *Plos One* 2013
15. Giffin SC. 14. van Berkel J, Boot CRL, Proper KI, Bongers PM, van der Beek AJ. Effectiveness of a worksite mindfulness-related multi-component intervention on lifestyle behaviours. Submitted Dealing with uncertainty in the economic evaluation of health care technologies. Thesis (PhD), University of York, 2010.
16. Leigh JP. Expanding research on the economics of occupational health. *Scand J Work Environ Health* 2006;32(1):1-4.
17. Drummond MF, Sculpher M.J., Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.
18. Cavallo D. Using return on investment analysis to evaluate health promotion programs: challenges and opportunities. *Health Promotion Economics Issue Briefs* 2006;1(3):1-4. RTI-UNC Center of Excellence. Available at: http://www.rti.org/pubs/IssueBrief_3.pdf.

5

19. Stone PW. Return-on-investment models*. Appl Nurs Res* 2005;18(3):186-189.

20. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011;12(12):1031-1049.

21. van Berkel J, Proper KI, Boot CRL, Bongers PM, van der Beek AJ. Mindful "Vitality in Practice": an intervention to improve the work engagement and energy balance among workers; the development and design of the randomised controlled trial. *BMC Public Health* 2011;11(1):736.

22. van der Zee KI, Sanderman R. *Het Meten van Gezondheidstoestand Met de RAND-36: een Handleiding*. Groningen, the Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 1993

23. Koppes LLJ, de Vroome EMM, Mol MEM, Janssen BJM, van den Bossche SNJ. *Nationale enquete arbeidsomstandigheden 2010: Methodologie en globale resultaten. 2010.* http://www.tno.nl/downloads/rapport_nea_20111.pdf

24. Tuomi K, Ilmarinen A, Jahkola A, Katajarinne L, Tulkki A. *Work Ability Index*. Helsinki: Finish Institute of Occupational Health; 1998.

25. Ahlstrom L, Grimby-Ekman A, Hagberg M, Dellve L. The work ability index and single-item question: associations with sick leave, symptoms, and health--a prospective study of women on long-term sick leave. *Scand J Work Environ Health* 2010;36(5):404-412.

26. Torgén M. Experiences of WAI in a random sample of the Swedish working population. International Congress Series 2005;1280(0):328-332.

27. Frick FD. Microcosting Quantity Data Collection Methods. Med Care 2009;47(7 Suppl 1):S76-S81.

28. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding Voor Kostenonderzoek. Methoden en Standaardkostprijzen *Voor Economische Evaluaties in de Gezondheidszorg*. Diemen, the Netherlands: College Voor Zorgverzekeringen; 2010.

29. Z-index. G-Standard. The Hague, The Netherlands: Z-Index BV; 2009.

30. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;14(2):171-189.

31. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE. Using the world health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46:S23-S37.

32. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med* 2003;45:156-174.

33. Statistics Netherlands. 2011. http://www.cbs.nl

34. Drummond MF, Sculpher M.J., Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.

35. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2010;30(4):377-399.

36. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;16(3):219-242.

37. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work*? Int J Methods Psychiatr Res* 2011;20(1):40-49.

38. Rubin DB. Multiple imputation for nonresponse in surveys. John Wiley & Sons: New York, 1987.

38. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;13(5):461-475.

40. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10(3):212-4.

41. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. *Health Econ* 2004;13(5):405-415.

42. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5(1):23-28.

43. Koopmanschap M, Meeding WJ, Evers S, Severens J, Burdorf A, Brouwer W. *Handleiding voor het gebruik van PRODISQ versie 2.1.* [Handbook on use of PRODISQ.] Rotterdam/Maastricht, Erasmus MC - Instituut voor Medical Technology Assessment, Instituut Maatschappelijke Gezondheidszorg, Universiteit van Maastricht - Beleid Economie en Organisatie van de Zorg; 2004.

44. Smith HL, Watkins LE. Managing manpower turnover costs. *Pers Adm* 1978;23(4):46-50.

45. van Berkel J, Boot CRL, Proper KI, Bongers PM, van der Beek AJ. Process evaluation of a workplace health promotion intervention aimed at improving work engagement and energy balance. *J Occup Environ Med* 2013:55(1):19-26.

46. Adams RL. Examining the effects of Mindfulness-Based Stress Reduction (MBSR) Training on Working Adults. Thesis (PhD), Institute of transpersonal psychology, 2011.

47. Allexandre D, Neuman A, Hunter J, Morledge T, Roizen M. P02.63. Efficacy of an 8-week online mindfulness stress management program in a corporate call center. *BMC Complementary and Alternative Medicine* 2012;12(Suppl 1):P119.

48. Bazarko DM. Mindfulness Meditation in a Corporate Setting. Proceedings of the 22nd international nursing research congress; 2011 Jul 11-14; Cancun, Mexico.

49. Gold E, Smith A, Hopper I, Herne D, Tansey G, Hulland C. Mindfulness-Based Stress Reduction (MBSR) for Primary School Teachers. *J Child Fam Stud* 2010;19(2):184-189.

50. Wolever DQ, McCabe K, Fekete E, Bobinet KJ, Mackenzie ER, Kusnick CA. Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. *J Occup Health Psychol* 2012;17(2):246-258.

51. Hofmann SG, Sawyer AT, Witt AA, Oh D. The Effect of Mindfulness-Based Therapy on Anxiety and Depression: A Meta-Analytic Review. *J Consult Clin Psychol* 2010;78(2):169-183.

52. Petrou S, Gray A Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ* 2011;342.

53. Salanova M, Agut S, Peiro JM. Linking organizational resources and work engagement to employee performance and customer loyalty: the mediation of service climate. *J Appl Psychol* 2005;90(6):1217-27.

54. van Dongen JM, Strijk JE, Proper KI, van Wier MF, van Mechelen W, van Tulder MW, et al. A cost-effectiveness and return-on-investment analysis of a worksite vitality intervention among older hospital workers: results of a randomized controlled trial. *J Occup Environ Med* 2013;55(3):337-346.

55. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, et al. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health* 2012;38(5):393-408.

56. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, Van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the contruction industry at risk for cardiovascular disease. *J Occup Environ Med* 2011;53[6]:610-617

57. Robroek SJW, Polinder S, Bredt FJ, Burdorf A. Cost-effectiveness of a long-term Internet-delivered worksite health promotion programme on physical activity and nutrition: a cluster randomized controlled trial. Health Educ Res 2012;27(3):399-410.

58. Oude Hengel OH, Bosmans JE, van Dongen JM, Bongers PM, van der Beek AJ, Blatter BM. Prevention program at construction worksites aimed at improving health and work ability is cost saving to the employer: results from a RCT. *Am J Ind Med* 2013.

Appendix 1: Turnover cost calculation

Cost categorie	Staff	Units	Unit prices	Total costs (Euro 2011)	Costs per worker (2011)
Separation costs					
Exit interviews	Departing employee	1 hour	€ 49.02/hour[a]	€ 89.07	€ 89.07
	Capital	1 hour	€ 1.56/hour[b]	€ 1.56	€ 1.56
	Senior employee	1 hour	€ 46.60/hour[c]	€ 49.02	€ 49.02
	Materials (Exit form)	1	€ 1.50/form	€ 1.50	€ 1.50
				Subtotal	**€ 141.15**
Administrative functions relating to termination	Employee human resources	0.25 hour	€ 46.60/hour[c]	€ 11.65	€ 11.65
	Secretary (department level)	1.25 hours	€ 30.74/hour[d]	€ 38.43	€ 38.43
	Capital	1.50 hours	€ 1.56/hour[b]	€ 2.34	€ 2.34
				Subtotal	**€ 52.41**
Severance pay	N.A.	N.A.	N.A.	N.A.	N.A.
				Subtotal	**€ 0.00**
Total separation costs					**€ 193.56**
Replacement costs					
Pre-employment administration	Employee human resources	0.50 hours	€ 46.60/hour[c]	€ 23.30	€ 23.30
Arranging job interviews,	Senior employee	4.25 hours	€ 59.70/hour[f]	€ 253.72	€ 253.72
reference checks,	Secretary (department level)	4.00 hours	€ 30.74/hour[d]	€ 122.96	€ 122.96
drafting job description etc.	Manager	0.5 hour	€ 89.07/hour[e]	€ 44.53	€ 44.53
	Capital	8.75 hours	€ 1.56/hour[b]	€ 13.56	€ 13.56
	Capital	0.5 hour	€ 3.12/hour[f]	€ 1.56	€ 1.56
				Subtotal	**€ 459.72**
Job postings	Budget per job posting	1	€ 603.18	€ 603.18	€ 603.18
				Subtotal	**€ 603.18**
Job interviews	*Round 1 (5 interviews)*				
	Senior employee	5 hours	€ 59.70/hour[g]	€ 298.49	€ 298.49
	Senior employee	5 hours	€ 59.70/hour[g]	€ 298.49	€ 298.49
	Capital	5 hours	€ 3.12/hour[f]	€ 15.60	€ 15.60

Cost categorie	Staff	Units	Unit prices	Total costs (Euro 2011)	Costs per worker (2011)
	Travelling expenses applicants	300 km	€ 0.21/km[h]	€ 63.00	€ 63.00
	Round 1 (2 interviews)				
	Manager	2 hours	€ 89.07/hour[e]	€ 178.14	€ 178.14
	Senior employee	2 hours	€ 59.70/hour[g]	€ 119.40	€ 119.40
	Capital	2 hours	€ 3.12/hour[f]	€ 6.24	€ 6.24
	Travelling expenses applicants	120 km	€ 0.21/km[h]	€ 25.20	€ 25.20
				Subtotal	**€ 1,004.55**
Costs of hiring decision meeting	Manager	0.5 hour	€ 89.07/hour[e]	€ 44.53	€ 44.53
	Senior employee	0.5 hour	€ 59.70/hour[g]	€ 29.85	€ 29.85
	Senior employee	0.5 hour	€ 59.70/hour[g]	€ 29.85	€ 29.85
	Capital	0.5 hour	€ 3.12/hour[f]	€ 1.56	€ 1.56
				Subtotal	**€ 105.79**
Costs of employment meeting	Employee human resources	1 hour	€ 46.60/hour[c]	€ 46.60	€ 29.85
	New employee	1 hour	€ 49.02/hour[a]	€ 49.02	€ 29.85
	Capital	1 hour	€ 3.12/hour[f]	€ 3.12	€ 1.56
				Subtotal	**€ 98.74**
Post-employment information gathering	Employee human resources	0.5 hours	€ 46.60/hour[c]	€ 23.30	€ 23.30
	Secretary (department level)	1.75 hours	€ 30.74/hour[d]	€ 53.80	€ 53.80
	Capital	2.25 hours	€ 1.56/hour[b]	€ 3.51	€ 3.51
				Subtotal	**€ 80.60**
Applicant testing	N.A.	N.A.	N.A.	N.A.	N.A.
				Subtotal	**€ 0.00**
Total replacement costs					**€ 2,352.59**
Training costs					
Information literature	Booklet	1 booklet	€ 10.00/booklet	€ 10.00	€ 10.00
				Subtotal	**€ 10.00**

Cost categorie	Staff	Units	Unit prices	Total costs (Euro 2011)	Costs per worker (2011)
Preparation general orientation	Employee human resource department	2 hours	€ 46.60/hour[c]	€ 93.20	€ 2.49
	Capital	2 hours	€ 1.56/hour[e]	€ 3.12	€ 0.08
				Subtotal	**€ 2.57**
General orientation tour	Employee human resource department	7.5 hours	€ 46.60/hour[c]	€ 349.48	€ 9.32[i]
	New employee	7.5 hours	€ 49.02/hour[a]	€ 367.65	€ 367.65
	Senior employees	5.5 hours	€ 59.70/hour[g]	€ 328.43	€ 8.76[i]
	Capital	7.5 hours	€ 10.40/hour[j]	€ 78.00	€ 2.08[i]
	Drinks and snacks		€ 12.37/participant[k]	€ 12.37	€ 12.37
				Subtotal	**€ 397.61**
Costs of job orientation	Senior employee	2 hours	€ 59.70/hour[g]	€ 119.40	€ 119.40
	Secretary (department level)	1.5 hours	€ 30.74/hour[d]	€ 46.11	€ 46.11
	New employee	7.5 hours	€ 49.02/hour[a]	€ 367.65	€ 367.65
	Capital	11 hours	€ 1.56/hour[b]	€ 43.42	€ 43.42
				Subtotal	**€ 567.48**
Total training costs					**€ 977.65**
Total costs per participant turning over					**€ 3523.80**

Abbreviations: N.A.: Not Applicable
[a] Valued using the average gross annual salary including holiday allowances and premiums of participants, [b] Valued assuming a surface area of $6m^2$, [c] Valued using the average gross annual salary including holiday allowances and premiums of human resource department employees, [d] Valued using the average gross annual salary including holiday allowances and premiums of administrative employees, [e] Valued using the average gross annual salary including holiday allowances and premiums of management employees, [f] Valued assuming a surface area of $12m^2$, [g] Valued using the average gross annual salary including holiday allowances and premiums of senior employees, [h] Valued in accordance with the Dutch manual of costing, [l] Divided by the average number of new employees participating per general orientation tour (n=37,5), [j] Assuming a surface area of $40m^2$, [k] Values using invoices

Appendix 2: Cost of the Mindful VIP intervention from the societal perspective, valued using a bottom-up micro-costing approach (Euros 2011)

	Cost categories	Staff	Units	Unit prices	Total costs (Euro 2011) (n=129)	Costs per worker (Euro 2011)
Mindfulness training	Labor costs	Mindfulness trainer	112 sessions (226.33 hours)	€ 41.66/hour	€ 9,429.66	€ 73.10
	Capital costs			€ 8.32/hour	€ 1.883.09	€ 14.60
	Travelling costs		4872 km	€ 0.21/km	€ 1007.80	€ 7.81
	Material costs		129 Hand-outs	€ 1.80/hand-out	€ 232.20	€ 1.80
			129 Exercise booklet + CD	€ 10.60/set	€ 1,367.40	€ 10.60
			VIP Webpage			
					Subtotal:	**€ 112.49**
Inter- and supervision mindfulness coaches	Labor costs	4 Mindfulness trainers	7 hours	€ 166.65/hour	€ 1,166.56	€ 9.04
	Capital costs			€ 8.32/hour	€ 58.24	€ 0.45
					Subtotal:	**€ 9.49**
E-coaching	Labor costs	Mindfulness trainer	194 sessions (64.67 hours)	€ 41.66/hour	€ 2.694.19	€ 20.89
	Capital costs			N.A.[a]		
					Subtotal:	**€ 20.89**
"VIP Webpage" hosting and maintenance	Labor costs	Computer specialist	16 hours	€ 35.45/hour	€ 567.27	€ 4.40
	Capital costs			€ 1.56/hour	€ 24.96	€ 0.19
					Subtotal:	€ **4.59**
Fruit and vegetables			157 fruit/vegetable boxes	€ 17.65/box	€ 2,770.57	€ 21.48
					Subtotal:	**€ 21.48**
Development intervention					€114,633.25	€ 6.53[b]
					Subtotal:	**€ 6.53**

Total intervention costs	**€ 170.89**

Abbreviations: N.A.: Not Applicable
[a] Capital costs are not applicable for the e-coaching trajectory as the mindfulness trainers provided these sessions from their own home.
[b] €114,633.25 was paid for the development of the Mindfulness VIP intervention. For calculating the development costs per participant, these were divided by the expected number of program users during the first five years after implementation (17,557). Knowledge workers make up 9.5% (702,240 workers) of the total Dutch workforce (http://www.cbs.nl). During the five year period, it was hypothesized that the intervention will be offered to 10% of Dutch knowledge workers and that 25% of them will participate to the intervention (15% during the first year after implementation and 2.5% during each consecutive year).

6

Cost-effectiveness and return-on-investment of a worksite intervention aimed at improving physical activity and nutrition among construction workers

Johanna M van Dongen
Laura Viester
Marieke F van Wier
Judith E Bosmans
Evert ALM Verhagen
Maurits W van Tulder
Paulien M Bongers
Allard J van der Beek

To be submitted

ABSTRACT

Objectives: To conduct a cost-effectiveness and return-on-investment (ROI) analysis of a worksite physical activity and nutrition program for construction workers in comparison with usual practice.

Methods: The intervention consisted of generic as well as tailored health information and personal health counseling. A total of 314 participants were randomized to the intervention (n=162) or control group (n=152). Data on body weight, waist circumference, musculoskeletal disorders (MSD), work-related vitality, and job satisfaction were collected at baseline, 6, and 12 months. Sickness absence data were collected from company records. Other cost data were collected with 3-monthly questionnaires. Missing data were imputed using multiple imputation. Cost-effectiveness analyses were conducted from both the societal and employer's perspective. A ROI analysis was performed from the employer's perspective. Bootstrapping techniques were used to assess the uncertainty of the results.

Results: Intervention costs per participant were €178 from the societal perspective (bottom-up micro-costed) and €287 from that of the employer (market prices). At 12-month follow-up, no statistically significant cost and effect differences were found. The probabilities of cost-effectiveness for body weight, waist circumference, and MSD gradually increased with an increasing ceiling ratio to 0.84 (willingness-to-pay = €21,000/kg), 0.77 (willingness-to-pay = €18,000/cm), and 0.84 (willingness-to-pay = €42,000/person prevented from having a MSD), respectively. The probabilities of cost-effectiveness for work-related vitality and job satisfaction were low at all ceiling ratios (≤0.54). Financial return estimates were positive, but their confidence intervals were rather wide and none of them was statistically significant.

Conclusion: The intervention's cost-effectiveness in improving weight-related outcomes and MSD depends on the societal and employer's willingness-to-pay for these effects and the probability of cost-effectiveness that they consider acceptable. From the employer´s perspective, the intervention was not cost-effective in improving work-related vitality and job satisfaction. Also, due to a high level of uncertainty, it cannot be concluded that the intervention was cost-beneficial to the employer.

INTRODUCTION

Excessive body weight and musculoskeletal disorders (MSD) have a serious impact on public health in many developed countries (1-5). In the Netherlands, the combined prevalence of overweight (Body Mass Index [BMI] 25 - 30 kg/m^2) and obesity (BMI ≥ 30 kg/m^2) is 48% among adults (6), and that of MSD is estimated to be 39% in adult men and 45% in adult women (7). Among construction workers, these prevalences are even higher (8;9). Both conditions not only reduce a person's well-being, but also impose a large economic burden on companies and society as a whole due to increased absenteeism, presenteeism (i.e. reduced productivity while at work), and healthcare consumption (10-12).

The workplace presents a useful setting to combat the high prevalence of excessive body weight and MSD, as it provides social and organizational support structures that can help improve risk behaviours and many companies have the infrastructure available to offer behaviour change interventions at relatively low costs (13). In addition, worksite physical activity and nutrition programs in particular, cannot only reduce body weight (14) and MSD prevalence (15), but may also generate cost savings to a company through reduced absenteeism (16) and presenteeism (17). Therefore, in the VIP in Construction study, a worksite physical activity and nutrition program was developed aimed at preventing and reducing overweight and MSD among construction workers (i.e. VIP in Construction intervention) (18). An evaluation of the intervention's effectiveness has been reported elsewhere (19;20).

Decisions about investments in worksite health promotion programs typically lie by the company management. In doing so, they are not just interested in the effectiveness of such interventions, but also in their impact on the company's bottom-line (21;22). To provide this information, return-on-investment (ROI) analyses can be performed in which the costs of an intervention are compared to the company's resulting financial savings (23;24). However, as health outcomes are not directly considered in a ROI analysis and other stakeholders may reap a large part of the benefits (e.g. health insurance companies), cost-effectiveness analyses (CEAs) and analyses from the broader societal perspective are of importance as well.

The present study aimed to conduct CEAs and a ROI analysis, in which the VIP in Construction intervention was compared to usual practice. CEAs were performed from both the societal and employer's perspective, and the ROI analysis from that of the employer.

METHODS

Study design

Analyses were conducted alongside a 12-month randomized controlled trial (RCT), which took place from 2010 to 2012. The study protocol was approved by the Medical Ethics Committee of the VU University Medical Center (18), and the trial has been registered in the Netherlands Trial Register (NTR2095).

Participants

All blue collar workers of a Dutch construction company who were invited for a voluntary periodical health screening at the occupational health service between February 2010 and October 2011 were recruited for the study. Workers who were on long-term sick leave (≥4 weeks) were excluded. At baseline, all workers who decided to participate in the study provided informed consent. After baseline measurements, participants were randomized to the intervention or control group. Randomization took place at the individual level and was performed by a research assistant using a computer-generated randomization sequence in SPSS (v15, Chicago, IL). The research assistant had no information on the participants to ensure allocation concealment (18).

Intervention and control condition

All participants received practice as usual. Additionally, intervention group participants received the VIP in Construction intervention. A detailed description of the intervention has been given elsewhere (18). In brief, the intervention consisted of generic as well as tailored health information (i.e. VIP in Construction toolbox) and personal health counseling (PHC). Participants with a healthy weight status (i.e. BMI<25 and waist circumference<94) and a healthy physical activity level (i.e. meeting physical activity recommendations (25;26)) only received the VIP in Construction toolbox; all others also received PHC.

The VIP in Construction toolbox consisted of health information brochures tailored to the participants' physical activity level and weight status, a calorie guide, a pedometer, a BMI card, a waist circumference measuring tape, a cookbook including healthy recipes and a knowledge test, "personal energy plan" forms, an overview of the health promotion facilities of the company, and an exercise card.

PHC intensity (i.e. number and duration of contacts) was tailored to the participants' stage-of-change for improving physical activity and nutrition (Table 1) (18;27). Face-to-face and telephone coaching contacts were provided during work hours and were given by physiotherapists specialized in lifestyle coaching (i.e. health coaches). Face-to-face coaching contacts took place at the worksite. A web-based system was used to register the participants' coaching contacts (i.e. date, time), as well as their content (i.e. goals, action plans).

Table 1: Personal health coaching (PHC) contact schedule

Stage-of-change(27)	PHC-group	2 weeks	1 month	2 months	3 months	4 months
Pre-contemplation stage *The participant does not intend to change his risk behavior(s)*	A	Intake (60 min face-to-face)	Follow-up 1: (30 min; telephone)	Follow-up 2: (15 min; telephone)		Follow-up 3: (15 min; telephone)
Contemplation/ Preparation stage *The participant wants to change his risk behavior(s), but does not know how*	B	Intake (60 min face-to-face)		Follow-up 1: (30 min; telephone)	Follow-up 2 (15 min; telephone)	
Action stage *The participant already started changing his risk behavior(s)*	C	Intake (30 min face-to-face)			Follow-up 1 (10 min telephone)	

Abbreviations: min: minutes

Effect measures

Primary and secondary outcomes were assessed at baseline, six, and 12 months.

Primary outcomes

Primary outcomes were body weight and waist circumference. Body weight was measured using a calibrated scale with participants wearing light clothes and no shoes. Waist circumference was measured midway between the lower rib margin and the iliac crest, and was rounded to the nearest 0.1cm. Measurements were performed in a standing position, over bare skin, and at the end of expiration (28). At baseline, these measurements were performed by occupational physicians or their assistants. At 6 and 12 months, they were performed by the research team.

Secondary outcomes

Secondary outcomes were MSD, work-related vitality, and job satisfaction. The prevalence of MSD was assessed using the “Dutch Musculoskeletal Questionnaire” (DMQ) (29). Participants were asked to rate the occurrence of pain or discomfort in the neck, shoulders, upper and lower back, elbows, wrists/hands, knees, and ankles/feet during the previous three months on a 4-point scale (never, sometimes, frequent, and prolonged). Participants who answered “frequent” or “prolonged” on one or more of the questions were classified as having MSD; all others as not having MSD. Work-related vitality was assessed using a subscale of the “Utrecht Work Engagement Scale” (i.e. UWES Vitality Scale). This scale included six items, scored on a 7-point scale ranging from “never”(0) to “always”(6). The UWES Vitality Score ranged from 0-6 (higher scores indicate a better work-related vitality) (30). Job satisfaction was assessed using a 1-item question of the “Netherlands Working Conditions Survey” (31). Participants were asked to rate their overall job satisfaction on a 5-point scale ranging from “very dissatisfied”(1) to “very satisfied”(5).

Resource use and valuation

Intervention costs

For the societal perspective, bottom-up micro-costing was used to quantify intervention costs (32). Intervention costs included those related to the development, implementation, and operation of the intervention. Frequency, duration, preparation time, and locations of coaching contacts were recorded by the coaches. Labor costs were valued by multiplying the intervention staff's time investments (hours) by their gross hourly salaries including overhead costs. Capital costs were valued using cost data collected from finance department staff. Material costs were estimated using invoices. Coaches' travelling costs were valued according to the Dutch manual of costing (33). As PHC contacts took place during work hours, the participants' lost productivity costs for the duration of the contacts were included as well, and were valued using the average salary (including overhead costs) of Dutch construction workers (Economic Institute of the Dutch construction industry, personal communication).

For the employer's perspective, intervention costs were valued using charges paid. Lost productivity due to PHC was valued using the average salary (including overhead costs) of blue collar workers of the participating company.

Healthcare costs

Healthcare utilization was assessed using 3-monthly retrospective questionnaires and included costs of primary healthcare (i.e. general practitioner, allied health professionals, complementary medicine), secondary healthcare (i.e. medical specialist, hospitalization), and both prescribed and over-the-counter medications. Dutch standard costs were used to value primary and secondary healthcare utilization (33). If unavailable, prices according to professional organizations were used. Medication use was valued using unit prices of the Royal Dutch Society of Pharmacy (34).

Occupational health costs

Occupational health costs consisted of gym membership subsidies, as provided by the employer. The duration of the memberships was assessed using 3-monthly retrospective questionnaires. The associated costs were calculated by multiplying the duration of the memberships (in months) by the height of the subsidy (i.e. €10/month).

Sports costs

Sports costs were assessed using 3-monthly retrospective questionnaires asking participants to report their sports membership fees and expenses on sports equipment during the previous three months.

Absenteeism costs

Baseline (i.e. one year prior to baseline) and follow-up sickness absence data were collected from company records. For the societal perspective, costs per sickness absence day were calculated by dividing the average annual salary of Dutch construction workers (including overhead costs) by the associated number of workable days (i.e. 214) (33). Absenteeism costs were estimated using the "Friction Cost Approach"(FCA) (35). A friction period of 23 weeks (i.e. period needed to replace a sick worker) and an elasticity of 0.8 (i.e. a 100% reduction in work time corresponds with an 80% reduction in productivity) were assumed (33;35). For the employer's perspective, costs per sickness absence day were calculated using the average annual salary of blue collar workers of the participating company (including overhead costs). Subsequently, absenteeism costs were estimated using the "Human Capital Approach"(HCA), in which absenteeism costs are neither truncated as in the FCA, nor is elasticity considered (33).

Presenteeism costs

Presenteeism was assessed on a 3-monthly basis using an item of "The World Health Organization Health and Work Performance Questionnaire"(WHO-HPQ) (36;37). In the WHO-HPQ, presenteeism is conceptualized as a measure of actual work performance in relation to "best performance", irrespective of the presence or absence of health complaints (37). Participants were asked to rate their overall work performance during the previous three months on an 11-point scale ranging from "worst performance"(0) to "best performance"(10). Their average work performance during follow-up (*Wown*) was estimated and the participants' level of presenteeism (P_{HPQ}) was calculated using the following formula:

$$P_{HPQ} = (10 - Wown)/10$$

Presenteeism days were calculated by multiplying the participants' P_{HPQ} by their number of days worked during follow-up; i.e. working days minus sickness absence days. Presenteeism days were valued using the average salary of Dutch construction workers (societal perspective) and that of blue collar workers of the participating company (employer's perspective).

Using consumer price indices, all costs were converted to 2011 Euros (38). Discounting of costs and effects was not necessary, because the follow-up of the trial was one year (39). Price weights used for valuing resource use are given in Appendix 1.

Data analysis

Analyses were performed according to the intention-to-treat method. Descriptive statistics were used to compare baseline characteristics between intervention and control group participants, and participants with complete and incomplete data. Missing data were imputed in IBM SPSS (v20, Chicago, IL) using Fully Conditional Specification and Predictive Mean Matching. An imputation model was constructed that included variables related to the "missingness" of data and those that predicted the outcome variables. The model included age, smoking status, baseline sickness absence, baseline effect measure values, and available midpoint and follow-up cost and effect measure values (6- and 12 months). Fifteen different data sets were created (Loss of Efficiency≤5%) (40). Each data set was analyzed separately as specified below. Pooled estimates were subsequently calculated using Rubin's rules (41). Data were imputed at the cost level. Therefore, a descriptive analysis of resource use was performed using the complete-cases only. T-tests were used for continuous variables and Chi-square tests for dichotomous variables. For skewed data, uncertainty was assessed using the bias-corrected accelerated (BCA) bootstrap method (5000 replications). Unless otherwise stated, data were analyzed in STATA (V12, Stata Corp, College Station, TX), with a level of significance of $p<0.05$.

Cost-effectiveness analysis

CEAs in terms of body weight and waist circumference were conducted from the societal perspective (i.e. all costs were taken into consideration regardless of who pays or benefits). CEAs in terms of work-related vitality, job satisfaction, and MSD were conducted from the employer's perspective (i.e. only the costs borne by employers were considered). Linear regression analyses were used to compare outcomes between the intervention and control group. Follow-up outcomes were adjusted for their baseline values. To compare costs between both groups, 95% confidence intervals (95%CIs) around the unadjusted mean differences in total and disaggregated costs were calculated using BCA bootstrapping (5000 replications). Seemingly unrelated regression (SUR) analyses were performed, in which effect differences were corrected for their baseline values and cost differences for baseline sickness absence and presenteeism scores (42). Incremental cost-effectiveness ratios (ICERs) were calculated by dividing the corrected cost differences by those in effects. Uncertainty was graphically illustrated by plotting bootstrapped incremental cost-effect pairs (CE-pairs) on cost-effectiveness planes (CE-planes) (43). A summary measure of the joint uncertainty of costs and effects was provided using cost-effectiveness acceptability curves (CEACs), which provide an indication of the intervention's probability of cost-effectiveness at different ceiling ratios (i.e. the maximum amount of money decision-makers are willing to pay per unit of effect) (44).

Return-on-investment analysis

The ROI analysis was performed from the employer's perspective, in which only employer costs and benefits were considered. Costs were defined as intervention costs. Benefits were defined as the difference in total monetized outcome measures (i.e. absenteeism, presenteeism, and occupational health costs) between the intervention and control group during follow-up, with positive benefits indicating reduced spending. The ROI analysis (costs and benefits) was conducted using SUR analyses, in which benefits were adjusted for baseline sickness absence and presenteeism scores. Three ROI metrics were calculated; 1) Net Benefits (NB), 2) Benefit Cost Ratio (BCR), and 3) Return On Investment (ROI) (23;24;45).

NB = Benefits – Costs
BCR = Benefits / Costs
*ROI = ((Benefits – Costs)/Costs)*100*

To quantify precision, 95% bootstrapped confidence intervals (5000 replications) were estimated around the benefits and ROI metrics using the percentile method. Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0% (23;24;45).

Sensitivity analyses

Five sensitivity analyses were conducted to test the robustness of the results. First, analyses were performed using the complete-cases only (SA1). Second, analyses were performed in which intervention costs were estimated under the assumption that the intervention took place outside work hours (SA2). Thus, the costs of lost productivity due to PHC were excluded. Third, analyses were performed in which absenteeism costs were valued using the HCA for the societal perspective and the FCA for the employer's perspective (SA3). Fourth, analyses were performed in which presenteeism costs were estimated using a slightly modified version of the "PROductivity and DISease Questionnaire" (PRODISQ) (46;47). In this version of the PRODISQ, presenteeism was conceptualized as reduced work performance due to health complaints and was valued by considering both the quantity and quality of labor input (SA4). Fifth, as overall consensus about whether or not to include presenteeism costs in economic evaluations does currently not exist, analyses were performed in which presenteeism costs were excluded (SA5).

RESULTS

Participants

After randomization, 162 participants were allocated to the intervention group and 152 to the control group. At baseline, intervention group participants had approximately four more sickness absence days than their control group counterparts. Also, the prevalence of MSD was higher in the intervention group (55.6%) than in the control group (49.3%) (Table 2). After 12 months, 32 intervention group (19.7%) and 22 control group participants (14.5%) were lost to follow-up, among others, because they lost their job or lost interest in the study (Figure 1). Complete data were obtained from 62.4% of participants on the effect measures (n=196; 101 intervention group participants and 95 control group participants) and 40.5% on the cost measures (n=127; 62 intervention group participants and 65 control group participants). Some differences were observed between participants with complete and incomplete data in both the intervention and control group (Table 2).

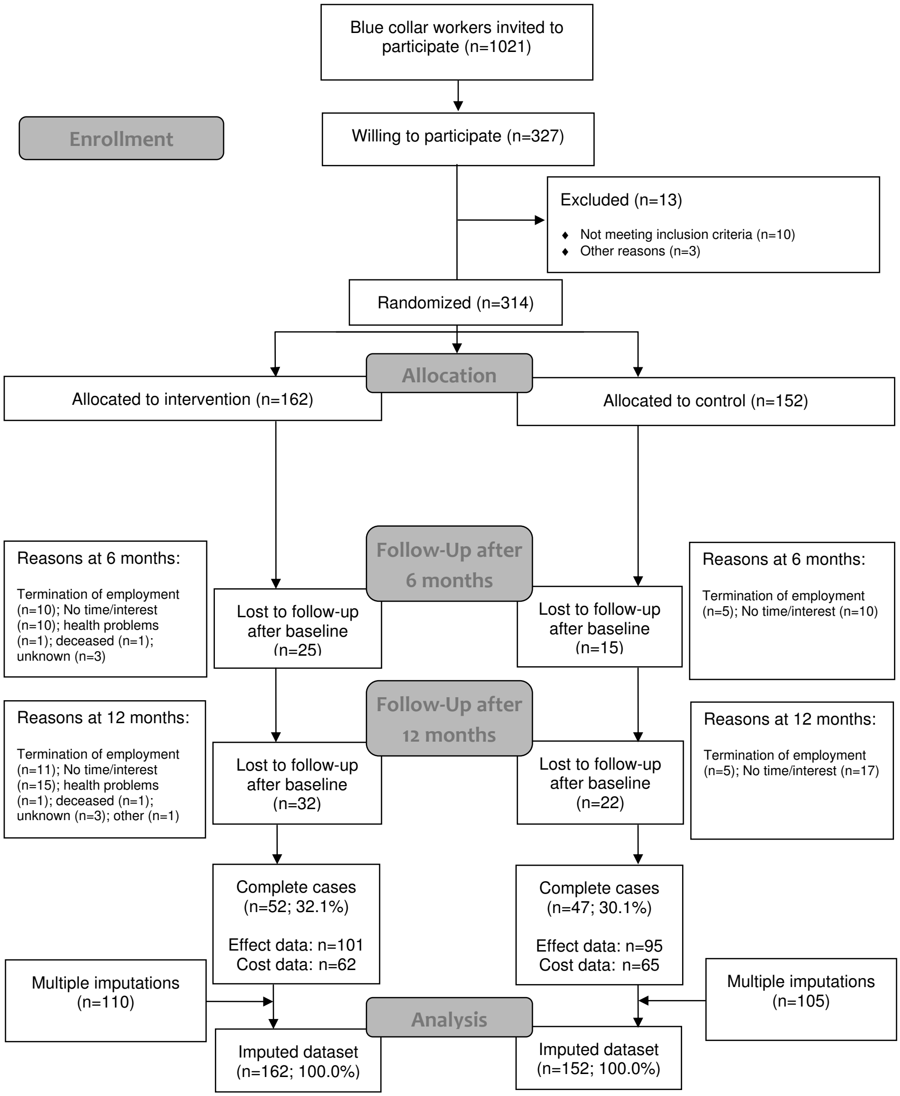

Figure 1: Flow chart of participants to the VIP in Construction study

Table 2: Baseline characteristics of the study population

	Intervention group			Control group		
Baseline characteristics	**All (n=162)**	**Complete (n=52)**	**Incomplete (n=110)**	**All (n=152)**	**Complete (n=47)**	**Incomplete (n=105)**
Male [n (%)]	162 (100); n=162	52 (100); n=52	110 (100); n=110	152 (100); n=152	47 (100); n=47	105 (100); n=105
Age (years) [mean (SD)]	46.3 (9.9); n=162	48.2 (9.2); n=52	45.3 (10.1); n=110	47.0 (9.5); n=151	47.5 (8.7); n=47	46.8 (9.9); n=104
Smokers [n (%)]	45 (27.8); n=155	12 (23.5); n=51	33 (31.7); n=104	44 (29.7); n=148	14 (31.1); n=45	30 (29.1); n=103
Body weight (kilograms) [mean (SD)]	88.7 (12.9); n=161	87.4 (11.8); n=52	89.3 (13.4); n=110	88.9 (14.4); n=152	89.9 (16.3); n=47	88.5 (13.5); n=105
Body Mass Index (kg/m^{-2}) [mean (SD)]	27.3 (3.5); n=161	27.2 (3.3); n=52	27.4 (3.6); n=109	27.4 (3.9); n=152	27.9 (4.4); n=47	27.2 (3.7); n=105
Waist circumference (centimetres) [mean (SD)]	99.0 (10.2); n=152	99.4 (10.1); n=52	98.9 (10.3); n=100	100.0 (11.8); n=133	100.3 (12.9); n=47	99.8 (11.2); n=86
Musculoskeletal disorders [n (%)]						
Yes	90 (55.6); n=162	30 (57.7); n=52	60 (54.5); n=110	75 (49.3); n=152	21 (44.7); n=47	54 (51.4); n=105
No	72 (44.4); n=162	11 (42.3); n=52	50 (45.5); n=110	77 (50.7); n=152	26 (55.3); n=47	51 (48.6); n=105
Work-related vitality (range: 0-6) [mean (SD)]	4.9 (1.0); n=157	5.0 (1.00); n=52	4.8 (1.1); n=105	5.0 (1.0); n=142	5.0 (1.0); n=47	5.0 (1.0); n=95
Job satisfaction (range: 1-5) [mean (SD)]	4.0 (0.7); n=157	4.0 (0.8); n=52	4.0 (0.7); n=105	3.9 (0.9); n=146	4.0 (0.9); n=47	3.9 (0.9); n=99
Sickness absence: number of sickness absence days during the year prior to baseline [mean (SD)]	14.0 (26.9); n=162	11.9 (24.7); n=52	15.0 (27.9); n=110	9.8 (20.6); n=152	11.1 (25.8); n=47	9.3 (17.8); n=105
Work performance: WHO-HPQ work performance score during a 4-week period prior to baseline [mean (SD)]	7.6 (1.1); n=154	7.7 (0.9); n=52	7.5 (1.2); n=102	7.9 (1.0); n=143	7.9 (1.0); n=47	7.9 (1.0); n=96

Abbreviations: n: number, SD: standard deviation, WHO-HPQ: World Health Organization Work Performance Questionnaire

Effectiveness

After 12 months, no statistically significant differences were found between the intervention and control group for body weight (-0.7; 95%CI: -2.2 to 0.7), waist circumference (-0.7; 95%CI: -2.5 to 1.1), MSD (-0.07; 95%CI -0.22 to 0.08), work-related vitality (-0.03; 95%CI: -0.39 to 0.33), and job satisfaction (-0.01; 95%CI: -0.34 to 0.32).

Resource use

Forty participants were allocated to PHC group A, 61 to PHC group B, 48 to PHC group C, and 13 only received the VIP in Construction toolbox (Table 1). During the intervention period, 126 face-to-face and 173 telephone counseling contacts were provided. Based on the complete-cases, intervention and control group participants did not significantly differ in terms of their average number of visits to a care provider (-2.4; 95%CI: -5.7 to 0.7), average number of days of hospitalization (-0.1; 95%CI: -0.4 to 0.2), average number of months of gym membership subsidies (0.5; 95%CI: -0.3 to 1.3), average number of sickness absence days (-2.7; 95%CI: -9.7 to 3.0), and average number of presenteeism days (-2.6; 95%CI: -9.6 to 4.1). However, significantly more intervention group participants (n=36) had sports costs than their control group counterparts (n=23; X^2: 5.3, p=0.02) (Appendix 1).

Costs

Average intervention costs per participant were € 178 (SD=77) from the societal perspective and €287 (SD=22) from the employer's perspective (Appendix 2). No statistically significant differences were found on all cost measures (Table 3).

Table 3: Mean costs per participant in the intervention and control group, and unadjusted mean cost differences between both groups during the 12-month follow-up period

Cost category	Intervention group n=162; mean (SEM)	Control group n=152; mean (SEM)	Mean cost difference (95%CI)
	Societal perspective		
Intervention costs	178 (6)	0 (0)	178 (166 to 190)
Medical costs	1499 (356)	1033 (174)	457 (-129 to 1434)
Occupational health costs	26 (4)	20 (3)	5 (-3 to 15)
Sports costs	461 (98)	265 (46)	156 (32 to 497)
Absenteeism costs	2214 (338)	2055 (345)	150 (-802 to 1094)
Presenteeism costs	9382 (550)	9663 (975)	-533 (-2449 to 1597)
Total	**13760 (725)**	**13037 (1025)**	**412 (-1572 to 3093)**
	Employer's perspective		
Intervention costs	287 (2)	0 (0)	287 (283 to 290)
Occupational health costs	26 (4)	20 (3)	5 (-3 to 15)
Absenteeism costs	2543 (447)	2217 (374)	306 (-742 to 1551)
Presenteeism costs	10088 (591)	10390 (1048)	-573 (-2634 to 1717)
Total	**12943 (616)**	**12626 (1111)**	**25 (-2005 to 2485)**

Abbreviations: n: number; SEM: Standard Error of the Mean, CI: Confidence Interval, NA: Not Applicable, SD: Standard Deviation
Note: Costs are expressed in 2011 Euros

Societal perspective: cost-effectiveness

The ICER for body weight was -371, indicating that society has to pay €371 for an additional kilogram body weight loss. An ICER in the similar direction was found for waist circumference (ICER:-392). In both cases, the majority of CE-pairs were located in the north-east quadrant (Table 4; Figure 2 (1a-b)). These results imply that the intervention was more costly and more effective than usual practice, but the wide distribution of CE-pairs around the quadrants of the CE-planes indicates that the uncertainty surrounding these estimates was large (Table 4; Figure 2 (1a-b)). The CEAC in Figure 2 (2a) indicates that if society is not willing to pay anything for a kilogram body weight loss, the probability of cost-effectiveness is 0.41. This probability increased with an increasing willingness-to-pay to 0.84 at a ceiling ratio of €21,000/kg. The CEAC for waist circumference showed a similar picture, with a 0.41 probability at a ceiling ratio of €0/cm and a maximum of 0.77 at a ceiling ratio of €18,000/cm (Figure 2(2b)).

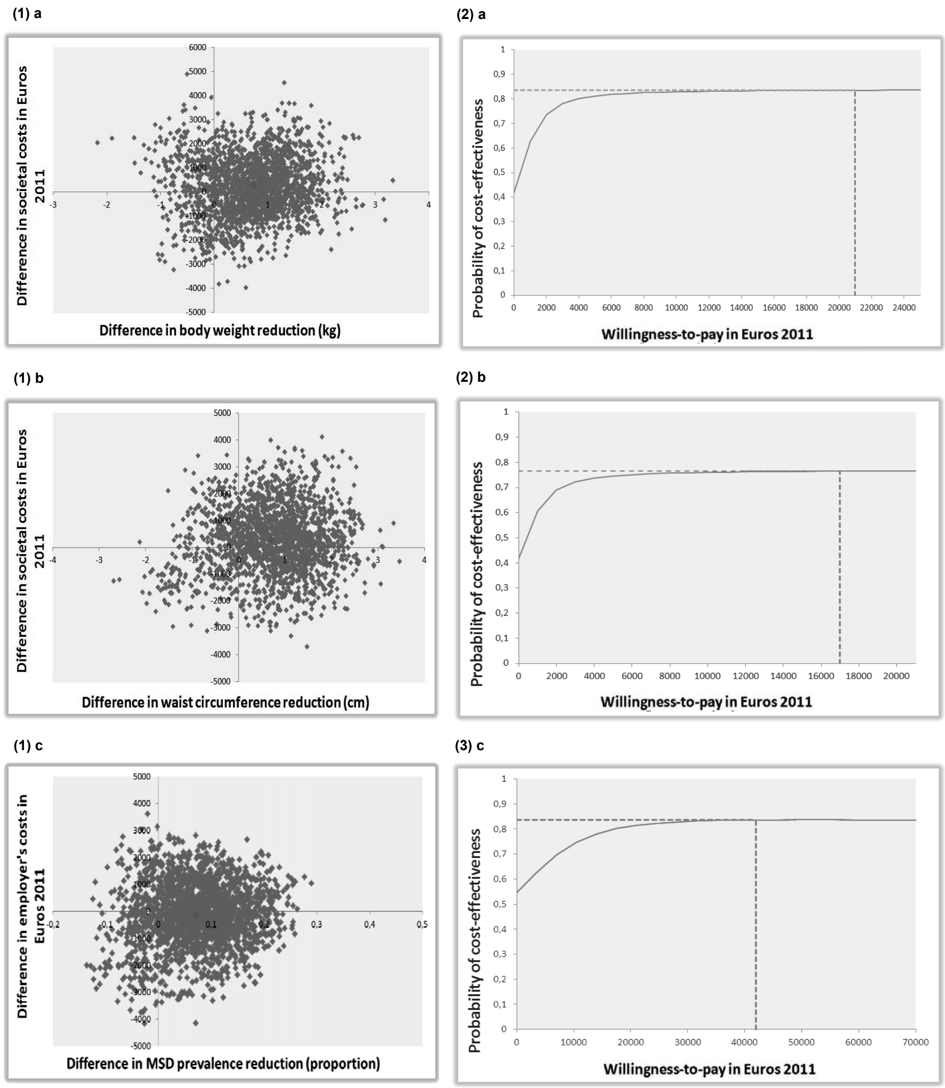

Figure 2: Cost-effectiveness planes indicating the uncertainty around the incremental cost-effectiveness ratios (1) and cost-effectiveness acceptability curves indicating the probability of the intervention being cost-effectiveness at different values (€) of willingness to pay per unit of effect gained (2) for weight loss (a), waist circumference (b), and MSD (c) (based on the imputed dataset).

Note: Effects are expressed in terms of kilogram body weight loss and waist circumference, and MSD prevalence reduction

Table 4: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes

Analysis		Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
				Societal perspective							
		Intervention	Control		€	Points	€/point	NE[1]	SE[2]	SW[3]	NW[4]
Main analysis	- Imputed dataset	162	152	Body weight	271 (-2155 to 2679)	-0.7 (-2.2 to 0.7)	-371	50.0	34.4	6.5	9.1
		162	152	Waist circumference	272 (-2140 to 2692)	-0.7 (-2.5 to 1.1)	-392	48.3	31.2	9.8	10.7
SA1	- Complete-cases	52	47	Body weight	-1228 (-3514 to 576)	-0.5 (-1.8 to 0.8)	2418	10.7	67.9	17.4	4.0
		52	47	Waist circumference	-1196 (-3400 to 602)	-1.1 (-3.0 to 0.8)	1068	13.7	74.4	10.5	1.4
SA2	- Outside work hours	162	152	Body weight	245 (-2181 to 2653)	-0.7 (-2.2 to 0.7)	-334	49.2	35.3	6.6	8.9
		162	152	Waist circumference	246 (-2168 to 2665)	-0.7 (-2.5 to 1.1)	-354	47.6	31.9	10.0	10.5
SA3	- HCA	162	152	Body weight	386 (-2011 to 2794)	-0.7 (-2.2 to 0.7)	-527	53.6	30.9	6.1	9.4
		162	152	Waist circumference	386 (-2001 to 2800)	-0.7 (-2.5 to 1.1)	-556	51.7	27.8	9.2	11.3
SA4	- PRODISQ	162	152	Body weight	-89 (-1586 to 1559)	-0.7 (-2.2 to 0.7)	122	39.2	45.3	9.5	6.1
		162	152	Waist circumference	-89 (-1586 to 1564)	-0.7 (-2.5 to 1.1)	128	36.0	43.5	11.2	9.3
SA5	- Excluding presenteeism costs	162	152	Body weight	799 (-430 to 2317)	-0.7 (-2.2 to 0.7)	-1093	74.5	9.9	2.1	13.5
		162	152	Waist circumference	796 (-433 to 2327)	-0.7 (-2.5 to 1.1)	-1147	69.6	9.9	2.2	18.4
				Employer's perspective							
		Intervention	Control		€	Points/ proportions	€/point	NE[1]	SE[2]	SW[3]	NW[4]
Main analysis	- Imputed dataset	162	152	MSD	-142 (-2674 to 2056)	-0.07 (-0.22 to 0.08)	2000	38.9	44.1	10.0	7.0
		162	152	Work-related vitality (range: 0-6)	-113 (-2583 to 2083)	-0.03 (-0.39 to 0.33)	3322	15.6	28.1	25.0	31.3
		162	152	Job satisfaction (range: 1-5)	-129 (-2610 to 2070)	-0.01 (-0.34 to 0.32)	16328	20.2	27.7	26.1	26.0
SA1	- Complete-cases	52	47	MSD	-1161 (-3027 to 706)	0.01(-0.19 – 0.18)	248800	5.6	45.8	40.4	8.2
		52	47	Work-related vitality (range: 0-6)	-1180 (-3300 to 496)	-0.05 (-0.36 to 0.25)	22121	3.1	33.1	53.5	10.3
		52	47	Job satisfaction (range: 1-5)	-1126 (-3266 to 550)	0.02 (-0.22 to 0.26)	-54230	4.4	52.5	34.4	8.6
SA2	- Outside work hours	162	152	MSD	-171 (-2702 to 2028)	-0.07 (-0.22 to 0.08)	2400	38.1	45.0	10.1	6.8
		162	152	Work-related vitality (range: 0-6)	-142 (-2611 to 2055)	-0.03 (-0.39 to 0.32)	4167	15.2	28.5	25.7	30.7
		162	152	Job satisfaction (range: 1-5)	-158 (-2638 to 2041)	-0.01 (-0.34 to 0.32)	19960	19.6	28.2	26.6	25.6

SA3	- FCA	162	152	MSD	-260 (-2824 to 1914)	-0.07 (-0.22 to 0.08)	3700	35.3	47.7	10.6	6.4
		162	152	Work-related vitality (range: 0-6)	-236 (-2742 to 1954)	-0.03 (-0.39 to 0.32)	9677	13.8	30.0	27.8	28.4
		162	152	Job satisfaction (range: 1-5)	-294 (-2761 to 1946)	-0.01 (-0.34 to 0.32)	30671	18.1	29.7	28.6	23.7
SA4	- PRODISQ	162	152	MSD	-556 (-1811 to 727)	-0.07 (-0.22 to 0.08)	7800	15.6	67.8	12.6	4.0
		162	152	Work-related vitality (range: 0-6)	-535 (-1798 to 760)	-0.03 (-0.39 to 0.32)	16464	8.5	35.5	43.9	12.2
		162	152	Job satisfaction (range: 1-5)	-544 (-1807 to 744)	-0.01 (0.34 to 0.32)	57512	8.4	39.2	40.5	11.8
SA5	- Excluding presenteeism	162	152	MSD	408 (-567 to 1487)	-0.07 (-0.22 to 0.08)	-5700	64.4	19.0	3.0	13.6
		162	152	Work-related vitality (range: 0-6)	422 (-559 to 1517)	-0.03 (-0.39 to 0.32)	-13155	34.9	9.1	12.2	43.7
		162	152	Job satisfaction (range: 1-5)	416 (-563 to 1504)	-0.01 (-0.34 to 0.32)	-43750	36.2	11.4	10.2	42.1

Abbreviations: CI: Confidence Interval, C: Costs, E: Effects, ICER: Incremental Cost-Effectiveness Ratio, CE-plane: Cost-Effectiveness plane, SA: Sensitivity Analysis, HCA: Human Capital Approach, FCA: Friction Cost Approach, MSD: Musculoskeletal Disorders

Note: Costs are expressed in 2011 Euros

[1] Refers to the northeast quadrant of the CE plane, indicating that the VIP in Construction intervention is more effective and more costly than usual practice

[2] Refers to the southeast quadrant of the CE plane, indicating that the VIP in Construction intervention is more effective and less costly than usual practice

[3] Refers to the northwest quadrant of the CE plane, indicating that the VIP in Construction intervention is less effective and more costly than usual practice

[4] Refers to the southwest quadrant of the CE plane, indicating that the VIP in Construction intervention is less effective and less costly than usual practice

Employer's perspective: cost-effectiveness

For MSD, an ICER of 2000 was found, indicating that employers save €2,000 per additional person prevented from having a MSD. Most CE-pairs were contained in the north-east quadrant (Table 4; Figure 2(1c)). This implies that the intervention was less costly and more effective than usual practice, but the level of uncertainty was large. The CEAC in Figure 2 (2c) indicates that the probability of cost-effectiveness was 0.55 at a ceiling ratio of €0/person, increasing to 0.84 at a ceiling ratio of €42,000/person. The ICERs for work-related vitality and job satisfaction were 3322 and 16328, respectively (Table 4). In both cases, the intervention was less costly and less effective than usual practice. CEACs showed that the associated maximum probabilities of cost-effectiveness were 0.54 for both outcomes, irrespective of the willingness-to-pay (Figures not shown).

Employer's perspective: financial return

Total benefits in terms of absenteeism, presenteeism, and occupational health costs were on average €424 (95%CI: -1789 to 2923) (Table 5). The NB was on average 138 (95%CI: -2073 to 2641), suggesting that the intervention resulted in a net saving to the employer of €138 per participant. The BCR (i.e. amount of money returned per Euro invested) and ROI (i.e. percentage of profit per Euro invested) were 1.48 (95%CI: -6.23 to 10.21) and 48% (95%CI: -723 to 921), respectively. However, their confidence intervals were rather wide and none of them was statistically significant.

Sensitivity analyses

The results of SA2 and SA3 were similar to those of the main analysis, whereas the outcomes of SA1 (complete-case analysis), SA4 (PRODISQ), and SA5 (Excluding presenteeism) differed in some aspects from those of the main analysis (Table 4; Table 5).

Table 5: Intervention costs, benefits, Net Benefits (NB), Benefit Cost Ratio (BCR), and Return-On-Investment (ROI) per participant

Analysis		Sample size		Costs	Benefits	Financial return		
		I	C	€	Total (95% CI)	NB[1] (95% CI)	BCR[2] (95% CI)	ROI (%)[3] (95% CI)
Main analysis	- Imputed dataset	162	152	287 (283 to 290)	424 (-1789 to 2923)	138 (-2073 to 2641)	1.48 (-6.23 to 10.21)	48 (-723 to 921)
SA1	- Complete dataset	52	47	289 (283 to 295)	1447 (-265 to 3530)	1158 (-757 to 2948)	5.00 (-1.64 to 11.20)	400 (-264 to 1020)
SA2	- Outside work hours	162	152	258 (258 to 258)	430 (-1783 to 2928)	172 (-2039 to 2677)	1.67 (-6.90 to 11.38)	67 (-790 to 1038)
SA3	- HCA	162	152	287 (283 to 290)	543 (-1697 to 3034)	257 (-1967 to 2769)	1.90 (-5.87 to 10.67)	90 (-687 to 967)
SA4	- PRODISQ	162	152	287 (283 to 290)	840 (-442 to 2099)	553 (-728 to 1814)	2.93 (-1.54 to 7.33)	193 (-254 to 633)
SA5	- Excluding presenteeism	162	152	287 (283 to 290)	-123 (-1142 to 910)	-410 (-1458 to 595)	-0.43 (-4.08 to 3.08)	-143 (-508 to 208)

Abbreviations: CI: Confidence Interval, NB: Net Benefit, BCR: Benefit Cost Ratio, ROI: Return-On-Investment, I: Intervention, C: Control, SA: Sensitivity Analysis, HCA: Human Capital Approach
Note 1: Costs are expressed in 2011 Euros
Note 2: Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0
[1] Indicates the amount of money returned after intervention costs are recovered
[2] Indicates the amount of money returned per Euro invested in the intervention
[3] Indicates the percentage of profit per Euro invested in the intervention

In SA1, total societal and employer's costs were lower in the intervention group than in the control group. All cost and effect differences were not statistically significant. CEACs differed from those of the main analysis (Figures not shown). Most notably, a 0.88 probability of cost-effectiveness was found for body weight at a ceiling ratio of €0/kg, increasing to 0.94 at €1,000/kg. In accordance with the main analysis, financial return estimates were positive, but their confidence intervals were rather wide and not statistically significant.

When using the PRODISQ (SA4), total societal and employer's costs were lower in the intervention group than in the control group. All cost and effect differences were not statistically significant. CEACs differed from those of the main analysis (Figure not shown). Most notably, a 0.54 probability of cost-effectiveness was found for body weight at a ceiling ratio of €0/kg, increasing to 0.84 at €4,000/kg. In accordance with the main analysis, financial return estimates were positive, but their confidence intervals were rather wide and not statistically significant.

When excluding presenteeism costs (SA5), total societal and employer's costs were higher in the intervention group than in the control group. All cost and effect differences were not statistically significant. CEACs differed from those of the main analysis (Figures not shown). Most notably, a 0.22 probability of cost-effectiveness was found for MSD at a ceiling ratio of €0/person, increasing to 0.82 at €100,000/person. In contrast to the main analysis, financial return estimates were negative, but statistically non-significant as well.

DISCUSSION

This study evaluated the cost-effectiveness and financial return of a worksite physical activity and nutrition program for construction workers. In comparison with usual practice, the intervention had no significant effect on all cost and effect measures. The probabilities of cost-effectiveness for body weight, waist circumference, and MSD increased with an increasing ceiling ratio to 0.84 (willingness-to-pay = €21,000/kg), 0.77 (willingness-to-pay = €18,000/cm), and 0.84 (willingness-to-pay = €42,000/person prevented from having MSD), respectively. The probabilities of cost-effectiveness for work-related vitality and job satisfaction were low at all ceiling ratios (≤0.54). Also, per Euro invested in the program, €1.48 was returned to the employer, but the uncertainty surrounding this estimate was large.

Effects and costs

Various reasons may explain the lack of significant effects at 12-month follow-up. First, as the intervention focused on both the prevention and treatment of excessive body weight and MSD, participation in the intervention was not restricted to high-risk individuals (e.g. employees were not pre-selected on high body weight). As a consequence, many participants were relatively healthy at baseline, leaving less room for improvement. Second, a lower than expected number of participants fully participated in the program; e.g. 39% of participants eligible for counselling did not complete the PHC program and most of the VIP in Construction toolbox materials were used by fewer than 50% of participants (48). Third, it is possible that the intensity of the intervention was too low to improve the participants' lifestyle behaviours in such a way that it translates in long-term health improvements. To illustrate, the intervention was previously found effective in reducing body weight at 6-month follow-up (19), but this effect was not sustained at the long-term. To sustain this effect, more counselling contacts and/or booster sessions after the termination of the intervention may be needed. As for the lack of significant cost differences, it is known that cost data are right skewed and therefore require relatively large sample sizes to detect relevant differences. Nonetheless, as in most trial-based economic evaluations, the sample size was based on one of the primary outcomes (i.e. body weight) (18), which likely underpowered it to detect relevant cost differences.

It is noteworthy that the present findings with respect to body weight-related outcomes (i.e. the primary outcomes) contrast those of previous studies. Two systematic reviews found worksite physical activity and nutrition programs to significantly reduce body weight by -1.3kg and -1.2kg (14;49). In addition, Groeneveld et al. (2010) showed in an RCT that a similar intervention for construction workers resulted in a statistically significant body weight loss of -1.8kg at 12-month follow-up (50). The difference in effect between both studies is likely explained by the fact that their intervention was more intensive than ours; i.e. three face-to-face and four telephone contacts versus a maximum of one face-to-face and three telephone contacts. Furthermore, their intervention was aimed at construction workers with an elevated risk of cardiovascular disease, whereas the present intervention was aimed at construction workers in general. This supports our reasoning that a more intensive program, aimed at high-risk individuals, may have been needed to produce better effects.

Societal perspective: Cost-effectiveness

The intervention's cost-effectiveness in improving weight-related outcomes depends on the societal willingness-to-pay for these effects and the probability of cost-effectiveness that society considers acceptable. Since both are unknown, however, strong conclusions cannot be made. Nonetheless, decision-makers themselves can use the present results to consider whether they perceive that the intervention provides "good value for money" at an acceptable probability of cost-effectiveness.

The aforementioned study of Groeneveld et al. (2011) also evaluated the societal cost-effectiveness of the worksite physical activity and nutrition program. They found an ICER of €145/kg body weight loss, a 0.60 probability of cost-effectiveness at a ceiling ratio of €250/kg, which increased to 0.95 at €2,000/kg (51). In contrast to the present study, however, presenteeism and occupational health costs were not included. If we would exclude both cost categories as well, an ICER of €1088/kg body weight loss would be found. Van Wier et al. (2013) evaluated the societal cost-effectiveness of an occupational health guideline aimed at preventing weight gain among Dutch employees. As the probabilities of cost-effectiveness were low for body weight and waist circumference (≤0.52), the intervention was not considered cost-effective (52). Most other studies that evaluated the cost-effectiveness of similar interventions in improving weight-related outcomes solely included intervention costs (53).

Employer's perspective: Cost-effectiveness

The intervention was not cost-effective in improving work-related vitality and job satisfaction (≤0.54 probabilities of cost-effectiveness). If employers are not willing to pay anything for preventing one person from having a MSD, there is a 0.55 probability of the intervention being cost-effective. This probability increased with an increasing willingness-to-pay to 0.84 at a ceiling ratio of €42,000/person. Again, however, strong conclusions about the intervention's cost-effectiveness in terms of this outcome cannot be made, and employers themselves should consider whether the intervention provides "good value for money" at an acceptable probability of cost-effectiveness.

To our knowledge, studies evaluating the employer's cost-effectiveness of similar interventions in improving work-related vitality and MSD are lacking. One study, however, evaluated the employer's cost-effectiveness in improving job satisfaction of a mindfulness-based worksite intervention aimed at improving work engagement and energy balance-related behaviours (54). Irrespective of the maximum willingness-to-pay, the intervention had a low probability of cost-effectiveness (≤0.25) and was therefore not considered cost-effective in improving job satisfaction either.

Employer's perspective: Financial return

On average, €1.48 was returned to the employer per Euro invested in the program. However, as the uncertainty surrounding the financial return estimates was large and none of them was statistically significant, it cannot be concluded that the intervention was cost-beneficial to the employer.

A systematic review found worksite physical activity and/or nutrition programs to result in positive financial returns in terms of absenteeism benefits according to non-randomized studies (BCR: 4.25), but negative financial returns according to RCTs (BCR: 0.51). If we would solely include absenteeism benefits, our results would be in line with those of the review (BCR: 0.41). The review also indicated that the current evidence on the financial return of such interventions is limited by the fact that few studies incorporate presenteeism benefits and none of them report on the uncertainty surrounding their results. The present findings underscore the importance of addressing these limitations. Namely, as financial return estimates were positive, but statistically non-significant, wrong conclusions would have been drawn if the level of uncertainty was not taken into account. Furthermore, the direction of the financial return estimates proved to be highly influenced by the in- or exclusion of presenteeism benefits; i.e. positive when included, but negative when excluded.

Robustness of the study results

In accordance with the main analysis, cost and effect differences as well as financial return estimates were not statistically significant in all sensitivity analyses. Also, the overall conclusions would not change when using the results of any of the sensitivity analyses. Nonetheless, it is important to mention that the results of the complete-case analysis (SA1) were much more favorable than those of the main analysis. Amongst others, relatively high probabilities of cost-effectiveness were found at ceiling ratios of €0; e.g. a 0.88 probability at a ceiling ratio of €0/kg body weight loss. However, as a post-hoc analysis indicated that participants with complete data had fewer sickness absence days during follow-up than those with incomplete data (i.e. 6.7 versus 13.3 in the intervention group and 9.5 versus 10.9 in the control group), self-selection of participants seems to have biased these results, and the results of the main analysis were considered more valid.

Strengths and limitations

An important strength of the present study is its pragmatic RCT design. The pragmatic aspect of the trial enabled us to evaluate the intervention's resource implications under "real world" circumstances. This facilitates the generalizability of the results (i.e. external validity), whereas the internal validity is guaranteed by the randomization of participants (55;56). Another strength concerns the use of state-of-the-art statistical methods that are not or infrequently used in occupational health research. Amongst others, multiple imputation was used to deal with missing data, SUR analyses were performed to account for the possible correlation between costs and effects/benefits, and bootstrapping was used to estimate the uncertainty surrounding cost differences as well as cost-effectiveness and financial return estimates. Furthermore, both absenteeism and presenteeism costs were included, whereas most previous studies solely included absenteeism costs (45;53). This is of importance because efforts to improve health seem to have a more immediate effect on presenteeism than on absenteeism (57).

Several limitations deserve attention as well. First, complete cost and effect data were only obtained from 40.5% and 62.4% of participants, respectively. To deal with this issue, missing values were imputed using multiple imputation. While having complete data is always preferred, multiple imputation is increasingly being acknowledged as a more valid and precise way to deal with missing data than a complete-case analysis (56;58).Complete-case analyses reduce the power of a study and ignore available information of participants who only have missing data on a few measurement points. Also, complete-case analyses only produce reliable estimates when there are no systematic differences between the missing and observed values, which, according to a post-hoc analysis, was probably not the case (40;58). Second, many cost and effect data were gathered using self-report of participants, which may have causes "social desirability bias" and/or "recall bias". Amongst others, we had to rely on self-reported values of healthcare utilization as health insurance claim data of participants are practically inaccessible in Dutch economic evaluations. Also, the period over which participants had to report their resource use was relatively long (i.e. 3 months). This might be a particular concern for presenteeism, as relatively short recall periods seem to be needed for this outcome (59). In future studies, mobile apps might provide a solution for this issue, as they can be used to collect data in a way that is relatively non-burdensome to participants. Third, the presence of MSD was assessed in terms of *"self-reported pain or discomfort in one or more body regions".* As discomfort can be regarded as an early manifestation of MSD, participants classified as having MSD may not necessarily have serious functional limitations and/or low levels of health-related welfare. This should be kept in mind while interpreting the results. It is also important to bear in mind that economic evaluation results are not directly transferable between countries or jurisdictions due to differences in healthcare and/or social security systems (60;61). In the Netherlands, for example, healthcare costs are generally borne by the government and/or health insurance companies, whereas in countries with employer-provided healthcare (e.g. The United States) they accrue to the employer. Furthermore, for the employer's perspective, the HCA was used for estimating absenteeism costs. This was done because Dutch employers are obliged to pay at least 70% of the salary of sick employees for a period of two years, and most of them top up the wage payments from 70% to 100% during the first year of sickness absence (62). Thus, although the initial productivity level of a Dutch company may be restored after the friction period, employers still bear the salary costs of a sick worker. Readers should keep in mind that alternative valuation methods may be more appropriate in other countries or jurisdictions (61).

Conclusion

The intervention's cost-effectiveness in improving weight-related outcomes and MSD depends on the societal and employer's willingness to pay for these effects and the probability of cost-effectiveness that they consider acceptable. From the employer's perspective, the intervention was not cost-effective in improving work-related vitality and job satisfaction. Also, due to a large degree of uncertainty, it cannot be concluded that the intervention is cost saving to the employer

Acknowledgements

This project is part of a research program called "Vitality In Practice", which is funded by Fonds Nuts Ohra (Nuts Ohra Foundation). The authors wish to thank Anneke van Paridon for her help with the data collection. The authors would also like to thank all participants and health coaches.

REFERENCES

1. Overweight, obesity, and health risk. National Task Force on the Prevention and Treatment of Obesity. *Arch Intern Med* 2000;160(7):898-904.
2. Anandacoomarasamy A, Fransen M, March L. Obesity and the musculoskeletal system. *Curr Opin Rheumatol* 2009;21:71-77.
3. Shiri R, Karppinen J, Leino-Arjas P, Solovieva S, Viikari-Juntura E. The association between obesity and low back pain: a meta-analysis. *Am J Epidemiol* 2010;171:135-154.
4. Woolff AD, Phleger B. Burden of major musculoskeletal conditions. *Bull World Health Org* 2003;81(9):646-56.
5. World Health Organization. Overweight and Obesity Fact sheet. http://www.who.int/mediacentre/factsheets/fs311/en/
6. Statistics Netherlands. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=81565NED&D1=a&D2=a&D3=a&D4=0&D5=30&HDR=T&STB=G1,G2,G3,G4&VW=T
7. Wijnhoven HAH, de Vet HCW, Picavet HS. Prevalence of Musculoskeletal Disorders Is Systematically Higher in Women Than in Men*.* *Clin J Pain* 2006;22(8):717-724.
8. Arbouw. Bedrijfstaksatlas 2012. http://www.arbouw.nl/pdf/tools/bedrijfstak-atlas-2012.pdf
9. Boschman J, van der Molen H, Sluiter J, Frings-Dresen M. Musculoskeletal disorders among construction workers: a one-year follow-up study. *BMC Musculoskeletal Disorders* 2012;13(1):196.
10. Council for Public Health and Health Care (Raad voor de Volksgezondheid en Zorg). Gezondheid en gedrag 2002.
11. Polder JJ, Takkern J, Meerding WJ, Kommer GJ, Stokx LJ. Cost of illness in the Netherlands Bilthoven: RIVM; 2002. Report No.: 270751005.
12. Lambeek LC, van Tulder MW, Swinkels ICS, Koppes LLJ, Anema JR, van Mechelen W. The trend in total cost of back pain in the Netherlands in the pseriod 2002 to 2007. *Spine (Phila Pa 1976)* 2011;36(13):1050-1058.
13. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Annu Rev Public Health* 2008;29:303-323.
14. Verweij LM, Coffeng J, Van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. Obes Rev 2011;12(6):406-429.
15. 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. *Clin J Sport Med* 2003;13(2):106-17.
16. Neovius K, Johansson K, Kark M, Neovius M. Obesity status and sick leave: a systematic review. *Obes Rev* 2009;10(1):17-27.
17. Cancelliere C, Cassidy JD, Ammendolia C, Cote P. Are workplace health promotion programs effective at improving presenteeism in workers? a systematic review and best evidence synthesis of the literature. *BMC Public Health* 2011;11(1):395.
18. Viester L, Verhagen EA, Proper KI, van Dongen JM, Bongers PM, van der Beek AJ. VIP in construction: systematic development and evaluation of a multifaceted health programme aiming to improve physical activity levels and dietary patterns among construction workers. *BMC Public Health* 2012;12(1):89.

6

19. Viester L, Verhagen EALM, Bongers PM, van der Beek AJ. Improvement of dietary and physical activity behaviour through a worksite intervention in construction workers: results of a randomised controlled trial. Submitted

20. Viester L, Verhagen EALM, Bongers PM, van der Beek AJ. The effect of a health promotion intervention for construction workers on work-related outcomes: a randomised controlled trial. Submitted

21. Miller P, Rossiter P, Nuttall D. Demonstrating the economic value of occupational health services. Occup Med 2002;52(8):477-483.

21. van Dongen JM, Tompa E, Clune LA, Sarnocinska-Hart A, Bongers PM, van Tulder MW, et al. Bridging the gap between the economic evaluation literature and daily practice in health and safety: a study into the information needs of decision makers in the healthcare sector. *Implement Sci* 2013;8:57

23. Cavallo D. Using return on investment analysis to evaluate health promotion programs: challenges and opportunities. *Health Promotion Economics Issue Briefs* 2006;1(3):1-4. RTI-UNC Center of Excellence. Available at: http://www.rti.org/pubs/IssueBrief_3.pdf.

24. Stone PW. Return-on-investment models. *Appl Nurs Res* 2005;18(3):186-189.

25. American College of Sports Medicine Position Stand. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Med Sci Sports Exerc* 1998;30(6):975-991.

26. Kemper HGC, Ooijendijk WTM, Stiggelbout M. Consensus over de Nederlandse Norm voor Gezond Bewegen. *Tijdschr Soc Gezondheidsz* 2000;78:180-183.

27. Ajzen I. The theory of planned behavior. *Orga Behav Hum Decis Process* 1991;50:179-211.

28. Lean ME, Han TS, Morrison CE. Waist circumference as a measure for indicating need for weight management. *BMJ* 1995;311:158-161.

29. Hildebrandt VH, Bongers PM, van Dijk FJ, Kemper HC, Dul J. Dutch Musculoskeletal Questionnaire: description and basic qualities. *Ergonomics* 2001;44:1038-1055.

30. Schaufeli WB, Bakker AB. Bevlogenheid: een begrip gemeten. *Gedrag en Organisatie* 2004;17:90-112.

31. Koppes LLJ, de Vroome EMM, Mol MEM, Janssen BJM, van den Bossche SNJ. *Nationale enquete arbeidsomstandigheden 2010: Methodologie en globale resultaten. 2010.* http://www.tno.nl/downloads/rapport_nea_20111.pdf

32. FD. Microcosting Quantity Data Collection Methods. *Med Care* 2009;47(7 Suppl 1):S76-S81.

33. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding Voor Kostenonderzoek. Methoden en Standaardkostprijzen *Voor Economische Evaluaties in de Gezondheidszorg*. Diemen, the Netherlands: College Voor Zorgverzekeringen; 2010.

34. Z-index. G-Standard. The Hague, The Netherlands: Z-Index BV; 2009.

35. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;14(2):171-189.

36. Kessler R, Ames M, Hymel P, Loeppke R, McKenas D, Richling D. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46:S23-S37.

37. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D. The world health organization health and work performance questionnaire (HPQ). *J Occup Environ Med* 2003;45:156-174.

38. Statistics Netherlands. 2011. http://www.cbs.nl
39. Drummond MF, Sculpher MJ, Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.
40. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statist Med* 2011;30(4):377-399.
41. Rubin DB. *Multiple imputation for nonresponse in surveys.* John Wiley & Sons: New York, 1987.
42. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data*. Health Econ* 2004;13(5):461-475.
43. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10(3):212-214.
44. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. *Health Econ* 2004;13(5):405-415.
45. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011;12(12):1031-1049.
46. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5:23-28.
47. Koopmanschap M, Meeding WJ, Evers S, Severens J, Burdorf A, Brouwer W. *Handleiding voor het gebruik van PRODISQ versie 2.1.* [Handbook on use of PRODISQ.] Rotterdam/Maastricht, Erasmus MC - Instituut voor Medical Technology Assessment, Instituut Maatschappelijke Gezondheidszorg, Universiteit van Maastricht - Beleid Economie en Organisatie van de Zorg; 2004.
48. Viester L, Verhagen EALM, Bongers PM, van der Beek AJ. Process evaluation of a multifaceted health programme aiming to improve physical activity levels and dietary patterns among construction workers. 2013
49. Anderson LM, Quinn TA, Glanz K, Ramirez G, Kahwati LC, Johnson DB, et al. The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: a systematic review. *Am J Prev Med* 2009;37(4):340-357.
50. Groeneveld IF, Proper KI, van der Beek AJ, van Mechelen W. Sustained body weight reduction by an individual-based lifestyle intervention for workers in the construction industry at risk for cardiovascular disease: results of a randomized controlled trial. *Prev Med* 2010;51:240-246.
51. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, Van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the contruction industry at risk for cardiovascular disease. *J Occup Environ Med* 2011;53(6):610-617.
52. van Wier MF, Verweij LM, Proper KI, Hulshof CTJ, van Tulder MW, van Mechelen W. Economic evaluation of an occupational health care guideline for prevention of weight gain among employees. J Occup Environ Med 2013; 55(9):1100-1109.
53. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, et al. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. Scand J Work Environ Health 2012;38(5):393-408.

54. van Dongen JM, van Berkel J, Boot CRL, Bosmans JE, Proper KI, Bonges PM, van der Beek AJ, van Tulder MW, van Wier MF. Cost-effectiveness and financial return of a mindfulness-based worksite intervention aimed at improving work engagement: results of a randomized controlled trial. Submitted

55. Tompa E, Dolinschi J, de Oliveira C. Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil* 2006;16:375-400.

56. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ* 2011;342: d1548.

57. Caverley N, Cunningham JB, MacGregor JN. Sickness Presenteeism, Sickness Absenteeism, and Health Following Restructuring in a Public Service Organization. *Journal of Management Studies* 2007;44(2):304-319.

58. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;29:338.

59. Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: A critical review. *Soc Sci Med* 2011;72(2):185-192.

60. Verbeek J, Pulliainen M, Kankaanpaa E, Taimela S. Transferring results of occupational safety and health cost-effectiveness studies from one country to another - a case study. *Scand J Work Environ Health* 2010;36(4):305-312.

61. Tompa E, Culyer AJ, Dolinschi J. *Economic evaluation of interventions for occupational health and safety: Developing good practice*. New York: Oxford University Press; 2008.

62. OECD. Sickness and Disability Schemes in the Netherlands: Country memo as a background paper for the OECD Disability Review. 2007. http://www.oecd.org/social/soc/41429917.pdf.

63. Groeneveld IF, Proper KI, van der Beek AJ, van Duivenbooden C, van Mechelen W. Design of a RCT evaluating the (cost-) effectiveness of a lifestyle intervention for male construction workers at risk for cardiovascular disease: the health under construction study. *BMC Public Health* 2008;8:1.

64. UWV. Brancheschets Bouw. Oktober 2012. https://www.werk.nl/pucs/groups/public/documents/document/wdo_007277.pdf

Appendix 1: Price weights used for valuing resource use and resources consumed by the intervention and control group participants during follow-up (based on the complete-cases)

Units [Units of measurement]	**Price weight**		**Resources consumed**	
	Societal perspective	**Employer's perspective**	**Intervention group (n=51)**	**Control group (n=48)**
Intervention costs	€ 177.77	€ 287.56		
Medical costs				
Visits to a care provider [No. of visits; mean (SD)]				
General practitioner				
Office consultation	€ 28.96[c]	N.A.	1.3 (1.9)	1.6 (2.2)
Telephone consultation	€ 14.48[c]	N.A.	0.2 (0.5)	0.2 (0.8)
House call	€ 44.47[c]	N.A.	0.0 (0.3)	0.0 (0.2)
Allied health professionals				
Psychologist	€ 82.47[c]	N.A.	0.8 (3.3)	0.2 (0.1)
Dietician	€ 27.93[c]	N.A.	0.0 (0.0)	0.0 (0.3)
Physical therapist	€ 37.23[c]	N.A.	0.7 (2.3)	3.8 (8.0)*
Other allied health professionals	Variable[c,d]	N.A.	0.7 (3.7)	0.5 (1.9)
Medical specialists				
Psychiatrist	€ 106.53[c]	N.A.	0.0 (0.0)	0.0 (0.0)
Other medical specialists	€ 74.47[c]	N.A.	0.8 (1.7)	0.8 (1.8)
Complementary medicine	Variable[c,d]	N.A.	0.2 (1.7)	0.4 (1.8)
Hospitalization [No. of days; mean (SD)]				
Ward	€ 472.66[c]	N.A.	0.2 (0.2)	0.3 (0.8)
Intensive care	€ 2257.82[c]	N.A.	0.0 (0.0)	0.0 (0.0)
Medications [No. of participants using medication; Number (%)]	Variable[e]	N.A.	30 (58.8)	25 (52.1)
Absenteeism costs				
Sickness absence [days; Mean (SD)]	198.20[f]	213.10[g]	6.7 (9.5)	9.4 (21.9)
Presenteeism costs				
Presenteeism [days; Mean (SD)]	198.20[f]	213.10[g]	43.7 (14.5)	46.3 (19.7)
Sports costs [No. of participants with sports costs; Number (%)]	Variable[h]	N.A.	36 (70.6)	23 (47.9)*
Occupational health costs				
In-company fitness [No. of months; mean (SD)]	€ 10.00[i]	€ 10.00[i]	0.9 (2.5)	0.4 (1.6)

* Significant at p<0.05
Abbreviations: n: Number, SD: Standard Deviation, N.A.: Not ApplicableNote: Costs are expressed in 2011 Euros
Price weight sources: [a] Bottum-up micro-costed, valued using tariffs and depleted sources (See Appendix 2); [b] Market prices, valued using invoices of contractors; [c] Dutch Manual of Costing; [d] Professional organizations; [e] Dutch Society of Pharmacy; [f] Average gross annual salary of Dutch construction workers including holiday allowances and premiums; [g] Average gross annual salary of blue collar workers of the participating construction company including holiday allowances and premiums; [h] Self-reported expenses on sports memberships and sports equipment; [i] Height of the employer's gym membership subsidy

6

Appendix 2: Cost of the VIP in Construction intervention from the societal perspective, valued using a bottom-up micro-costing approach (Euros 2011)

	Cost categories	Intervention staff / Worker	Units	Unit Prices	Total Costs (Euros 2011)	Mean costs per worker (SD) (Euros 2011)
			VARIABLE COSTS			
Personal Health Coaching (PHC)						
Face-to-face contacts	Labor costs	Health coach	297.2 hours[a]	€ 35.72/ hour	€ 10,615.98	€ 65.53 (35.93)
		Construction worker	104.0 hours	€ 27.53/ hour	€ 2,863.12	€ 17.67 (11.11)
	Capital costs		114.5 hours	€ 0.44/ hour	€ 50.07	€ 0.31 (0.19)
	Travelling costs		14563.6 kilometres	€ 0.21/ km	€ 3,077.52	€ 18.95 (10.15)
Telephone contacts	Labor costs	Health coach	65.9 hours	€ 35.72/ hour	€ 2,353.86	€ 14.53 (16.12)
		Construction worker	52.3 hours	€ 27.53/ hour	€ 1,440.18	€ 8.89 (10.22)
	Capital costs	Health coach	65.9 hours	€ 0.44/ hour	€ 29.16	€ 0.18 (0.20)
	Capital costs	Construction worker	52.3 hours	€ 0.44/ hour	€ 22.68	€ 0.14 (0.16)
	Phone costs		3132 minutes	€ 0.09/ min	€ 281.88	€ 1.74 (2.00)
					Subtotal:	**€ 127.96 (76.81)**
Health information brochure						
PHC-group A			40 brochures	€ 4.68/ brochure	€ 187.08	€ 1.16 (2.02)
PHC-group B			61 brochures	€ 4.68/ brochure	€ 285.30	€ 1.76 (2.28)
PHC-group C			48 brochures	€ 3.20/ brochure	€ 153.60	€ 0.95 (1.47)
No counselling			13 brochures	€ 4.80/ brochure	€ 62.40	€ 0.39 (1.31)
					Subtotal:	**€ 4.25 (0.69)**
TOTAL VARIABLE COSTS						**€ 132.21 (76.88)**

			FIXED COSTS			
Quick scan processing	Labor costs	Occupational physician	27 hours	€ 168.47/ hour	€ 5,548.69	€ 28.08
	Capital costs			€ 3.90/ hour	€ 105.30	€ 0.65
	Material costs		162 quick scans	€ 0.20/ quick scan	€ 32.40	€ 0.20
					Subtotal:	**€ 28.93**
VIP in Construction toolbox						
Waist circumference measuring tape / BMI card			162 packages	€ 3.10/ package	€ 502.66	€ 3.10
Calorie guide			162 calorie guides	€ 0.74/ guide	€ 119.63	€ 0.74
Pedometer			162 pedometer	€ 2.00/ pedometer	€ 324.00	€ 2.00
Cookbook			162 cookbooks	€ 0.80/ cookbook	€ 129.60	€ 0.80
Exercise card / Personal energy plan forms			162 packages	€ 0.74/ package	€ 119.63	€ 0.74
					Subtotal:	**€ 7.38**
Recruitment folder			162 folders	€ 0.77/ folder	€ 123.74	€ 0.77
					Subtotal:	**€ 0.77**
Development intervention[b]					€ 116,107.80	€ 8.48[c]
					Subtotal:	**€ 8.48**
TOTAL FIXED COSTS						**€ 45.56**
TOTAL INTERVENTION COSTS						**€ 177.77 (76.88)**

Abbreviations: SD: Standard Deviation, PHC: Personal Health Coaching, BMI: Body Mass Index

[a] The time investment of the health coaches includes travelling, preparation, and coaching time

[b] €116,107.80 was paid for the development of the VIP in Construction intervention. For calculating the development costs per participant, these were divided by the expected number of program users during the first five years after implementation (13,695). In the Netherlands, 221,250 construction workers are employed by a medium- (10-100 employees) or large-scale (>100 employees) construction company(63;64). During the five year period, it was hypothesized that the intervention will be offered to 20% of these construction workers and that 31% of them will participate to the intervention(48).

7

Cost-effectiveness and return-on-investment analysis of a combined social and physical environmental intervention in office employees

Johanna M van Dongen
Jennifer Coffeng
Marieke F van Wier
Ingrid JM Hendriksen
Cécile RL Boot
Willem van Mechelen
Paulien M Bongers
Allard J van der Beek
Judith E Bosmans
Maurits W van Tulder

To be submitted

ABSTRACT

Objective: To evaluate the cost-effectiveness and financial return of a combined social and physical environmental intervention in office employees in comparison with usual practice, and of both intervention conditions separately. Moreover, the probabilities of the interventions being cost-effective in comparison with each other were explored.

Methods: In a 2X2 factorial design, 412 employees were allocated at the department level to the combined intervention (n=92), social environmental intervention (n=118), physical environmental intervention (n=96), or control group (n=106). The social environmental intervention consisted of group motivational interviewing. The physical environmental intervention consisted of environmental modifications to the workplace. Control group participants received usual practice. Data on need for recovery (NFR), general vitality, and job satisfaction were collected at baseline, six, and 12 months. Salary and sickness absence data were collected from company records. Data on healthcare utilization and expenses on sports memberships, sports equipment and active commuting equipment were collected using 3-monthly questionnaires. Using linear multilevel analyses, cost-effectiveness analyses were conducted from the societal and employer's perspective, and return-on-investment analyses from that of the employer. Uncertainty was assessed using bootstrapping techniques, and shown in cost-effectiveness planes and cost-effectiveness acceptability curves.

Results: At 12 months, combined intervention group participants significantly improved their NFR compared with the control group (-8.4; 95%CI -14.6 to -2.2). Their total employer's costs, however, were significantly higher than those of the control group (3102; 95%CI 598 to 5969). All other between-group differences in costs and effects were not significant. For NFR, the combined intervention became the preferred option in comparison with the other interventions at willingness-to-pay values of €170 (societal perspective) and €300 (employer's perspective) per point improvement, after which its probability of cost-effectiveness gradually increased to 0.85. For general vitality and job satisfaction, the probabilities that the interventions were cost-effective in comparison with each other were low at all ceiling ratios (≤0.55), as were their probabilities of financial return (≤0.41).

Conclusion: Depending on the societal and employer's willingness-to-pay and the probability of cost-effectiveness that they consider acceptable, the combined intervention may be considered cost-effective in improving NFR. Both separate interventions were not cost-effective in improving this outcome. Moreover, all interventions were neither cost-effective in improving general vitality (societal perspective) and job satisfaction (employer's perspective), nor cost saving to the employer.

INTRODUCTION

During the last decades, the pressure at work has increased substantially (1). Currently, 36% of Dutch workers *"regularly have to work at a high work pace"* and 30% *"regularly have to work under high time pressure"* (2). As a consequence, many workers experience higher levels of work stress compared to a couple of decades ago (3). Work stress is defined as the psychological and physical state that results when individual resources are insufficient to cope with the demands and pressures of work (4). If stress persists, there may be changes in immunological, neuroendocrine, cardiovascular, and autonomic functioning, leading to mental and physical ill health (e.g. mental disorders, cardiovascular disease) and an associated increase in healthcare and productivity-related costs (4-6).

Need for recovery (NFR) from work-related fatigue seems to be an important intermediate factor in the relation between short-term work stress and longer-term mental and physical ill health (7-10). Previous research indicates that the ability to recover from work may be enhanced by improving a worker's level of physical activity and relaxation (11-14). Therefore, in the Be Active & Relax "Vitality In Practice" (VIP) study, a worksite health promotion program was developed aimed at reducing NFR among office employees by improving their physical activity and relaxation (15). The intervention was developed in close cooperation with stakeholders of the participating company and consisted of both a social and physical environmental component. Such a social ecological intervention approach was chosen because interventions targeted at both the individual and environmental determinants of behaviours are expected to be more effective in achieving health behaviour change than those that are solely targeted at individual determinants (15-17).

Evaluations of the intervention's effectiveness in terms of health- and work-related outcomes have been reported elsewhere (18;19). However, as resources for occupational health are scarce, employers are not just interested in the effectiveness of worksite health promotion programs, but also in their impact on the company's bottom-line (20). This can be determined with a return-on-investment (ROI) analysis, in which the costs of an intervention are compared to its resulting financial benefits to the employer (21;22). Various program outcomes, however, are hard to monetize (e.g. health outcomes, job satisfaction) and can therefore not be included in a ROI analysis. Therefore, cost-effectiveness analyses (CEAs), in which the incremental costs of an intervention are compared to its incremental effects, are also important (23).

The present study aimed to evaluate the cost-effectiveness and financial return of the combined social and physical environmental intervention in office employees in comparison with usual practice, as well as those of both intervention conditions separately. Additionally, the probabilities of the interventions being cost-effective in comparison with each other were explored. CEAs were performed from both the societal and employer's perspective, and the ROI analysis from that of the employer. The combined intervention was hypothesized to produce the most favourable results.

METHODS

Study design and study population

This study was performed alongside a 12-month trial with a pragmatic 2X2 factorial design. Data collection took place in the Netherlands from September 2011 up until December 2012. Full details of the study design, development and content of the intervention, as well as the sample size calculation have been published elsewhere (15). The study design and informed consent procedure were approved by the Medical Ethics Committee of the VU University Medical Center, Amsterdam, the Netherlands.

In September 2011, all 1,182 office employees of 24 departments of a financial service provider were invited to participate in the study (i.e. they received an invitation letter, information on the study, informed consent form, and a baseline questionnaire). Those who were on sickness absence during the previous four weeks were not eligible to participate. A total of 412 employees (response: 35%) from 19 departments signed the informed consent form and completed the baseline questionnaire. Subsequently, their respective departments were either stratified to the "physical environmental intervention" or the "no physical environmental intervention" group. Within these strata, departments were subsequently randomized to either the "social environmental intervention" or the "no social environmental intervention" group by means of tossing a coin. This resulted in four research groups: 1) combined social and physical environmental intervention group; 2) social environmental intervention group only; 3) physical environmental intervention group only; 4) control group (usual practice). Group allocation was performed at the department level, because the interventions under study acted on the group-level rather than on the individual-level as well as to minimize contamination between study groups. As a result of the nature of the intervention, blinding of participants and intervention providers was not possible.

Social and physical environmental intervention conditions

Social environmental intervention condition

The social environmental intervention condition consisted of Group Motivational Interviewing (GMI). GMI was delivered by the team leaders of the departments after receiving a 2-day GMI-training course, which was provided by a GMI-professional. During the intervention period, team leaders also participated in two 90-minute GMI-coaching sessions. These sessions took place at the workplace and during work hours, and were provided by a GMI-professional as well. Within a period of six weeks (i.e. three weeks between sessions), team leaders provided three 90-minute GMI-sessions to their own team. Two months after the final session, a booster GMI-session was provided. All GMI-sessions took place at the workplace and during work hours. GMI-sessions were supported by a GMI-session workbook and a web-based social media platform.

Physical environmental intervention condition

As part of the physical environmental intervention condition, several so-called “VIP zones” were created at the workplace, including: 1) the VIP Coffee Corner Zone – the coffee corner was modified by adding a bar with bar chairs, a large plant, and a giant relaxing wall poster, 2) the VIP Open Office Zone – the office was modified by introducing exercise balls and curtains to divide desks in order to reduce background noise, 3) the VIP Meeting Zone – conference rooms were modified by placing a standing table and a giant relaxing wall poster, and 4) the VIP Hall Zone - table tennis tables were placed and lounge chairs were introduced in the hall for informal meetings. Moreover, footsteps were placed on the floor of the entrance hall to promote stair walking. All environmental modifications were promoted through banners in the VIP Coffee Corner and digital flyers.

Effect measures

Primary and secondary outcomes were assessed at baseline, six, and 12 months.

Primary outcome

The primary outcome was NFR. NFR was assessed using a subscale of the "Dutch Questionnaire on the Experience and Evaluation of Work", which consists of 11 dichotomous items (yes/no). The NFR Score ranges from 0-100, with lower scores indicating a better NFR (24).

Secondary outcomes

Secondary outcomes were general vitality and job satisfaction. General vitality was assessed using the RAND-36 Vitality Scale, which includes four items assessing a participant's general vitality during the previous four weeks. Items were scored on a 6-point scale ranging from "all of the time"(1) to "none of the time"(6). The RAND-36 Vitality Score ranges from 0-100, with higher scores indicating a better general vitality (25). Job satisfaction was assessed using a 1-item question of the "Netherlands Working Conditions Survey". Participants were asked to rate their overall job satisfaction on a 5-point scale ranging from "very dissatisfied"(1) to "very satisfied"(5) (26).

Measurement and valuation of resource use

Intervention costs

For the societal perspective, a bottom-up micro-costing approach was used for estimating intervention costs, meaning that detailed data were collected regarding the quantity of resources consumed as well as their unit prices (27). Intervention costs included those related to the development, implementation, and operation of the intervention (conditions) (e.g. costs for recruiting participants, GMI-training courses, GMI-sessions, GMI-coaching sessions, GMI-website hosting and maintenance, printed materials, VIP zones). Frequency and duration of GMI-training courses, GMI-sessions, and GMI-coaching sessions were registered by the team leaders and GMI-professionals. Labor costs of intervention providers were valued by multiplying their total time investments by their gross hourly salaries including overhead costs. VIP zone costs were based on invoices and were linearly depreciated over a period of 5 years with a scrap value of zero. Capital costs were valued using cost data collected from finance department staff. Printed material and website hosting costs were estimated using invoices. Development costs were estimated by dividing the total costs related to the development of the intervention by the expected number of program users during the first five years after implementing it broadly. For the employer's perspective, intervention costs were valued using charges paid.

Healthcare costs

Healthcare utilization was assessed using 3-monthly questionnaires and included primary healthcare (e.g. general practitioner, allied health professionals, complementary medicine), secondary healthcare (e.g. medical specialist, hospitalization), and both prescribed and over-the-counter medications. Primary and secondary healthcare utilization were valued using Dutch standard costs (28). If unavailable, prices according to professional organizations were used. Medication use was valued using unit prices derived from the Dutch Royal Society of Pharmacy (29).

Absenteeism costs

Baseline (i.e. a one year period prior to baseline) and follow-up sickness absence data as well as gross annual salaries of participants were collected from company records.

Costs associated with one sickness absence day were calculated per participant by dividing their gross annual salary (including overhead costs) by their total number of workable days per year (28). If the societal perspective was applied, absenteeism costs were estimated using the "Friction Cost Approach" (FCA), with a friction period of 23 weeks and an elasticity of 0.8 (30). For the employer's perspective, absenteeism costs were estimated using the "Human Capital Approach" (HCA), in which costs are neither truncated to the friction period, nor is an elasticity factor applied.

Presenteeism costs

Presenteeism was assessed on a 3-monthly basis using an item of "The World Health Organization Health and Work Performance Questionnaire"(WHO-HPQ) (31;32). In the WHO-HPQ, presenteeism is conceptualized as a measure of actual work performance in relation to "best performance", irrespective of the presence or absence of health complaints. Participants were asked to rate their overall work performance during the previous three months on an 11-point scale (range: "worst performance" (0) to "best performance" (10)). Subsequently, their average work performance during follow-up (*Wown*) was estimated and the participants' level of presenteeism (P_{HPQ}) was calculated using the following formula:

$$P_{HPQ} = (10 - Wown)/10$$

The total number of days lost due to presenteeism were calculated by multiplying the participants' P_{HPQ} by their number of days worked during follow-up; i.e. working days minus sickness absence days. These days were subsequently valued using the participants´ gross annual salaries (including overhead costs).

Sports costs

Participants' expenses on sports memberships and sports equipment were assessed using 3-monthly questionnaires.

Active commuting equipment costs

Participant's expenses on active commuting equipment (i.e. a bike) were assessed using 3-monthly questionnaires.

All costs were converted to 2011 Euros using consumer price indices (33). Discounting of costs and effects was not necessary, because the follow-up of the trial was one year (23). An overview of the cost prices used for valuing resource use can be found in Table 1.

Table 1: Price weights used for valuing resource use in the Be Active & Relax VIP study

Resource use categories	Price weight	
	Societal perspective	Employer's perspective
Intervention costs		
Social and physical environmental intervention	€ 427.96[a]	€ 465.92[b]
Social environmental intervention	€ 392.28[a]	€ 430.25[b]
Physical environmental intervention	€ 71.65[a]	€ 71.46[b]
Medical costs		
Visits to a care provider		
General practitioner		
Office consultation	€ 28.96[c]	N.A.
Telephone consultation	€ 14.48[c]	N.A.
House call	€ 44.47[c]	N.A.
Allied health professionals		
Psychologist	€ 82.47[c]	N.A.
Dietician	€ 27.93[c]	N.A.
Physical therapist	€ 37.23[c]	N.A.
Other allied health professionals	Variable[c,d]	N.A.
Medical specialists		
Psychiatrist	€ 106.53[c]	N.A.
Other medical specialists	€ 74.47[c]	N.A.
Complementary medicine	Variable[c,d]	N.A.
Hospitalization		
Ward	€ 472.66[c]	N.A.
Intensive care	€ 2257.82[c]	N.A.
Medications	Variable[e]	N.A.
Absenteeism costs		
Sickness absence	Variable[f]	Variable[f]
Presenteeism costs		
Presenteeism	Variable[f]	Variable[f]
Sports costs	Variable[g]	N.A.
Active commuting equipment costs	Variable[h]	N.A.

Abbreviations: N.A.: Not Applicable
Note: Costs are expressed in 2011 Euros
Cost price sources: [a] Bottum-up micro-costed, valued using tariffs and depleted sources (See Appendix 1); [b] Market prices, valued using invoices; [c] Dutch Manual of Costing; [d] Professional organizations; [e] Dutch Society of Pharmacy; [f] Gross annual salaries of office employees including overhead costs; [g] Self-reported expenses on sports memberships and sports equipment; [h] Self-reported expenses on active commuting equipment

Potential confounders

At baseline, several potential confounders were assessed by questionnaire, including gender (female/male), age (years), having a partner (yes/no), Dutch nationality (yes/no), education level (low=elementary school or less, intermediate=secondary education, and high=college/university), working hours per week, general health (range: 1-5), job demands (range: 1-5), and supervisor support (range: 1-5). Of these, only age and education level were found to be a confounder for both costs and effects. That is, the interventions' effects changed by more than 10% after adding these potential confounders to the crude models.

Statistical analyses

Analyses were performed according to the intention-to-treat principle. Baseline characteristics of intervention and control group participants as well as those of participants with complete and incomplete data were compared using descriptive statistics. Missing data were imputed using multiple imputation. The imputation model included age, gender, number of working days, baseline sickness absence, baseline work performance, baseline effect measure values, and available midpoint and follow-up cost and effect measure values (i.e. 6- and 12 months). Imputations were performed per study group. Using Predictive Mean Matching and Fully Conditional Specification, 15 complete data sets were created in IBM SPSS (v20, Chicago, IL) (Loss of Efficiency ≤5%) (34;35). All datasets were analysed separately as specified below, after which pooled estimates were calculated using Rubin's rules (36). Except for the multiple imputation, analyses were performed using Stata (V12, Stata Corp, College Station, TX). Statistical significance was set at $p<0.05$.

Cost-effectiveness analysis

CEAs with NFR and general vitality were conducted from the societal perspective, in which all costs and consequences related to the interventions were considered, regardless of where they occur. CEAs with NFR and job satisfaction were also conducted from the employer's perspective, in which only costs and consequences relevant to Dutch employers were taken into account.

Effectiveness at 12-month follow-up was analyzed using linear multilevel analyses, adjusted for baseline values and confounders (i.e. age, education level). Three levels were identified: employees (n=412), team leaders (n=49), and departments (n=19). Unadjusted cost differences between study groups were calculated for total as well as disaggregated costs. 95%CIs around these cost differences were estimated by means of bias-corrected (BC) intervals, with 5000 replications. Adjusted cost differences at 12-month follow-up were estimated using linear multilevel analyses (37). These cost differences were corrected for baseline sickness absence, baseline work performance, age, and education level. The 95%CIs around the adjusted cost differences were estimated by means of BC intervals as well (5000 replications). To account for the clustering of data, bootstrap replications were stratified for team leaders (38). Incremental cost-effectiveness ratios (ICERs) were calculated by dividing the adjusted cost differences by those in effects. To graphically illustrate the uncertainty around the ICERs, bootstrapped incremental cost-effect pairs (CE-pairs) were plotted on cost-effectiveness planes (CE-planes) (39). A summary measure of the joint uncertainty of costs and effects was presented using cost-effectiveness acceptability curves (CEACs), which show the probability that each of the interventions is more cost-effective than the others at different ceiling ratios (i.e. the maximum amount of money decision-makers are willing to pay per unit of effect) (40).

ROI analysis

The ROI analysis was performed from the employer's perspective. Costs were defined as intervention costs. Benefits were defined as the difference in monetized outcome measures (i.e. absenteeism and presenteeism costs) between study groups during follow-up, with positive benefits indicating reduced spending. Using linear multilevel analyses, benefits were adjusted for baseline sickness absence, baseline work performance, age, and education level. Subsequently, three ROI-metrics were calculated; 1) Net Benefits (NB), 2) Benefit Cost Ratio (BCR), and 3) Return On Investment (ROI) (21;22).

NB = Benefits – Costs
BCR = Benefits / Costs
*ROI = ((Benefits – Costs)/Costs)*100*

The NB indicates the amount of money gained after costs are recovered (i.e. net-loss or net-savings). The BCR indicates the amount of money returned per Euro invested. The ROI indicates the percentage of profit per Euro invested. To quantify precision, bootstrapped 95%CIs around the NB, BCR, and ROI were estimated using the percentile method, with 5000 replications. Again, bootstrap replications were stratified for team leaders (38). In addition, the probability of financial return was estimated by determining the proportion of bootstrapped financial return estimates was positive (i.e. NB>0, BCR>1, and ROI>0%) (21;22).

Sensitivity analyses

To assess the robustness of the results, five univariate sensitivity analyses were performed. The first sensitivity analysis (SA1) was restricted to participants with complete cost and effect data at all measurement points (i.e. complete-case analysis). In the second sensitivity analysis (SA2), a slightly modified version of the "PROductivity and DISease Questionnaire" (PRODISQ) was used for estimating presenteeism costs, in which presenteeism was conceptualized as reduced work performance due to health complaints (41;42). In the third sensitivity analysis (SA3), absenteeism costs were valued using the HCA for the societal perspective and the FCA for the employer's perspective. As overall consensus about the inclusion of presenteeism costs in economic evaluations does currently not exist, presenteeism costs were excluded in a fourth sensitivity analysis (SA4). Finally, a fifth sensitivity analysis (SA5) was performed, in which absenteeism and presenteeism were valued using age- and gender-specific Dutch price weights (28).

RESULTS

Participants

Of the participants, 92 were allocated to the combined social and physical intervention group, 118 to the social environmental intervention group, 96 to the physical environmental intervention group, and 106 to the control group (Figure 1). At baseline, some meaningful differences were found between study groups in age, education level, and sickness absence days (Table 2). A total of 83 participants (20%) were lost to follow-up (combined: n=29, social: n=20, physical: n=24, control: n=9).

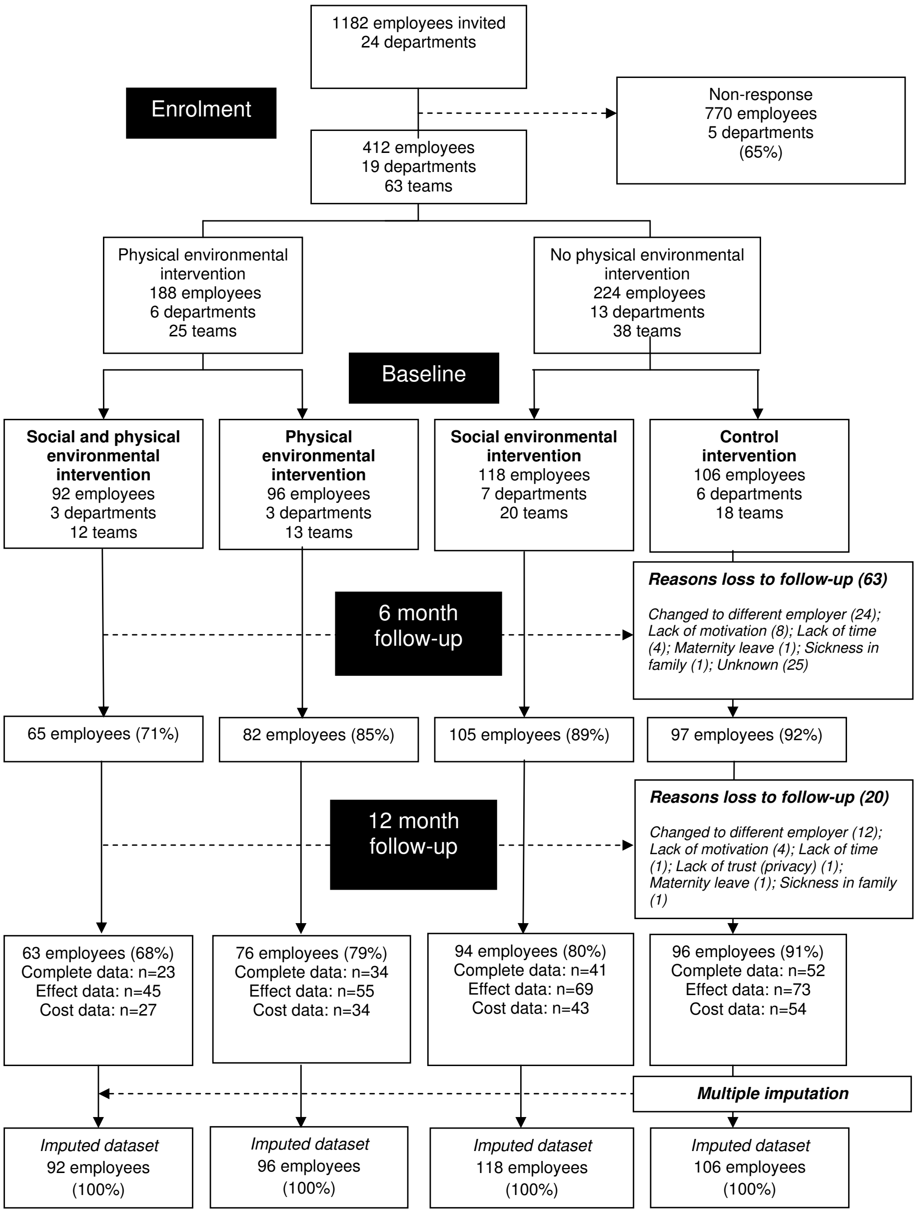

Figure 1: Flow diagram of participants

7

Table 2: Baseline characteristics of participants

Baseline characteristics	Social and physical environmental intervention group			Social environmental intervention group		
	All (n=92)	Complete (n=23)	Incomplete (n=69)	All (n=118)	Complete (n=41)	Incomplete (n=77)
Male [n (%)]	51 (55.4)	17 (73.9)	34 (49.3)	73 (61.9)	25 (61.0)	48 (62.3)
Age (years) [mean (SD)]	38.0 (10.5)	42.8 (9.8)	36.4 (10.3)	43.6 (10.3)	46.8 (9.2)	42.0 (10.5)
Having a partner [n (%)]	74 (80.4)	20 (87.0)	54 (78.3)	91 (77.1)	31 (75.6)	60 (77.9)
Dutch nationality [n (%)]	82 (89.1)	22 (95.7)	60 (87.0)	106 (89.9)	38 (92.7)	68 (88.3)
Education level [n (%)]						
Low	17 (18.5)	3 (13.0)	14 (20.3)	39 (33.1)	14 (34.1)	25 (32.5)
Intermediate	19 (20.7)	4 (17.4)	15 (21.7)	23 (19.5)	8 (19.5)	15 (19.5)
High	55 (59.8)	16 (69.6)	39 (56.5)	56 (47.5)	19 (46.3)	37 (48.1)
Working hours per week [mean (SD)]	35.1 (6.1)	36.0 (5.6)	34.7 (6.5)	36.9 (4.1)	37.1 (4.0)	36.9 (4.2)
General health (range: 1-5) [mean (SD)]	3.8 (0.9)	3.8 (0.8)	3.8 (0.9)	3.8 (0.7)	3.8 (0.8)	3.8 (0.7)
Job demands (range: 1-5) [mean (SD)]	2.6 (0.3)	2.6 (0.2)	2.6 (0.3)	2.7 (0.2)	2.7 (0.2)	2.7 (0.2)
Supervisor support (range: 1-5) [mean (SD)]	2.8 (0.5)	2.7 (0.8)	2.9 (0.4)	2.9 (0.5)	2.8 (0.4)	3.0 (0.5)
Need for recovery (range: 0-100) [mean (SD)]	33.3 (29.9)	21.4 (24.3)	37.2 (30.7)	31.8 (28.7)	25.7 (28.6)	35.1 (28.3)
General vitality (range: 0-6) [mean (SD)]	59.7 (18.0)	65.7 (13.5)	57.6 (18.9)	63.9 (18.3)	66.0 (19.4)	62.8 (17.8)
Job satisfaction (range: 1-5) [mean (SD)]	3.9 (0.7)	3.7 (0.6)	3.9 (0.7)	3.9 (0.8)	3.9 (0.7)	3.9 (0.8)
Sickness absence (days) [mean (SD)]	6.8 (18.5)	9.3 (31.5)	5.8 (9.3)	7.0 (14.2)	11.4 (18.7)	4.5 (10.2)
Work performance (range: 0-10) [mean (SD)]	7.5 (1.0)	7.6 (1.0)	7.5 (1.1)	7.6 (0.8)	7.5 (0.8)	7.7 (0.8)

Physical environmental intervention group			Control group		
All (n=96)	Complete (n=34)	Incomplete (n=62)	All (n=106)	Complete (n=52)	Incomplete (n=54)
60 (62.5)	21 (61.8)	39 (62.9)	65 (61.3)	34 (65.4)	31 (57.4)
42.2 (10.5)	42.7 (10.4)	42.0 (10.7)	40.7 (9.2)	41.3 (9.1)	40.2 (9.3)
82 (85.4)	29 (85.3)	53 (85.5)	85 (80.2)	40 (76.9)	45 (83.3)
87 (90.6)	31 (91.2)	56 (90.3)	95 (89.6)	47 (90.4)	48 (88.9)
16 (16.7)	4 (11.9)	12 (19.4)	21 (19.8)	10 (19.2)	11 (20.4)
20 (20.8)	8 (23.5)	12 (19.4)	24 (22.6)	15 (28.8)	9 (16.7)
60 (62.5)	22 (64.7)	38 (61.3)	61 (57.5)	27 (51.9)	34 (63.0)
35.7 (5.6)	35.1 (5.9)	36.0 (5.4)	36.2 (5.3)	36.1 (6.0)	36.3 (4.6)
3.8 (0.7)	3.8 (0.7)	3.7 (0.7)	3.8 (0.7)	3.7 (0.7)	3.9 (0.6)
2.6 (0.3)	2.7 (0.3)	2.6 (0.3)	2.7 (0.3)	2.7 (0.2)	2.6 (0.3)
2.9 (0.4)	2.9 (31.4)	2.9 (0.4)	2.9 (0.5)	2.9 (0.4)	2.8 (0.6)
33.7 (31.3)	31.2 (31.4)	35.1 (31.5)	30.4 (27.7)	28.9 (27.6)	31.9 (27.9)
63.4 (17.1)	62.2 (19.5)	64.1 (15.9)	66.5 (18.7)	66.4 (19.9)	66.6 (17.6)
4.1 (0.6)	4.1 (0.7)	4.0 (0.5)	4.0 (0.7)	4.0 (0.8)	4.0 (0.6)
11.0 (29.2)	11.2 (39.7)	10.8 (20.7)	3.7 (6.5)	2.9 (4.7)	4.5 (7.9)
7.7 (0.8)	7.9 (0.8)	7.6 (0.8)	7.7 (0.9)	7.8 (1.1)	7.6 (0.8)

The main reasons for loss to follow-up were lack of motivation and changing jobs. After 12 months, complete data were obtained from 59% of participants on the effect measures (combined: 49%, social: 58%, physical: 57%, control: 69%) and from 38% on the cost measures (combined: 29%, social: 36%, physical: 35%, control: 51%). Some meaningful differences were observed between participants with complete and incomplete data (Table 2). These characteristics were included in the imputation model.

Effectiveness

During follow-up, NFR statistically significantly improved among participants of the combined intervention group compared to the control group (-8.4; 95%CI -14.6 to -2.2), whereas this was not the case for the social (0.1; 95%CI -8.8 to 9.0) and physical environmental intervention group (-1.2; 95%CI -9.1 to 6.6). No statistically significant between-group differences were found for general vitality and job satisfaction.

Use of the interventions

During the intervention period, two GMI-training courses for team leaders, 72 GMI-sessions (combined: 24 sessions, social: n 48 sessions), and four GMI-coaching sessions were provided. Also, 19 VIP zones were created; i.e. six VIP Coffee Corner Zones, six VIP Open Office Zones, five VIP Meeting Zones, and two VIP Hall Zones.

Costs

From the societal perspective, the costs of the combined, social environmental, and physical environmental intervention were €428, €392 and €72 per employee, respectively (Appendix 1). From the employer's perspective, these costs were €466 (combined), €430 (social), and €72 (physical). Active commuting equipment costs were statistically significantly lower in all intervention groups as compared to the control group. Moreover, combined intervention group participants had statistically significantly lower sports costs than their control group counterparts. All other disaggregate cost differences were not statistically significant (Table 3). Total employer's costs in the combined intervention group were statistically significantly higher than in the control group (3102; 95%CI 598 to 5969), but this was not the case for both the social and physical environmental intervention group. Societal costs in all intervention groups were higher than in the control group, but these differences were not statistically significant (Table 4).

Table 3: Mean costs per participant in the intervention and control groups, and unadjusted mean cost differences between study groups during the 12-month follow-up period (based on the imputed dataset)

Cost category	Control group n=106; mean (SEM)	Social and physical environmental group n=92; mean (SEM)	Δ Costs (95%CI)	Social environmental group n=118; mean (SEM)	Δ Costs (95%CI)	Physical environmental group n=96; mean (SEM)	Δ Costs (95%CI)
				Societal perspective			
Intervention costs	0 (NA)	428 (NA)	428 (NA)	392 (NA)	392 (NA)	72 (NA)	72 (NA)
Healthcare costs	602 (116)	502 (74)	-99 (-276 to 95)	742 (114)	140 (-56 to 358)	741 (183)	139 (-88 to 406)
Absenteeism costs	2345 (715)	3729 (1146)	1384 (-774 to 3737)	2901 (665)	556 (-1312 to 2201)	2691 (677)	347 (-1394 to 1817)
Presenteeism costs	18068 (919)	19289 (1579)	1221 (-1028 to 3416)	17880 (923)	-188 (-1864 to 1507)	19284 (1276)	1216 (-638 to 3102)
Sports costs	799 (124)	372 (39)	-427 (-630 to -245)	635 (116)	-164 (-388 to 58)	184 (64)	-109 (-331 to 107)
Active commuting equipment costs	589 (149)	11 (5)	-578 (-755 to -428)	291 (124)	-298 (-516 to -80)	690 (125)	-405 (-595 to -214)
Total	**22402 (1155)**	**24331 (1626)**	**1929 (-846 to 4624)**	**22841 (1088)**	**438 (-1883 to 2739)**	**237663 (1464)**	**1260 (-1176 to 3634)**
				Employer's perspective			
Intervention costs	0 (NA)	466 (NA)	466 (NA)	430 (NA)	430 (NA)	72 (NA)	72 (NA)
Absenteeism costs	2520 (843)	4410 (1507)	1889 (-803 to 4964)	3073 (778)	553 (-1646 to 2536)	2691 (677)	171 (-1862 to 1752)
Presenteeism costs	18068 (919)	19289 (1579)	1221 (-1028 to 3415)	17880 (923)	-188 (1864 to 1507)	19284 (1276)	1216 (-638 to 3102)
Total	**20588 (1189)**	**24164 (1783)**	**3577 (510 to 6699)**	**21384 (1141)**	**796 (-1715 to 3179)**	**22047 (1388)**	**1459 (-1056 to 3830)**

Abbreviations: n: number; SEM: Standard Error of the Mean, CI: Confidence Interval, NA: Not Applicable
Note: Costs are expressed in 2011 Euros

7

Table 4: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes for the main analysis

Perspective	Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
			Social and physical environmental intervention							
	Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
Societal perspective	92	106	Need for recovery	1647 (-653 to 4257)	-8.4 (-14.6 to -2.2)	-197	81.4	18.2	0.0	0.4
	92	106	General Vitality	1647 (-653 to 4257)	3.4 (-9.3 to 16.2)	479	56.0	16.2	2.0	25.7
Employer's perspective	92	106	Need for recovery	3102 (598 to 5969)	-8.4 (-14.6 to -2.2)	-370	94.9	0.5	0.0	0.4
	92	106	Job satisfaction	3102 (598 to 5969)	-0.1 (-0.6 to 0.5)	-49595	36.2	3.7	1.0	59.2
			Social environmental intervention							
	Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
Societal perspective	118	106	Need for recovery	187 (-1895 to 2253)	0.1 (-8.8 to -9.0)	1784	22.6	26.7	19.0	31.7
	118	106	General Vitality	187 (-1895 to 2253)	7.3 (-3.8 to 18.4)	26	44.5	43.8	2.4	9.3
Employer's perspective	118	106	Need for recovery	447 (-1609 to 2472)	0.1 (-8.8 to 9.0)	4256	26.0	23.3	16.3	34.4
	118	106	Job satisfaction	447 (-1609 to 2472)	-0.2 (-0.9 to 0.5)	-2004	14.8	6.0	33.6	45.5
			Physical environmental intervention							
	Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
Societal perspective	96	106	Need for recovery	479 (-1757 to 2779)	-1.2 (-9.1 to -6.6)	-382	46.7	27.7	12.4	13.3
	96	106	General Vitality	479 (-1757 to 2779)	5.7 (-5.6 to 17.1)	84	45.7	37.1	3.0	14.2
Employer's perspective	96	106	Need for recovery	922 (-1384 to 3156)	-1.2 (-9.1 to 6.6)	-763	52.3	21.1	9.6	17.1
	96	106	Job satisfaction	922 (-1384 to 3156)	-0.1 (-0.5 to 0.4)	-17846	24.9	22.9	7.8	44.5

Abbreviations: CI: Confidence Interval, C: Costs, E: Effects, ICER: Incremental Cost-Effectiveness Ratio, CE-plane: Cost-Effectiveness plane
Note: Costs are expressed in 2011 Euros
[1] Refers to the northeast quadrant of the CE-plane, suggesting that the intervention is more effective and more costly compared to usual practice
[2] Refers to the southeast quadrant of the CE-plane, suggesting that the intervention is more effective and less costly compared to usual practice
[3] Refers to the northwest quadrant of the CE-plane, suggesting that the intervention is less effective and more costly compared to usual practice
[4] Refers to the southwest quadrant of the CE-plane, suggesting that the intervention is less effective and less costly compared to usual practice

Societal perspective: cost-effectiveness

For NFR, an ICER of -197 was found for the combined intervention group in comparison with the control group. This indicates that for every 1-point improvement in NFR, the intervention costs €197 in comparison with usual practice (Note that this ICER is negative, as lower scores indicate a better NFR). An ICER in the similar direction was found for the physical environmental intervention group (ICER: -382). In both cases, the majority of incremental CE-pairs were located in the northeast quadrant of the CE-plane (Table 4, Figure 2-1a), suggesting that both intervention conditions were more costly and more effective in improving NFR than usual practice. For the social environmental intervention group, an ICER of 1784 was found. This indicates that the intervention costs €1784 per point decline in NFR in comparison with usual practice (Table 4, Figure 2-1a). For general vitality, ICERs of 479, 26, and 84 were found for the combined, social, and physical environmental intervention group, respectively. In all cases, the majority of incremental CE-pairs were located in the northeast quadrant of the CE-plane (Table 4, Figure 2-1b). This suggests that all intervention conditions were more costly and more effective in improving general vitality than usual practice.

A summary measure of the joint uncertainty of costs and effects is presented by the CEACs in Figure 2-2a and Figure 2-2b. These CEACs indicate that the probabilities of cost-effectiveness of both separate intervention conditions and usual practice were about 0.3 at societal willingness-to-pay values of €0/point improvement in NFR and general vitality, while that of the combined intervention was lower (i.e. 0.09). For NFR, the separate intervention conditions' probabilities of cost-effectiveness as well as that of usual practice decreased with an increasing willingness-to-pay, while that of the combined intervention gradually increased to 0.85 at a ceiling ratio of €3900. For general vitality, on the other hand, the social environmental intervention condition's probability of cost-effectiveness increased with an increasing willingness-to-pay, whereas that of all other intervention conditions remained about the same.

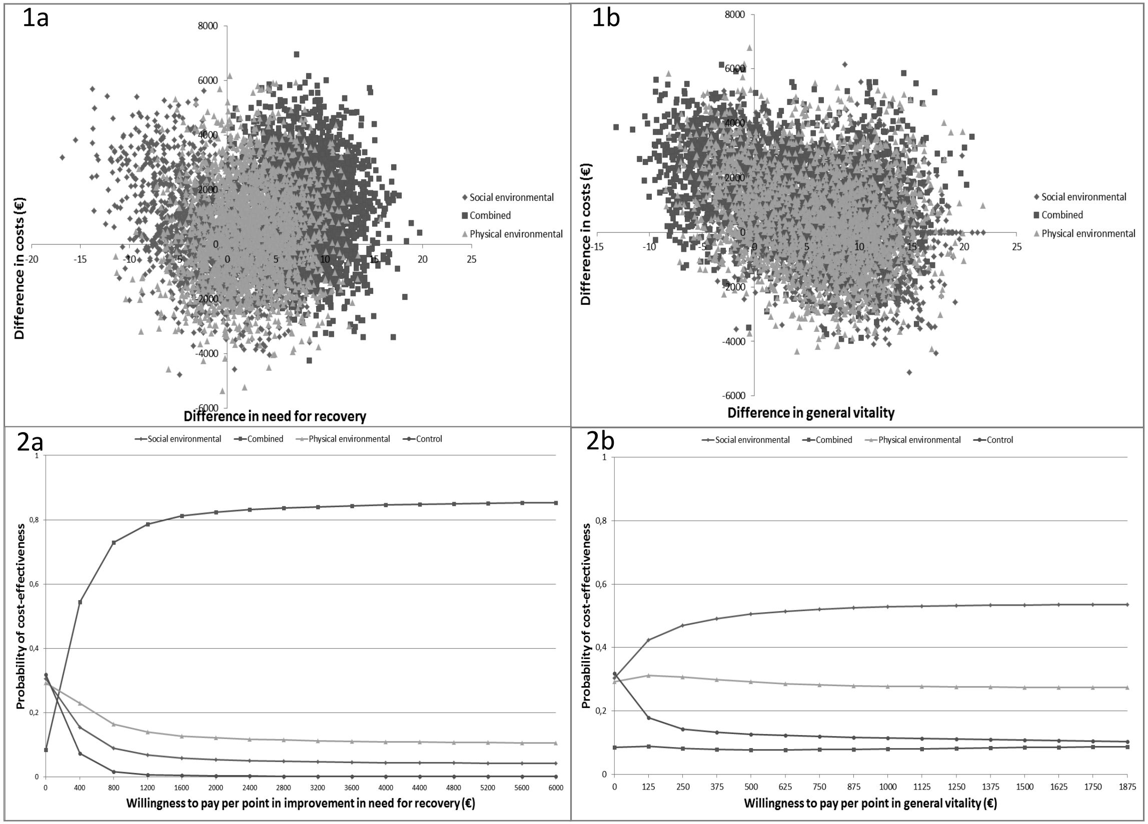

Figure 2: Cost-effectiveness planes indicating the uncertainty around the incremental cost-effectiveness ratios in comparison with usual practice (1) and cost-effectiveness acceptability curves indicating the probability of cost-effectiveness for different values (€) of willingness to pay per unit of effect gained (2) for need for recovery (a) and general vitality (b) (societal perspective, based on the imputed dataset)

Employer's perspective: cost-effectiveness

For NFR, an ICER of -370 was found for the combined intervention group, suggesting that the intervention was associated with an additional cost to the employer of €370 per point improvement in NFR in comparison with usual practice. For the physical environmental intervention group, an ICER in the similar direction was found (ICER: -763). In both cases, the majority of incremental CE-pairs were located in the northeast quadrant of the CE-plane (Table 4, Figure 3-1a). For the social environmental intervention group, on the other hand, an ICER of 4256 was found. This indicates that the intervention costs €4256 to the employer in comparison with usual practice per point decline in NFR. For job satisfaction, ICERs of -49595, -2004, and -17846 were found for the combined, social, and physical environmental intervention group, respectively (Table 4). All of these intervention conditions were more costly and less effective than usual practice (Table 4, Figure 3-1b).

The CEACs presented in Figure 3-2a and Figure 3-2b indicate that the probabilities of cost-effectiveness of all intervention conditions were lower than that of usual practice at employer's willingness-to-pay values of €0/point improvement in NFR and job satisfaction. For NFR, the separate intervention conditions' probabilities of cost-effectiveness as well as that of usual practice decreased with an increasing willingness-to-pay, while that of the combined intervention increased to 0.85 at a ceiling ratio of €6000. For job satisfaction, on the other hand, the probability of cost-effectiveness of all intervention conditions remained lower than that of usual practice, irrespective of the employer's willingness-to-pay (Figure 3-2b).

Employer's perspective: financial return

During follow-up, total employer's benefits, NBs, BCRs, and ROIs were negative for all intervention conditions, indicating that the investments were larger than the benefits (Table 5). Moreover, the probabilities of financial return of all intervention conditions were low (combined: 0.05, social: 0.41, environmental: 0.30).

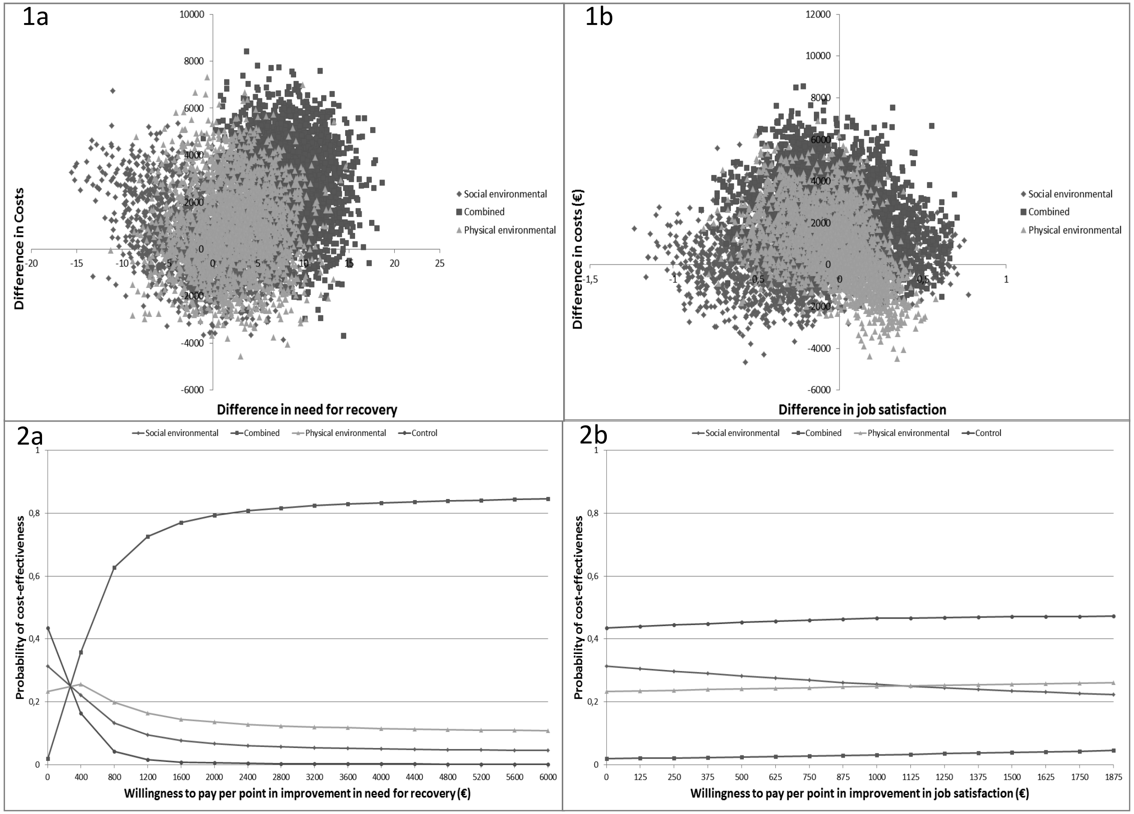

Figure 3: Cost-effectiveness planes indicating the uncertainty around the incremental cost-effectiveness ratios in comparison with usual practice (1) and cost-effectiveness acceptability curves indicating the probability of cost-effectiveness for different values (€) of willingness to pay per unit of effect gained (2) for need for recovery (a) and job satisfaction (b) (employer's perspective, based on the imputed dataset)

Table 5: Intervention costs, benefits, Net Benefits (NB), Benefit Cost Ratio (BCR), and Return-On-Investment (ROI) per participant for the main analyses

Intervention group	Sample size		Costs (€)	Benefits	Financial return			
	Intervention	Control	Total (95% CI)	Total (95% CI)	NB[1] (95% CI)	BCR[2] (95% CI)	ROI (%)[3] (95% CI)	Probability
Social and physical environmental intervention	92	106	466 (NA)	-2636 (-5724 to 90)	-3102 (-5897 to -93)	-5.7 (-11.7 to 0.8)	-666 (-1266 to -20)	0.05
Social environmental intervention	118	106	430 (NA)	-16 (-2335 to 2198)	-447 (-2715 to 1809)	-0.0 (-5.3 to 5.2)	-104 (-631 to 421)	0.41
Physical environmental intervention	96	106	72 (NA)	-850 (-3113 to 1834)	-922 (-4703 to 2466)	-11.8 (-65 to 35)	-1286 (-6564 to 3442)	0.30

Abbreviations: CI: Confidence Interval, NB: Net Benefit, BCR: Benefit Cost Ratio, ROI: Return-On-Investment, I: Intervention, C: Control: NA: Not Applicable
Note: Costs are expressed in 2011 Euros
Note: Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0
[1] Indicates the amount of money returned after intervention costs are recovered
[2] Indicates the amount of money returned per Euro invested in the intervention
[3] Indicates the percentage of profit per Euro invested in the intervention

7

Sensitivity analyses

Results of SA3 were similar to those of the main analysis. The outcomes of SA1 (complete-cases), SA2 (PRODISQ), SA4 (excluding presenteeism), and SA5 (age- and gender-specific price weights) differed in some aspects from the main analysis (Appendix 2, Appendix 3). Four differences stand out. First, in the main analysis, NFR statistically significantly decreased among combined intervention group participants in comparison with the control group, whereas the difference in NFR scores between both groups was not statistically significant among the complete-cases (SA1). Second, in the main analysis, total societal and employer's costs were higher among participants to all intervention groups in comparison with the control group, whereas they were lower when using a sightly modified version of the PRODISQ (SA2). Third, the probability of financial return was low for all intervention conditions in the main analysis, whereas that of the combined intervention condition (0.90) and that of the physical environmental intervention condition (0.93) were relatively high when using the PRODISQ (SA2). Fourth, in the main analysis, total employer's costs were statistically significantly higher among participants to the combined intervention in comparison with the control group, whereas this difference was not statistically significant when presenteeism costs were excluded (SA4).

DISCUSSION

This study aimed to evaluate the cost-effectiveness and financial return of a combined social and physical environmental intervention in office employees in comparison with usual practice, and of both intervention conditions separately. Additionally, the probabilities of the interventions being cost-effective in comparison with each other were explored. The combined intervention statistically significantly improved NFR in comparison with usual practice, whereas both separate intervention conditions did not. No statistically significant between-group differences were found for general vitality and job satisfaction. Employer's costs were statistically significantly higher in the combined intervention group compared with the control group, whereas all other societal and employer's cost differences were not statistically significant. Whether the combined intervention can be regarded as cost-effective in improving

NFR from both the societal and employer's perspective depends on the respective decision-makers' willingness-to-pay per point improvement as well as the probability of cost-effectiveness that they consider acceptable. However, as both are currently unknown, strong conclusions cannot be made. Nonetheless, societal and company decision-makers can use the present results to consider whether they perceive that the intervention provides "good value for money" at an acceptable probability of cost-effectiveness. Both separate intervention conditions, on the other hand, cannot be regarded as cost-effective in improving NFR, because their probabilities of cost-effectiveness in comparison with the other study groups were low, regardless of the maximum willingness-to-pay. None of the intervention conditions seemed to be cost-effective in improving general vitality from the societal perspective, nor in improving job satisfaction from that of the employer. Moreover, the probability of financial return was low for all intervention conditions, indicating that none of them generated cost savings to the employer.

Comparison with existing literature

Until now, few studies evaluated the effectiveness of comparable interventions in improving NFR, general vitality, and/or job satisfaction. Meijer et al. (2009), evaluated the effect of a so-called innovative office concept (e.g. open-office plan, flexible workplaces) on NFR among Dutch office employees. No significant improvements were found at 15 months follow-up (43). Their study, however, did not include a comparison group and the content of their intervention differed from the intervention conditions evaluated here. Using an RCT, Strijk et al. (2012) evaluated the effectiveness of a worksite vitality intervention aimed at improving physical activity, nutrition, and relaxation among older Dutch hospital employees versus usual practice. Even though the intervention statistically significantly improved NFR at 6-month follow-up, this effect was not sustained at the long-term (44;45). Moreover, the intervention did not improve general vitality at 6- and 12-month follow-up (46). Again, however, the content of the worksite vitality intervention (i.e. yoga and aerobic exercising, fruit, and individual counselling) differed from that of the present intervention conditions and the intervention was not specifically targeted at office employees. This study also evaluated the societal cost-effectiveness of the worksite vitality intervention

in improving NFR and general vitality. The intervention was not considered cost-effective in improving both outcomes, because a substantial amount of money had to be paid by society to reach a reasonable probability of cost-effectiveness (45).
To our knowledge, studies evaluating the employer's cost-effectiveness of comparable interventions in improving NFR are lacking. One study, however, evaluated the employer's cost-effectiveness in improving job satisfaction of a mindfulness-based worksite health promotion program in comparison with usual practice (41). Irrespective of the maximum willingness-to-pay, the intervention had a low probability of cost-effectiveness (i.e. ≤0.25) and was therefore not considered to be cost-effective either.
A systematic review of the financial return of worksite physical activity and/or nutrition programs indicated that such programs may generate positive financial returns through reduced absenteeism costs according to non-randomized studies (BCR: 4.25), whereas they do not according to RCTs (BCR: 0.51) (47). When we solely included absenteeism costs in SA4, our results were in line with those of the review (BCR-combined: -2.2, BCR-social: 0.0, BCR-physical: 0.5). Moreover, a recent review of U.S. worksite health promotion studies published after 2000 found that only one of the seven studies showing cost savings utilized a randomized design. Based on these findings, the authors concluded that strong evidence of cost savings of worksite health promotion programs is currently lacking (48).

Explanation of findings

The finding that the combined intervention statistically significantly improved NFR in comparison with usual practice, whereas both separate intervention conditions did not, is in line with our hypothesis that the combined intervention would be most effective. Moreover, it is noteworthy that even though the combined intervention had a statistically significant positive effect on NFR, total employer's costs were statistically significantly higher among combined intervention group participants compared to their control group counterparts. This is striking, as absenteeism costs accounted for more than half of the difference in total employer's costs, while improvements in NFR were previously found to be related to lower absenteeism costs (49). Our finding might have resulted from the fact that worksite health promotion programs, such as ours, may positively affect NFR at the short-term, while the related improvements in productivity occur at the long-term. However, further research is needed to confirm this. Furthermore, even though the intervention (conditions) were aimed at improving physical activity and active commuting, sports costs and active commuting equipment costs were lower in all intervention groups as compared with the control group. A possible explanation for this finding may be that control group participants were aware of the content and/or aims of the intervention conditions, and purchased sports memberships, sports equipment, and/or bicycles in an effort to compensate for the fact that they solely received usual practice.

Robustness of the results

Results of the sensitivity analyses differed in some aspects from those of the main analysis. Most notably, the combined intervention's effect on NFR was statistically significant in the main analysis (for which data were imputed), whereas this was not the case when participants with missing data were eliminated from the analyses. This probably resulted from the large difference in baseline NFR scores between combined intervention group participants with complete (mean: 33.3) and incomplete data (mean: 21.4). This indicates that the complete-case analysis is likely to be biased by self-selection of participants. Moreover, when presenteeism costs were estimated using a slightly modified version of the PRODISQ, the results were much more favourable than those of the main analysis (for which the WHO-HPQ was used). Both instruments likely produced different results, because they conceptualize presenteeism in a slightly different way (WHO-HPQ: reduced overall work performance, PRODISQ: reduced work performance due to health complaints). The WHO-HPQ was used in the main analysis, because worksite health promotion programs are not just hypothesized to indirectly affect presenteeism through individual health improvements, but also directly from program impact (50).

Strengths and limitations

This study has several strengths. First, this study was the first to evaluate the cost-effectiveness and financial return of a combined social and physical environmental intervention, as well as that of both intervention conditions separately. Second,

the use of randomization for allocating departments to the "social environmental intervention" and "no social environmental intervention" group reduced the possible influence of selection bias, while the study's external validity was improved by its pragmatic design. Third, to minimize contamination between study groups, group allocation was performed at the department-level. Moreover, to account for the possible clustering of data resulting from this design, this study was one of the first to use linear multilevel analyses for assessing the intervention conditions' cost-effectiveness and financial return. The latter is of great importance, as most economic evaluations alongside clustered studies ignore the possible clustering of data, whereas those that do seem to underestimate the statistical uncertainty and are likely to have inaccurate point estimates (51;52).
The study also had some limitations. First of all, the generalizability of the present findings to other companies, work settings, and/or the general working population may be hampered by the fact that the study was performed among office workers within a single company. Another limitation concerns the relatively large amount of missing data: i.e. 41% of participants had some missing effect data and 62% had some missing cost data. Even though missing data are generally inevitable in trial-based economic evaluations and multiple imputation techniques were used for filling in missing values, a 100% compete dataset would have produced more valid and reliable results. Therefore, the present results should be treated with caution and extensive efforts ought to be made in future studies to reduce the amount of missing data. Moreover, all effect measures and some resource use categories were assessed using retrospective questionnaires. This may have induced "recall bias". Nonetheless, as it is seems highly unlikely that the extent of impairment in recall systematically differed between study groups, we do not expect that our use of such questionnaires severely biased our results (53).

Conclusion

Depending on the societal and employer's willingness-to-pay and the probability of cost-effectiveness that they consider acceptable, the combined intervention may be considered cost-effective in improving NFR. Both separate interventions were not cost-effective in improving this outcome. Moreover, all interventions were neither cost-effective in improving general vitality (societal perspective) and job satisfaction (employer's perspective), nor cost saving to the employer.

Acknowledgements
This project is part of a research program called "Vitality In Practice", which is funded by Fonds Nuts Ohra (Nuts Ohra Foundation). The authors wish to thank Ruben Kraaijeveld and Robine van der Starre for their help with the data collection. The authors would also like to thank all participants and intervention providers.

REFERENCES

1. World Health Organization. Mental health and well-being at the workplace - protection and inclusion in challenging times. Copenhagen; 2010.
2. Koppes LLJ, de Vroome EEM, Mol MEM, Janssen BJM, van Zwieten MHJ, van de Bossche SNJ. Nationale enquete arbeidsomstandigheden [Netherlands Working Conditions Survey] 2011. http://www.tno.nl/downloads/rapport_nea_20111.pdf
3. European working conditions observatory. Work-related stress. http://www.eurofound.europa.eu/ewco/studies/tn1004059s/tn1004059s.htm
4. Michie S. Causes and management of stress at work. *Occup Environ Med* 2002;59(1):67-72.
5. Rosch PJ. The quandary of job stress compensation. *Health and Stress* 2001;3:1-4.
6. Tennant C. Life events, stress and depression: a review of recent findings. *Australian and New Zealand Journal of Psychiatry* 2002;36(2):173-182.
7. Geurts SAE, Sonnentag S. Recovery as an explanatory mechanism in the relation between acute stress reactions and chronic health impairement. *Scand J Work Environ Health* 2006;32(6):482-492.
8. Sluiter JK, de Croon EM, Meijman TF, Frings-Dresen MH. Need for recovery from work related fatigue and its role in the development and prediction of subjective health complaints. *Occup Environ Med* 2003;60(Suppl 1):i62-i70.
9. Sluiter JK, Frings-Dresen MHW, van der Beek AJ, Meijman TF. The relation between work-induced neuroendocrine reactivity and recovery, subjective need for recovery, and health status. *J Psychosom Res* 2001;50(1):29-37.
10. van Amelsvoort LG, Kant IJ, Bultmann U, Swaen GM. Need for recovery after work and the subsequent risk of cardiovascular disease in a working population. *Occup Environ Med* 2003;60(Suppl 1):i83-i87.
11. Rook JW, Zijlstra FRH. The contribution of various types of activities to recovery. *European Journal of Work and Organizational Psychology* 2006;15:218-240.
12. Siltaloppi M, Kinnunen U, Feldt T, Tolvanen A. Development of need for recovery from work over one year: a person-centered approach. *Anxiety Stress Coping* 2012;25:23-42.
13. Sonnentag S. Work, recovery activities, and individual well-being: A diary study. *J Occup Health Psychol* 2001;6:196-210.
14. Sonnentag S, Natter E. Flight attendants' daily recovery from work: Is there no place like home? *Int J Stress Manag* 2004;11:366-391.
15. Coffeng JK, Hendriksen IJ, Duijts SF, Proper KI, van Mechelen W, Boot CR. The development of the Be Active & Relax "Vitality in Practice" (VIP) project and design of an RCT to reduce the need for recovery in office employees. *BMC Public Health* 2012;12(1):592.
16. McLaren L, Hawe P. Ecological perspectives in health research. J Epidemiol Community Health 2005;59(1):6-14.
17. Richard L, Gauvin L, Raine K. Ecological Models Revisited: Their Uses and Evolution in Health Promotion Over Two Decades. Annu Rev Public Health 2011;32(1):307-326.

18. Coffeng JK, Hendriksen IJM, Twisk JWR, Van Mechelen W & Boot CRL. Effectiveness of a social and physical environmental intervention on work-related outcomes in office employees. Submitted
19. Coffeng JK, Boot CRL, Twisk JWR, Van Mechelen W & Hendriksen IJM. A combined social and physical environmental intervention at work reduces exhaustion, and increases frequency of small breaks and active commuting in office employees: results from a randomised controlled trial. Submitted
20. van Dongen JM, Tompa E, Clune LA, Sarnocinska-Hart A, Bongers PM, van Tulder MW, et al. Bridging the gap between the economic evaluation literature and daily practice in health and safety: a study into the information needs of decision makers in the healthcare sector. *Implement Sci* 2013;8:57
21. Stone PW. Return-on-investment models. *Appl Nurs Res* 2005;18(3):186-189.
22. Cavallo D. Using return on investment analysis to evaluate health promotion programs: challenges and opportunities. *Health Promotion Economics Issue Briefs* 2006;1(3):1-4. RTI-UNC Center of Excellence. Available at: http://www.rti.org/pubs/IssueBrief_3.pdf.
23. Drummond MF, Sculpher MJ, Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.
63. Krol M. Productivity costs in economic evaluations Erasmus. Thesis (PhD), Universiteit Rotterdam, 2012.
24. van Veldhoven M, Broersen S. Measurement quality and validity of the "need for recovery scale". *Occup Environ Med* 2003;60:3-9.
25. van der Zee KI, Sanderman R. *Het Meten van Gezondheidstoestand Met de RAND-36: een Handleiding*. Groningen, the Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 1993.
26. Koppes LLJ, de Vroome EMM, Mol MEM, Janssen BJM, van den Bossche SNJ. *Nationale enquête arbeidsomstandigheden 2010: Methodologie en globale resultaten. 2010.* http://www.tno.nl/downloads/rapport_nea_20111.pdf
27. Frick FD. Microcosting Quantity Data Collection Methods. *Med Care* 2009;47(7 Suppl 1):S76-S81.
28. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. *Handleiding Voor Kostenonderzoek. Methoden en Standaardkostprijzen Voor Economische Evaluaties in de Gezondheidszorg*. Diemen, the Netherlands: College Voor Zorgverzekeringen; 2010.
29. Z-index. G-Standard. The Hague, The Netherlands: Z-Index BV; 2009.
30. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;14(2):171-189.
31. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D. The world health organization health and work performance questionnaire (HPQ). *J Occup Environ Med* 2003;45:156-174.
32. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE. Using the world health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46:S23-S37.
33. Statistics Netherlands. 2011. http://www.cbs.nl
34. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res* 2007;16(3):219-242.

35. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statist Med* 2011;30(4):377-399.

36. Rubin DB. *Multiple imputation for nonresponse in surveys.* John Wiley & Sons: New York, 1987.

37. Gomes M, Grieve R, Nixon R, Ng ES, Carpenter J, Thompson SG. Methods for covariate adjustments in cost-effectiveness analysis that use cluster randomised trials. *Health Econ* 2012;21(9):1101-1118.

38. van der Leeden R, Meijer E, Busing FMTA. *Resampling multilevel models.* In: de Leeuw J, Meijer E, editors. Handbook of Multilevel Analysis. Springer Science: New York, NY, 2008.

39. Black WC. The CE plane: a graphic representation of cost-effectiveness. *Med Decis Making* 1990;10(3):212-214.

40. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. *Health Econ* 2004;13(5):405-415.

41. van Dongen JM, van Berkel J, Boot CRL, Bosmans JE, Proper KI, Bonges PM, van der Beek AJ, van Tulder MW, van Wier MF. Cost-effectiveness and financial return of a mindfulness-based worksite intervention aimed at improving work engagement: results of a randomized controlled trial. Submitteda

42. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5:23-28.

43. Meijer EM, Frings-Dresen MHW, Sluiter JK. Effects of office innovation on office workers' health and performance. *Ergonomics* 2009;52(9):1027-1038.

44. Strijk JE, Proper KI, van der Beek AJ, van Mechelen W. A worksite vitality intervention to improve older workers' lifestyle and vitality-related outcomes: results of a randomised controlled trial. *J Epidemiol Community Health* 2012;66(11):1071-1078.

45. van Dongen JM, Strijk JE, Proper KI, van Wier MF, van Mechelen W, van Tulder MW, et al. A cost-effectiveness and return-on-investment analysis of a worksite vitality intervention among older hospital workers: results of a randomized controlled trial. *J Occup Environ Med* 2013;55(3):337-346.

46. Strijk JE, Proper KI, van Mechelen W, van der Beek AJ. A worksite vitality intervention for older hospital workers to improve vitality, work engagement, productivity and sick leave: results of a randomized controlled trial. *Scand J Work Environ Health* 2013;39(1):66-75

47. de Croon EM, Sluiter JK, Frings-Dresen MHW. Need for recovery after work predicts sickness absence: A 2-year prospective cohort study in truck drivers. *J Psychosom Res* 2003;55(4):331-339.

48. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011;12(12):1031-1049.

49. Osilla KC, van Busum K, Schnyer C, Larkin JW, Eibner C, Mattke C. Systematic review of the impact of worksite wellness programs. *Am J Manag Care* 2012;18(2):e68-e81.

50. Anderson DR, Sexner SA, Gold DB. Conceptual framework, critical questions, and practical challenges in conducting research on the financial impact of worksite health promotion. *Am J Health Promot* 2001;15(5):281-288.

51. Grieve R, Nixon R, Thompson SG. Bayesian hierachical models for cost-effectiveness analyses that use data from cluster randomized trials. *Med Decis Making* 2010;30(2):163-175.
52. Gomes M, Grieve R, Nixon R, Edmunds WJ. Statistical Methods for Cost-Effectiveness Analyses That Use Data from Cluster Randomized Trials: A Systematic Review and Checklist for Critical Appraisal. *Med Decis Making* 2012;32(1):209-220.
53. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol* 1990;43(1):87-91.

Appendix 1: Cost of the Be Active & Relax VIP interventions (societal perspective)

	Cost categories	Description	Units	Unit Prices	Total Costs (Euros 2011)	Mean costs per employee (Euros 2011)
			Intervention development			
Development costs [a]					€ 108,793.35	€ 3.31
					Subtotal:	**€ 3.31**
			Intervention kick-off			
Management / teamleader kick-off meetings	Labor costs	Intervention staff	16.5 hours	€ 31.43 / hour	€ 518.66	€ 1.26
		Intervention staff	3 hours	€ 117.74 / hour	€ 353.23	€ 0.86
		Intervention staff	2 hours	€ 64.03 / hour	€ 128.03	€ 0.31
		Managers	54.5 hours	€ 78.44 / hour	€ 4,275.04	€ 10.38
		Teamleaders	69 hours	€ 62.58 / hour	€ 4,317.92	€ 10.48
	Capital costs		16.5 hours	variable	€ 346.70	€ 0.84
	Material costs	Recruitment materials		variable	€ 3,518.42	€ 8.54
					Subtotal:	**€ 32.67**
			Social environmental			
			Group motivational interviewing (GMI)			
GMI-training	Labor costs	GMI-professional	36 hours	€ 49.12 / hour	€ 1768,20	€ 8.42
		Intervention staff	36 hours	€ 31.43 / hour	€ 1131,62	€ 5.39
		Teamleaders	318.5 hours	€ 62.58 / hour	€ 19,931.29	€ 94.91
	Capital costs		28 hours	€ 8.55 / hour	€ 239.40	€ 1.14
	Material costs	GMI-workbooks (teamleaders)	25 workbooks	€ 13.49 / workbook	€ 337.17	€ 1.61
		Lunch	50 lunches	€ 9.60 / lunch	€ 480.00	€ 2.29
GMI-coaching	Labor costs	GMI-professional	4 hours	€ 49.12 / hour	€ 196.47	€ 0.94
		Teamleaders	21 hours	€ 62.58 / hour	€ 1,314.15	€ 6.26
	Capital costs		4 hours	€ 8.55 / hour	€ 34.20	€ 0.16
GMI-sessions	Labor costs	Teamleaders	108 hours	€ 62.58 / hour	€ 6,758.49	€ 32.18
		Employees	746.8 hours	€ 48.22 / hour	€ 36,015.93	€ 171.50
	Capital costs		108 hours	€ 8.55 / hour	€ 923.40	€ 4.40
	Material costs	GMI-workbooks (employees)	185 workbooks	€ 5.95 / workbook	€ 1,100.75	€ 5.24
		Other materials		variable	€ 3,920.38	€ 18.67
					Subtotal:	**€ 353.10**
			Social media platform			
Website hosting					€ 672.92	€ 3.20
					Subtotal:	**€ 3.20**

	Cost categories	Description	Total purchase costs (Euros 2011)	Total annual depreciaton Costs (Euros 2011)	Mean costs per employee (Euros 2011)
		Physical environmental			
VIP zones	Material costs	Giant wall poster	€ 847.12	€ 169.42	€ 0.90
		Bar / standing tables	€ 2,632.00	€ 526.40	€ 2.80
		Bar chairs	€ 2877.00	€ 575.40	€ 3.06
		Plants	€ 2137.32	€ 427.46	€ **2.27**
		Exercise balls	€ 1844.77	€ 368.95	€ 1.96
		Curtains	€ 659.73	€ 131.95	€ 0.70
		Table tennis	€ 607.50	€ 120.50	€ 0.65
		Lounge chairs	€ 2000.00	€ 400.00	€ 2.13
		Footsteps	€ 880.60	€ 176.12	€ 0.94
		VIP banners	€ 1071.00	€ 214.20	€ 1.14
		Other / office rebuilding	€ 17,977.55	€ 3,595.51	€ 19.13
				Subtotal:	**€35.86**
TOTAL COST OF THE SOCIAL ENVIRONMENTAL INTERVENTION					**€ 392.28**
TOTAL COST OF THE PHYSICAL ENVIRONMENTAL INTERVENTION					**€ 71.65**
TOTAL COST OF THE SOCIAL AND PHYSICAL ENVIRONMENTAL INTERVENTION					**€ 427.96**

[a] € 108,793.35 was paid for the development of the Be Active & Relax intervention. For calculating the development costs per participant, these were divided by the expected number of program users during the first five years after implementation (17,557). Workers in the business and financial service sector make up 19.5% (1461000 workers) of the total Dutch workforce (http://www.cbs.nl). During the five year period, it was hypothesized that the intervention will be offered to 5% of these workers and that 45% of them will participate to the intervention (35% during the first year after implementation and 2.5% during each consecutive year).

Appendix 2: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes for the sensitivity analyses

Sensitivity analysis	Perspective	Sample size		Outcome	ΔC (95% CI)	ΔE (95% CI)	ICER	Distribution CE-plane (%)			
Social and physical environmental intervention											
		Intervention	Control		€	Points	€/point	NE[1]	SE[2]	SW[3]	NW[4]
SA1 – complete cases	Societal perspective	23	52	Need for recovery	2801 (862 to 6119)	-0.5 (-6.5 to 5.6)	-5921	53.0	5.3	7.2	41.0
		23	52	General Vitality	2801 (862 to 6119)	1.4 (-3.0 to 5.8)	2008	71.3	22.7	7.2	5.3
SA1 – complete cases	Employer's perspective	23	52	Need for recovery	3120 (-36 to 5129)	-0.5 (-6.5 to 5.6)	-6606	56.0	2.3	0.3	41.4
		23	52	Job satisfaction	3120 (-36 to 5129)	0.0 (-0.2 to 0.2)	635626	72.5	2.0	0.6	24.9
SA2 - PRODISQ	Societal perspective	92	106	Need for recovery	-4168 (-7319 to 989)	-8.4 (-14.6 to -2.2)	498	1.4	98.1	0.4	0.0
		92	106	General Vitality	-4168 (-7319 to 989)	3.4 (-9.3 to 16.2)	-1211	0.6	71.6	26.9	0.8
SA2 - PRODISQ	Employer's perspective	92	106	Need for recovery	-2750 (-6148 to 652)	-8.4 (-14.6 to -2.2)	328	9.3	90.3	0.4	0.0
		92	106	Job satisfaction	-2750 (-6148 to 652)	-0.1 (-0.6 to 0.5)	43961	2.9	37.0	53.7	6.5
SA3 – HCA	Societal perspective	92	106	Need for recovery	1867 (-596 to 4754)	-8.4 (-14.6 to -2.2)	-223	82.4	17.2	0.0	0.4
		92	106	General Vitality	1867 (-596 to 4754)	3.4 (-9.3 to 16.2)	542	57.0	15.2	2.0	25.8
SA3 - FCA	Employer's perspective	92	106	Need for recovery	2888 (550 to 5503)	-8.4 (-14.6 to -2.2)	-345	94.9	4.7	0.0	0.4
		92	106	Job satisfaction	2888 (550 to 5503)	-0.1 (-0.6 to 0.5)	-46180	36.3	3.5	1.1	59.0
SA4 – excluding presenteeism	Societal perspective	92	106	Need for recovery	86 (-1907 to 2325)	-8.4 (-14.6 to -2.2)	-10	95.0	4.5	0.0	0.4
		92	106	General Vitality	86 (-1907 to 2325)	3.4 (-9.3 to 16.2)	25	33.4	38.9	13.8	14.0
SA4 – excluding presenteeism	Employer's perspective	92	106	Need for recovery	1483 (-872 to 4103)	-8.4 (-14.6 to -2.2)	-177	83.3	16.3	0.0	0.4
		92	106	Job satisfaction	1483 (-872 to 4103)	-0.1 (-0.6 to 0.5)	-23714	33.3	6.5	9.8	50.4
SA5 – age- and gender-specific price weights	Societal perspective	92	106	Need for recovery	-14 (-1274 to 1557)	-8.4 (-14.6 to -2.2)	2	40.2	59.4	0.1	0.2
		92	106	General Vitality	-14 (-1274 to 1557)	3.4 (-9.3 to 16.2)	-4	22.7	49.6	10.0	17.8
SA5 – age- and gender-specific price weights	Employer's perspective	92	106	Need for recovery	1436 (130 to 3180)	-8.4 (-14.6 to -2.2)	-171	87.3	12.3	0.0	0.3
		92	106	Job satisfaction	1436 (130 to 3180)	-0.1 (-0.6 to 0.5)	-22959	31.3	8.5	3.8	56.3
Social environmental intervention											
		Intervention	Control		€	Points	€/point	NE[1]	SE[2]	SW[3]	NW[4]
SA1 – complete cases	Societal perspective	41	52	Need for recovery	105 (-2148 to 2404)	-3.3 (-10.9 to 4.3)	-32	47.1	43.4	3.9	5.6
		41	52	General Vitality	105 (-2148 to 2404)	3.5 (-0.6 to 7.5)	30	46.3	4.4	1.9	43.6
SA1 – complete cases	Employer's perspective	41	52	Need for recovery	359 (-1880 to 2901)	-3.3 (-10.9 to 4.3)	-110	51.0	39.5	3.6	5.9
		41	52	Job satisfaction	359 (-1880 to 2901)	0.0 (-0.2 to 0.2)	-281037	23.7	17.0	26.1	33.2

SA2 - PRODISQ	Societal perspective	118	106	Need for recovery	-1812 (-5602 to 681)	0.1 (-8.8 to 9.0)	-17268	17.2	32.1	25.4	15.3
		118	106	General Vitality	-1812 (-5602 to 681)	7.3 (-3.8 to 18.4)	-248	26.7	61.3	6.2	5.5
SA2 - PRODISQ	Employer's perspective	118	106	Need for recovery	-1538 (-4974 to 1110)	0.1 (-8.8 to 9.0)	-14652	18.1	31.2	34.4	16.4
		118	106	Job satisfaction	-1538 (-4974 to 1110)	-0.2 (-0.9 to 0.5)	6898	7.7	13.1	52.4	26.8
SA3 – HCA	Societal perspective	118	106	Need for recovery	77 (-2114 to 2178)	0.1 (-8.8 to 9.0)	729	9.3	90.3	0.4	0.0
		118	106	General Vitality	77 (-2114 to 2178)	7.3 (-3.8 to 18.4)	10.5	42.4	45.9	2.6	9.1
SA3 - FCA	Employer's perspective	118	106	Need for recovery	558 (-1381 to 2559)	0.1 (-8.8 to 9.0)	5321	27.3	22.0	15.3	35.4
		118	106	Job satisfaction	558 (-1381 to 2559)	-0.2 (-0.9 to 0.5)	-2505	15.5	5.3	32.0	47.2
SA4 – excluding presenteeism	Societal perspective	118	106	Need for recovery	121 (-1565 to 1704)	0.1 (-8.8 to 9.0)	-1150	27.7	21.6	20.4	30.3
		118	106	General Vitality	121 (-1565 to 1704)	7.3 (-3.8 to 18.4)	17	49.0	39.3	3.4	8.3
SA4 – excluding presenteeism	Employer's perspective	118	106	Need for recovery	414 (-1241 to 1934)	0.1 (-8.8 to 9.0)	3941	32.7	16.6	15.3	35.4
		118	106	Job satisfaction	414 (-1241 to 1934)	-0.2 (-0.9 to 0.5)	-1855	14.3	6.5	25.3	53.8
SA5 – age- and gender-specific price weights	Societal perspective	118	106	Need for recovery	67 (-1226 to 1475)	0.1 (-8.8 to 9.0)	639	19.6	29.7	21.5	29.2
		118	106	General Vitality	67 (-1226 to 1475)	7.3 (-3.8 to 18.4)	9	38.3	50.0	2.6	9.1
SA5 – age- and gender-specific price weights	Employer's perspective	118	106	Need for recovery	416 (-751 to 1789)	0.1 (-8.8 to 9.0)	3968	25.7	23.5	16.2	34.5
		118	106	Job satisfaction	416 (-751 to 1789)	-0.2 (-0.9 to 0.5)	-1868	15.1	5.7	34.0	45.2
Physical environmental intervention											
		Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
SA1 – complete cases	Societal perspective	34	52	Need for recovery	1338 (-2247 to 4942)	-1.6 (-9.0 to 5.7)	-824	61.1	18.9	5.3	14.7
		34	52	General Vitality	1338 (-2247 to 4942)	3.8 (-0.1 to 7.7)	351	74.5	1.3	3.2	43.6
SA1 – complete cases	Employer's perspective	34	52	Need for recovery	1250 (-2162 to 4648)	-1.6 (-9.0 to 5.7)	-770	61.0	18.9	5.6	14.4
		34	52	Job satisfaction	1250 (-2117 to 4616)	0.0 (-0.3 to 0.2)	-60588	32.7	15.0	9.6	42.7
SA2 - PRODISQ	Societal perspective	96	106	Need for recovery	-3917 (-6916 to -913)	-1.2 (-9.1 to 6.6)	3241	3.3	70.1	24.8	1.8
		96	106	General Vitality	-3917 (-6916 to -913)	5.7 (-5.6 to 17.1)	-692	1.8	80.0	14.9	3.3
SA2 - PRODISQ	Employer's perspective	96	106	Need for recovery	-3506 (-6528 to -543)	-1.2 (-9.1 to 6.6)	2901	3.9	69.4	24.2	2.5
		96	106	Job satisfaction	-3506 (-6528 to -543)	-0.1 (-0.5 to 0.4)	67903	1.4	46.4	47.3	5.0
SA3 – HCA	Societal perspective	96	106	Need for recovery	436 (-1899 to 2790)	-1.2 (-9.1 to 6.6)	-361	43.8	29.5	12.9	13.8
		96	106	General Vitality	436 (-1899 to 2790)	5.7 (-5.6 to 17.1)	77	42.3	39.4	3.0	15.3
SA3 - FCA	Employer's perspective	96	106	Need for recovery	1092 (-1091 to 3300)	-1.2 (-9.1 to 6.6)	903	54.6	18.7	8.8	17.8
		96	106	Job satisfaction	1092 (-1091 to 3300)	-0.1 (-0.5 to 0.4)	-21145	26.5	21.3	6.2	46.0
SA4 – excluding presenteeism	Societal perspective	96	106	Need for recovery	-317 (-1985 to 1359)	-1.2 (-9.1 to 6.6)	263	25.0	48.3	16.9	9.8
		96	106	General Vitality	-317 (-1985 to 1359)	5.7 (-5.6 to 17.1)	-56	26.5	55.2	10.0	8.3

7

		Physical environmental intervention									
		Intervention	**Control**		**€**	**Points**	**€/point**	**NE[1]**	**SE[2]**	**SW[3]**	**NW[4]**
SA4 – excluding presenteeism	Employer's perspective	96	106	Need for recovery	36 (-1739 to 1691)	-1.2 (-9.1 to 6.6)	-30	35.5	37.8	11.9	14.8
		96	106	Job satisfaction	36 (-1739 to 1691)	-0.1 (-0.5 to 0.4)	-696	23.6	24.2	25.5	26.8
SA5 – age- and gender-specific	Societal perspective	96	106	Need for recovery	181 (-1184 to 1835)	-1.2 (-9.1 to 6.6)	-150	38.7	34.6	14.7	12.0
price weights		96	106	General Vitality	181 (-1184 to 1835)	5.7 (-5.6 to 17.1)	32	36.8	45.0	4.3	13.9
SA5 – age- and gender-specific	Employer's perspective	96	106	Need for recovery	546 (-766 to 2058)	-1.2 (-9.1 to 6.6)	-452	49.5	23.9	10.3	16.4
price weights		96	106	Job satisfaction	546 (-766 to 2058)	-0.1 (-0.5 to 0.4)	-10570	24.6	23.1	11.0	41.3

Abbreviations: CI: Confidence Interval, C: Costs, E: Effects, ICER: Incremental Cost-Effectiveness Ratio, CE-plane: Cost-Effectiveness plane, SA: Sensitivity Analysis, PRODISQ: "PROductivity and DISease Questionnaire", FCA: Friction Cost Approach, HCA: Human Capital Approach

Note: Costs are expressed in 2011 Euros

[1] Refers to the northeast quadrant of the CE-plane, suggesting that the intervention is more effective and more costly compared to usual practice

[2] Refers to the southeast quadrant of the CE-plane, suggesting that the intervention is more effective and less costly compared to usual practice

[3] Refers to the northwest quadrant of the CE-plane, suggesting that the intervention is less effective and more costly compared to usual practice

[4] Refers to the southwest quadrant of the CE-plane, suggesting that the intervention is less effective and less costly compared to usual practice

Appendix 3: Intervention costs, benefits, Net Benefits (NB), Benefit Cost Ratio (BCR), and Return-On-Investment (ROI) per participant for the sensitivity analyses

Sensitivity analysis	Sample size		Costs (€)	Benefits	Financial return			
	Intervention	Control	Total (95% CI)	Total (95% CI)	NB[1] (95% CI)	BCR[2] (95% CI)	ROI (%)[3] (95% CI)	Probability
Social and physical environmental intervention								
SA1 – complete cases	23	52	466 (NA)	-2654 (-6958 to 758)	-3120 (-6947 to 662)	-5.7 (-13.7 to 2.4)	-670 (-1491 to 142)	0.06
SA2 – PRODISQ	92	106	466 (NA)	3216 (-486 to 6787)	2750 (-731 to 6559)	6.9 (-0.6 to 15.1)	590 (-157 to 1407)	0.90
SA3 – FCA	92	106	466 (NA)	-2377 (-5155 to -2377)	-2843 (-5408 to -69)	-5.1 (-10.6 to 0.9)	-610 (-1161 to -15)	0.05
SA4 – Excluding presenteeism	92	106	466 (NA)	-1017 (-3844 to 1502)	-1483 (-4178 to 1156)	-2.2 (-8.0 to 3.5)	-318 (-897 to 248)	0.16
SA5 – age- and gender-specific price weights	92	106	466 (NA)	-970 (-2820 to 466)	-1436 (-2832 to 470)	-2.1 (-5.1 to 2.0)	-308 (-608 to 101)	0.13
Social environmental intervention								
SA1 – complete cases	41	52	430 (NA)	71 (-3415 to 3457)	-359 (-3584 to 3272)	0.2 (-7.3 to 8.6)	-84 (-833 to 761)	0.44
SA2 – PRODISQ	118	106	430 (NA)	1968 (-978 to 5672)	1538 (-2167 to 4485)	4.6 (-4.0 to 11.4)	357 (-504 to 1042)	0.66
SA3 – FCA	118	106	430 (NA)	128 (-2410 to 1970)	-558 (-2727 to 1662)	-0.3 (-5.3 to 4.8)	-130 (-634 to 386)	0.39
SA4 – Excluding presenteeism	118	106	430 (NA)	17 (-1652 to 1738)	-414 (-2102 to 1293)	0.0 (-3.9 to 4.0)	-96 (-488 to 300)	0.33
SA5 – age- and gender-specific price weights	118	106	430 (NA)	14 (-1543 to 1297)	-416 (-1696 to 1140)	0.0 (-2.9 to 3.6)	-97 (-394 to 265)	0.40
Physical environmental intervention								
SA1 – complete cases	34	52	72 (NA)	-1250 (-5232 to 2740)	-1321 (-5536 to 2415)	-17.5 (-76.5 to 34.8)	-1849 (-7747 to 3381)	0.23
SA2 – PRODISQ	96	106	72 (NA)	3578 (576 to 6976)	3506 (433 to 6820)	50.1 (7.0 to 94.4)	4907 (605 to 9543)	0.93
SA3 – FCA	96	106	72 (NA)	-1007 (-3228 to -1007)	-1078 (-3480 to 1280)	-14.1 (-47.7 to 18.9)	-1509 (-4870 to 1791)	0.27
SA4 – Excluding presenteeism	96	106	72 (NA)	36 (-1661 to 2049)	-36 (-1933 to 1768)	0.5 (-23.7 to 28.1)	-50 (-2705 to 2474)	0.49
SA5 – age- and gender-specific price weights	96	106	72 (NA)	-474 (-2008 to 1072)	-546 (-2007 to 1076)	-6.6 (-27.1 to 16.1)	-764 (-2809 to 1505)	0.33

Abbreviations: CI: Confidence Interval, NB: Net Benefit, BCR: Benefit Cost Ratio, ROI: Return-On-Investment, I: Intervention, C: Control, NA: Not Applicable, SA: Sensitivity Analysis, PRODISQ: "PROductivity and DISease Questionnaire", FCA: Friction Cost Approach

Note: Costs are expressed in 2011 Euros

Note: Financial returns are positive if the following criteria are met: NB>0, BCR>1, and ROI>0

[1] Indicates the amount of money returned after intervention costs are recovered

[2] Indicates the amount of money returned per Euro invested in the intervention

[3] Indicates the percentage of profit per Euro invested in the intervention

8

Trial-based economic evaluations in occupational health: principles, methods, and recommendations

Johanna M van Dongen
Marieke F van Wier
Emile Tompa
Paulien M Bongers
Allard J van der Beek
Maurits W van Tulder
Judith E Bosmans

J Occup Environ Med, accepted for publication

ABSTRACT

Resources for occupational health are scarce. To allocate available resources as efficiently as possible, decision-makers need information on the relative economic merits of occupational health and safety (OHS) interventions. Economic evaluations can provide this information by comparing the costs and consequences of two or more alternative interventions. However, only a few of the studies that consider the effectiveness of OHS interventions take the extra step of considering whether they are efficient in terms of their resource implications. Moreover, the methodological quality of economic evaluations in the occupational health literature is generally poor.

Effectiveness trials are commonly used as a vehicle for economic evaluations, since they provide a unique opportunity to reliably estimate the resource implications of a new intervention without substantially higher research expenses. The present paper aims to help occupational health researchers conduct high quality trial-based economic evaluations by discussing the theory and methodology that underlie them, and by providing recommendations for good practice regarding their design, analysis, and reporting. The present paper will also help consumers of this literature with understanding and critically appraising trial-based economic evaluations of OHS interventions.

INTRODUCTION

Resources for occupational health are scarce (1;2). Therefore, decision-makers in this field increasingly call upon advisors and researchers to not only demonstrate that occupational health and safety (OHS) interventions are effective, but also efficient in terms of their resource implications. Economic evaluations provide information on the relative efficiency of two or more alternative interventions and are defined as *"the comparative analysis of alternative courses of action in terms of both their costs and consequences"* (1). The main aspects of any economic evaluation are to identify, measure, value, and compare the costs and consequences of alternatives (1).

In the healthcare sector, economic evaluations are increasingly being conducted and play an important role in many countries when deciding whether (new) treatments should be covered by public funding (1). However, only a few of the studies that consider the effectiveness of OHS interventions take the extra step of considering whether they are efficient in terms of their resource implications (3). Moreover, the methodological quality of those that do is generally poor (4-7). Reasons for this may be the distinct challenges that confront researchers when trying to identify the resource implications of OHS interventions as well as a lack of recommendations on how to deal with these issues (3). Many economic evaluation text books and articles are designed for use in healthcare settings and may therefore be difficult to adapt to the occupational health context (4).

Effectiveness trials are a commonly used vehicle for economic evaluations, as they provide a unique opportunity to reliably estimate the resource implications of a new intervention without substantially higher research expenses. Although some efforts have been undertaken to improve the quality of (trial-based) economic evaluations in occupational health (3;8;9), more needs to be done to accomplish this. Therefore, the present paper aims to help occupational health researchers conduct high quality trial-based economic evaluations by discussing the theory and methodology that underlie them, and by providing recommendations for good practice regarding their design, analysis, and reporting.

DESIGN OF AN ECONOMIC EVALUATION

Kind of economic evaluations

Choosing the appropriate kind of economic evaluation for a particular occupational health decision context can be a challenge as a result of the relative complexity of the decision-making context that generally includes multiple stakeholders (e.g. workers, employers, insurance companies, public policy makers). Four kinds of economic evaluations are distinguished. There are similarities across the four kinds. The main difference is the metric used to measure the key outcome (health and/or safety, in the case of OHS interventions) (10).

1) Cost-effectiveness analysis (CEA): Costs and some consequences (e.g. productivity, healthcare utilization implications) are measured in monetary units, whereas the key outcome is measured in natural units (1).
2) Cost-benefit analysis (CBA): Both costs and consequences are measured in monetary units. In business administration, CBAs are sometimes describes as return-on-investment analyses.
3) Cost-utility analysis (CUA): Costs and some consequences are measured in monetary terms, whereas the key outcome is measured in utility units. Utilities are often expressed in terms of Quality Adjusted Life Years (QALYs) (1).
4) Cost-minimisation analysis (CMA): Only costs are considered across alternatives, as it is assumed that the consequences are similar. CMAs are considered inappropriate if there is uncertainty regarding a possible difference in the magnitude of consequences (1).

Which kind of economic evaluation is most appropriate depends on the stakeholders involved and the question being asked. Generally, employers are most interested in CBAs that can provide insight into the impact of an intervention on a company's bottom-line, whereas public policy makers may be more interested in CEAs and CUAs, particularly if monetary measures do not adequately capture important health outcomes (1;8;11). Therefore, it is recommended that researchers conduct various kinds of economic evaluations within the same study in order to inform all relevant stakeholders (3).

When to undertake an economic evaluation

Economic evaluations are often conducted alongside ("piggybacked" onto) trials evaluating the effectiveness of OHS interventions. Various design aspects are therefore typically determined by the requirements of the effectiveness trial (e.g. alternatives, outcome measures). However, to ensure that all relevant economic data is collected in a valid, reliable, and efficient way, it is important to consider the requirements for the economic evaluation at the earliest possible stage (12-14).

Debate exists as to whether an economic evaluation should be included in a trial before the effectiveness of a new intervention is established. However, not including an economic evaluation would risk losing the opportunity to simultaneously collect cost and effect data (14). Also, the absence of statistically significant consequence/effect differences between the alternatives being compared does not necessarily imply that the new alternative is not cost-effective and/or cost-beneficial. Economic evaluations are about the joint distribution of costs and consequences, and could demonstrate clear cost-effectiveness/cost-benefit when neither cost nor consequence differences are individually significant (14). Also, cost savings might occur in the absence of health improvements and could thus be missed if an economic evaluation is not performed.

Trial design

Pragmatic randomized controlled trials (RCTs) are generally acknowledged as the best vehicle for economic evaluations, because they enable the evaluation of the resource implications of OHS interventions under "real life" conditions. This setup increases the external validity of results, while the internal validity is guaranteed by the randomization of participants (4;14). Within the occupational health setting, however, participant-level randomization may not always be feasible (e.g. when interventions include organizational components). In such cases, randomization at the level of departments or locations might provide a more feasible approach (i.e. cluster-RCTs) (3).

To ensure that the results of an economic evaluation are generalizable to occupational health practice, trial conditions should resemble daily practice as much as possible. For example, participants should be similar to those who will experience the intervention if it is implemented broadly, monitoring should be done under routine circumstances, and interventions should be compared to usual practice.

Perspective

An essential aspect of an economic evaluation is its perspective. Perspective refers to the "point of view" taken to identify relevant costs and consequences for inclusion in the evaluation. The chosen perspective may be that of any relevant stakeholder or an aggregate of stakeholders such as a societal perspective. The perspective determines which costs and consequences are included. In the societal perspective, for example, all costs and consequences are considered irrespective of who pays or benefits, whereas only those borne by employers are included when the employer's perspective is applied. Given this fact, the perspective is a critical element in an analysis and should therefore be stated explicitly (1).

OHS interventions are typically initiated by company management; either to comply with the law, in an effort to save money (i.e. reduced sickness absence costs), or for moral reasons (11). Consequently, most economic evaluations of such interventions are performed from the employer's perspective (4-7;15), but other perspectives may also be relevant; e.g. worker's, insurer's, and societal perspective. When the employer's perspective is applied, key worker outcomes, such as the value of worker health, are often not included in the analysis, but simply the health-related expenses incurred by an employer (e.g. productivity implications). This is a critical oversight, as occupational health is essentially about worker health. A societal perspective is particularly useful to consider as the perspective in a study, as it provides insight into the net effect across all stakeholders. Hereby, it better ensures that the societal costs of an intervention are less than the benefits experienced by all stakeholders, rather than simply the company's costs being less than its benefits (3). This information will ensure that there is a net societal benefit, rather than simply cost shifting from one stakeholder to another. In addition, the disaggregated information on costs and consequences from a societal perspective provides a good sense of their distribution across stakeholders. Such information can be the launch pad for bargaining between them (1). This may be of particular importance in countries with dual-payer (e.g. The Netherlands) and universal healthcare systems (e.g. The United Kingdom), since employers generally bear most of the costs of OHS interventions, whereas in such jurisdictions the healthcare system and/or government reaps a large part of their benefits (i.e. reduced medical spending) (16). Therefore, it is recommendable to supplement findings from the employer's perspective with those from other relevant perspectives, particularly the societal one.

Analytic time frame

Researchers also need to decide about the time frame over which costs and consequences are analysed. The analytic time frame ought to cover the entire period over which costs and consequences flow from the alternatives under consideration (12). This time frame generally extends beyond the follow-up needed to establish the effectiveness of a new intervention. To illustrate, the follow-up of an effectiveness trial may be terminated after the occurrence of the clinical event of interest (e.g. incidence of Repetitive Strain Injury (RSI)). If this follow-up would be used for the economic evaluation, all costs and consequences incurred during the course of the disorder or its recurrences would not be taken into account (e.g. RSI-related medication and/or operation costs), leading to an underestimation of the total costs and consequences. Although the optimal follow-up period is generally unknown, researchers and readers should at least feel confident that the most important costs and consequences are covered by the chosen analytic time frame. Additionally, future costs and consequences that occur after the measurement period can be estimated using information and data from various sources. This is particularly important to do if future costs and consequences are expected to be substantial (e.g. many of the (health) benefits of preventive interventions are thought to occur in the future).

Identification, measurement, and valuation of resource use

In economic evaluations, costs and some consequences are expressed in monetary units. For this purpose, relevant resource use categories should be identified, measured, and valued. As discussed earlier, relevant resource use categories for inclusion in an economic evaluation depend on its perspective. Other factors that might determine the relevance of a resource use category are, amongst others, the country or jurisdiction in which the study is undertaken and the nature of the alternatives being compared.

After relevant resource use categories are identified, researchers should determine how to cost them. Costing generally involves three steps; 1) the measurement of quantities of resources consumed *(Q)*, 2) the assignment of unit prices *(P),* and 3) the valuation of resources consumed by multiplying their quantities by their respective unit prices *(Q*P)* (1). These estimates should be reported separately so that readers can judge the relevance of these measures to his or her setting (17).

Measurement of quantities of resources consumed

Resource use data are ideally collected prospectively through a data collection process that is fully integrated into the effectiveness trial (1;13). Also, when collecting self-reported resource use data, researchers have to balance recall bias against completeness of information. Shorter recall periods reduce the risk of participants forgetting important information. However, collecting data with relatively short recall periods (e.g. a couple of weeks) over a longer period of time may be overly burdensome to participants and may thus increase the risk of missing data and drop-outs. Therefore, it may be better to maximize completeness at the cost of some recall bias (14); e.g. by using 2- to 3-month recall periods in a trial with a long-term follow-up (≥12 months) (18). Also, care should be taken to collect resource use data continuously during follow-up and to avoid the need for extrapolation of resource use estimates between measurement periods.

Assignment of unit prices

Unit prices used for valuing resource use ought to reflect opportunity costs; i.e. *"the value of a resource in its most highly valued alternative use"* (8). In a world of perfect markets, such costs are revealed by the market price of a good or service. However, if a competitive market does not exist for a good or service, market prices often are an inaccurate measure of its value. For example, if a premium is paid for a good or service due to restricted market entry, market prices may overestimate the opportunity costs at the societal level. When the societal perspective is applied, an adjustment should therefore be made to the market price; e.g. by using the price of a comparable good or service (8). For the employer's perspective, the actual purchase costs incurred by the employer may be more appropriate, as they better represent the sum of money that is not available to the employer for its best alternative use (12;19). Thus, appropriate unit prices may vary between perspectives, and researchers should ensure that they reflect the true resource implications to the decision-maker at hand (8).

A brief description of the methods used for measuring and valuing the most frequently used resource use categories in economic evaluations of OHS interventions is provided below. The most frequently used resource use categories are; intervention, healthcare, productivity, and worker's compensation costs (4-7;15).

Intervention costs

Information on the market price of an intervention may be derived from vendors or company and/or research project records. Many trials, however, assess novel interventions that either have no predefined price weights associated with them or for which the use of market prices is inappropriate (e.g. when the societal perspective is applied) (12). In such cases, the actual intervention costs can be assessed using a bottom-up micro-costing approach, in which detailed data regarding the quantities of resources consumed as well as their unit prices are collected per intervention component separately. Such resources may include intervention staff hours, materials used, depreciation, overhead activities, square feet of office space, and traveling (1;3;12). Also, workers may be taken away from their regular production activities to participate in the intervention and this should be accounted for as well. Costs associated with the intervention's evaluation should not be included unless it is a condition of implementation (8).

Quantities of resources consumed can be measured using administrative databases, expert panels, surveys or interviews with intervention participants and/or providers, intervention operation logs, or observations (20). Unit prices may be collected from administrative databases, scientific literature, vendors, and/or costing manuals (e.g. (21)).

Healthcare costs

Ideally, all healthcare service use is measured to reduce the likelihood that (unexpected) shifts in healthcare utilization rates are missed. Although this approach will increase the validity of the results, it may not always be feasible. An alternative strategy is to limit data collection to those healthcare services that are related to the alternatives and/or condition under study (12). A description of the care path for the condition under study might provide researchers with a clear picture of what those healthcare services are. In all cases, care should be taken to include the most important cost drivers.

Healthcare utilization can be measured through a variety of means, including retrospective questionnaires, prospective resource use diaries (i.e. cost diaries), and insurance or hospital databases. Databases, however, may not always contain all required data, and their validity and reliability may not be very high (10). Moreover, healthcare costs borne by participants (e.g. co-payments, over-the-counter medication) are typically not included in these databases. Therefore, researchers are often dependent on self-report data to measure these healthcare utilization items. To value healthcare utilization, unit prices may either be estimated using a micro-costing approach, or based on predefined price weights, prices according to professional organizations, or tariffs. Typically, several methods are used simultaneously (10;19).

Productivity costs

For employers, an important benefit of OHS interventions are the resulting changes in productivity loss. Productivity loss can be defined as the company's output loss corresponding to reduced labour input (i.e. time and efforts/skills of the workforce). According to this definition, to value productivity loss is to value the output loss (22). Unfortunately, however, objective measurement of the true impact of reduced labour input on a company's output is often impossible to estimate. Therefore, researchers typically use proxies of productivity loss, which are often estimated using (self-reported) data on the participants' level of absenteeism (i.e. sickness absence) and/or presenteeism (i.e. reduced performance while at work). The methodologies used for measuring and valuing absenteeism and presenteeism are a fiercely debated topic in the field of economic evaluations. Below, a brief description of the most frequently used methods is provided. For more information about the main debates and developments regarding the identification, measurement, and valuation of productivity we refer to other publications; (22;23).

The two main methods for estimating absenteeism costs are the Human Capital Approach (HCA) and the Friction Cost Approach (FCA). For both methods, the number of sickness absence days has to be collected, for which administrative databases, self-report (questionnaires), or reports by others can be used (9). For the FCA, it is also important to identify the number and duration of different absence periods. According to the HCA, absenteeism costs are equal to the amount of money participants would have earned had they not been injured or ill (4;21). Therefore, in the HCA, sickness absence days are typically valued using actual wage rates of participants (including employment overheads and benefits) and represent losses for the entire duration of absence (1;19;24). It is argued that the HCA overestimates the true societal cost of sickness absence, as the possible replacement of workers with long-term sickness absence is not taken into account (1;4). Therefore, the FCA was developed, in which production losses are assumed to be confined to the time-span companies need to replace a sick worker by a formerly unemployed person to restore the company's initial production level (i.e. friction period) (23). In the FCA, absenteeism is typically valued using age-, gender- and/or education-specific price weights (25). The length of the friction period depends on the state (i.e. the unemployment rate) and efficiency of the labour market. As such, friction periods typically differ between countries and should be estimated per country separately (1). If there are important changes in the economic climate, it may be necessary to estimate the friction period anew. In the Netherlands, a friction period of 23 weeks is currently assumed (21). Thus, if a sickness absence period exceeds 23 weeks, absenteeism costs are truncated at the costs of 23 weeks. Furthermore, as a reduction of labour input is often assumed to cause a less than proportional reduction in productivity, Koopmanschap et al. (1995) also proposed the application of an elasticity factor of 0.8, which is often used in economic evaluations that apply the FCA. This elasticity factor implies that a 100% loss of labor input corresponds with an 80% reduction in productivity (25).

In the economic evaluation literature, the need to consider presenteeism as a component of the costs incurred from productivity loss is increasingly being recognized (9). Presenteeism is typically estimated using participant self-report or report by others. For this purpose, various instruments have been developed, including both generic (26-29) and disease-specific questionnaires (30;31). Most of these questionnaires measure work performance in terms of points, percentages, or proportions (32). These responses can then be used to estimate the total number of working days lost due to presenteeism by using the equation:

$$P = (E - A) * p$$

where P is full working days lost due to presenteeism, E is total working days, A is sickness absence days, and p is the proportion of lost work performance estimated by the instrument used in the study (22). To value the number of lost working days due to presenteeism, actual wage rates of participants, or age-, gender-, and/or job-specific price weights can be used. Researchers should be aware, however, that the estimated number of work days lost due to presenteeism may vary widely between instruments. This suggests a lack of comparability among instruments, but it is still unclear which instrument provides the best presenteeism estimate (22). Given its significance, however, ignoring presenteeism may lead to severe underestimations (22). Therefore, researchers are recommended to include this resource use category whenever possible. To assess the possible influence of the choice of instrument, sensitivity analyses can be performed (See below).

Workers' compensation costs

Workers' compensation is an insurance program, offered in some countries (e.g. Canada, United States), through which workers may receive wage replacement and/or medical benefits in the event of an occupational injury or disease. Funding usually comes from premiums paid by employers (8). To estimate workers' compensation costs, total claim costs per participant can be obtained from company and/or workplace insurance records. It is generally inadequate, however, to use workers' compensation costs as the sole cost category, as they do not reflect the full extent of work-related injuries and illnesses (4). Many compensable injuries and illnesses go unreported and others are not compensable (4). When supplementing healthcare and/or productivity costs with workers' compensation costs, double counting should be avoided. Also, insurance premium-related wage replacement benefits should be excluded for the societal perspective, as they constitute "transfer payments" from the employer via the insurer to the worker rather than depleted sources (1;4)

Identification, measurement, and valuation of outcomes

As noted before, CEAs have the key outcome measured in natural units. The most appropriate outcome used for this purpose depends on the nature of the alternatives being compared, the condition under study, and/or the applied perspective. Sometimes, there may be some concern about whether the chosen outcome captures all relevant consequences. If this is a concern, it is advisable to conduct multiple CEAs using different outcomes (8). In CUAs, the key outcome is measured in utility units, generally known as QALYs. They capture both the duration of survival and health-related quality of life in a single measure (1;12;14). An advantage of QALYs is that they provide a general index score that allows decision-makers to compare the consequences of a range of interventions for different health issues (1;10). However, even though QALYs are the preferred outcome measure when healthcare interventions for patients are evaluated from the societal perspective (13;21;33), they have not yet been frequently used in economic evaluations of OHS interventions (4;6;7;34). This may be due to the fact that QALYs may not reflect what occupational health decision-makers feel is most important in terms of outcomes. In the case of a workplace safety programs, for example, outcomes such as worker safety may be more meaningful to decision-makers than a utility-weighted health measure (11). Moreover, occupational health decision-makers are generally unfamiliar with QALYs, and QALYs seem to lack sensitivity to mild conditions that are often the focus of OHS interventions (e.g. of worksite health promotion programs) (35). Therefore, more sensitive utility measures are warranted for economic evaluations of OHS interventions and/or utility measures that are more applicable to the occupational health setting; e.g. the recently conceptualized "Disease-Adjusted Working Years", which aims to express the amount of working years lost due to poor working conditions and associated illness (36;37).

ANALYSIS OF AN ECONOMIC EVALUATION

Below, we discuss some important issues in the analysis of trial-based economic evaluations. To illustrate some of them, data is used from an economic evaluation that was previously performed alongside a 12-month pragmatic RCT, in which construction workers at risk for cardiovascular disease either received a lifestyle intervention or usual practice. A CEA in terms of kilogram body weight loss was performed from the societal perspective and a CBA from that of the employer. Resource use categories included intervention, healthcare, absenteeism, and sports costs and were expressed in 2008 Euros. More detailed information about this trial-based economic evaluation can be found elsewhere; (38).

Sample size

Ideally, economic outcomes are used in the sample size calculation of a trial (13). However, although various techniques have been proposed to estimate the appropriate sample size for economic endpoints (39-42), sample size calculations are typically performed based on primary outcomes (10;13;14). This is due to the fact that cost data are right skewed and therefore require larger sample sizes to detect relevant differences than (health) outcome data. However, a large sample size may neither be feasible nor ethically acceptable (14;43). Also, a large number of parameters has to be specified to perform sample size calculations for economic endpoints (e.g. variance parameters of effectiveness measures, cost measures, incremental cost-effectiveness ratios), many of which are hard to predict a priori (39;41;42). Consequently, trial-based economic evaluations are typically underpowered for economic outcomes (10). Low powered studies have imprecise and uncertain cost estimates and should be interpreted with caution (43). Moreover, if studies are likely to be underpowered, researchers are recommended to use estimation rather than hypothesis testing (i.e. by using confidence intervals rather than p-values) (47).

Adjusting for differential timing

Interventions may have different time profiles of costs and consequences. Within occupational health, intervention costs are generally incurred immediately, while consequences such as productivity costs might extend into the future (44). Two types of adjustments should be made to account for these differences in timing. The first concerns the adjustment of cost data for inflation; i.e. *"the general upward price movement of goods and services"* (12). Due to inflation, prices drawn from different years are generally not comparable (8). All prices should therefore be adjusted to the same reference year using consumer price indices and the applied reference year should be stated explicitly (17). The second adjustment concerns the adjustment of cost and outcome data for time preferences of individuals when they are collected over a period of more than one year (12). Even within a world with zero inflation, individuals have a preference for receiving benefits today rather than in the future (1). Therefore, costs and consequences incurred in different years have to be discounted at some rate to estimate their present value (44). The appropriate discount rate depends on the borrowing cost of money and other contextual factors. Guidelines for discount rates used in public sector projects are provided by some jurisdictions. For example, in the Netherlands, cost data should be discounted at 4% and health outcomes at 1.5%, while both should be discounted at 3.5% in the United Kingdom (21;33).

Intention-to-treat and missing data

Guidelines for conducting trials prescribe that all participants should be included in the analyses, all retained in the group to which they were allocated (i.e. intention-to-treat analysis) (45). However, true intention-to-treat analyses are often hampered by missing data, which are generally inevitable in trials. For economic evaluations, this problem is even more pronounced, because total costs are typically the sum of numerous cost components. As such, cost data will already be incomplete if one component is missing (13). Missing data itself may have no relation to observed and unobserved factors among participants (MCAR: Missing Completely At Random), may only have a relationship to observed factors (MAR: Missing At Random), or may also have a relationship to unobserved factors (MNAR: Missing Not At Random) (See Box 1 for a more detailed description) (46). Historically, complete-case analyses (i.e. eliminating cases with missing data) were used to deal with missing data and this is still an often used approach in trial-based economic evaluations (47). However, complete-case analyses reduce the power of a study and lead to biased estimates if missing data are not MCAR (12;13). If the rate of missing data is smaller than 5%, complete-case analyses may be considered. If more than 5% of data are missing, researchers should use imputation techniques to fill in missing values. Nowadays, multiple imputation is generally recommended to impute missing data (13;14). When using multiple imputation, multivariate regression techniques are used to predict missing values on the basis of observed factors (12;14). To account for the uncertainty about the missing data, several different imputed datasets are created (46). As a rule of thumb, White et al. (2011) suggested that the number of datasets should at least be equal to the percentage of incomplete cases (48). The imputed datasets are subsequently analysed separately to obtain a set of parameter estimates, which can then be pooled using Rubin's rules to obtain overall estimates, variances, and 95% confidence intervals (95%CIs) (46;48;49). Multiple imputation leads to unbiased estimates if missing data are MAR (12). Researchers should bear in mind, however, that cost and consequence estimates derived using multiple imputation are less reliable and precise than those based on a 100% complete dataset (14). Every endeavor should therefore be made to minimize the amount of missing data.

Box 1: Types of missing data (46)

1) Missing Completely At Random (MCAR): The "missingness" of data has no relationship to observed and unobserved factors among participants. For example, sickness absence data may be missing because of problems with the registration of this data due to a temporary computer problem.

2) Missing At Random (MAR): The "missingness" of data has a relationship to observed factors among participants, but not to unobserved factors. For example, missing sickness absence durations may be longer than available sickness absence durations but only because older employees may be more likely to have missing sickness absence data.

3) Missing Not At Random (MNAR): Even after the observed data are taken into account, systematic differences remain between the missing values and the observed values. This means that the "missingness" of data also has a relationship to unobserved factors. For example, in trials relying on self-reported sickness absence, participants with longer sickness absence durations may be more likely to forget to return their cost diaries because they are not feeling well.

Incremental analysis of costs and consequences

After costs and consequences have been quantified, their mean differences between the intervention and control group(s) as well as the statistical significance of these differences need to be assessed (12).

As mentioned above, cost data are typically right skewed. This is caused by the fact that only a small proportion of participants incur high costs and costs are naturally bound by zero (See Figure 1) (1).

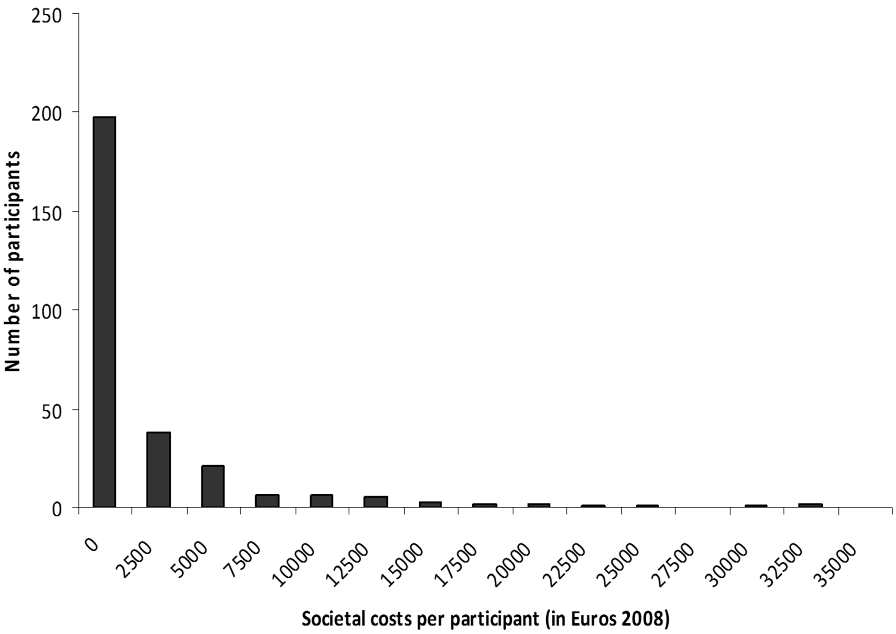

Figure 1: Distribution of the societal costs per participant in a trail-based economic evaluation of a lifestyle intervention for construction workers at risk for cardiovascular disease compared to usual practice (38)

The skewed cost distribution complicates the analysis of cost data, as it violates the assumptions of standard statistical tests, such as independent t-tests and linear regression analyses. A standard approach to describe skewed data is to provide a summary measure of the distribution in the form of a median. However, this is inappropriate for cost data as decision-makers need to be able to estimate the total

8

cost of implementing a new intervention (total implementation costs = mean costs per participant * number of participants). As such, the arithmetic mean is generally viewed as the most informative measure to describe cost data (1;14;50). Various methods are currently used to compare cost data between study arms, including standard non-parametric tests (e.g. Mann-Whitney U), t-tests on log-transformed data, and non-parametric bootstrapping. Standard non-parametric tests compare the distribution of the data instead of means and are therefore inappropriate. Transformations to normalize the distribution are not straightforward and are often sensitive to departures from distributional assumptions (13). Moreover, back-transformations are often complicated. Therefore, researchers increasingly favour the non-parametric bootstrap (13;50), which can be used to estimate 95%CIs around mean cost differences while avoiding distributional assumptions (Box 2) (51).

Box 2: Non-parametric bootstrapping

With non-parametric bootstrapping, statistical analyses are based on repeatedly sampling with replacement from the observed data. In short, a sample of *N* participants is repeatedly drawn with replacement from both the intervention and control group separately, where *N* equals the number of participants per study arm. Every resample (i.e. bootstrap sample) is the equivalent of a repetition of the trial. Since resamples have been drawn with replacement (i.e. per sample, participants can be drawn more than once), these bootstrap samples differ from one another. Per bootstrap sample, the statistics of interest is calculated (e.g. the difference in arithmetic mean costs and effects, incremental-cost effectiveness ratios, cost-benefit estimates). By doing so multiple times, a distribution for the statistics of interest is generated that provides an approximation of its population sampling distribution, which can then be used to estimate confidence intervals (12). At least 2000 bootstrap samples are recommended, and preferably more (52). Various methods have been proposed to estimate bootstrapped 95% confidence intervals, of which the bias corrected and accelerated method is currently the preferred one (51;53). Non-parametric bootstrapping is available in many software packages, including SPSS, SAS, STATA, and R.

Comparing incremental costs and consequences

The core of any economic evaluation is the analysis of the relation between the costs and consequences of alternatives. The preferred methods for conducting such analyses differ between the types of economic evaluations and are discussed below.

Cost-effectiveness analysis and cost-utility analysis

In CEAs and CUAs, an incremental cost-effectiveness ratio (ICER) is calculated by dividing the mean difference in cost (Δ Cost) between study arms by that in effect (Δ Effect). The ICER indicates the additional costs of a new intervention in comparison with a control condition per unit of effect gained (1;12).

$$\frac{Cost_{intervention} - Cost_{control}}{Effect_{intervention} - Effect_{control}} = \frac{\Delta\ Cost}{\Delta\ Effect} = ICER$$

To illustrate, a description of the calculation and interpretation of the example trial's ICER is provided in Box 3.

Box 3: Calculation and interpretation of the incremental cost-effectiveness ratio (ICER) of a lifestyle intervention for construction workers at risk for cardiovascular disease compared to usual practice (38)

During follow-up, intervention group participants significantly decreased their body weight by 2.02 kilogram compared to the control group (Δ Effect). Mean societal costs per participant were non-significantly higher in the intervention group compared to the control group by €293 (Δ Cost). Using this information, the ICER can be calculated;

ICER: €293/2.02 = €145

This ICER indicates that society has to pay €145 per participant in the intervention group for each additional kilogram body weight loss compared to usual practice.

ICERs are generally hard to interpret. For example, negative ICERs might represent reduced costs and positive effects indicating a win-win situation or increased costs and negative effects indicating a lose-lose situation (14). Therefore, ICERs are often graphically illustrated on cost-effectiveness planes (CE-planes), in which incremental effects are plotted on the X-axis and incremental costs on the Y-axis (Figure 2) (54;55). If an ICER is located either in the South East Quadrant (SE-Q) or the North West Quadrant (NW-Q), the choice between alternatives is clear (assuming there is no uncertainty surrounding the ICER). In the SE-Q, the new intervention is more effective and less costly than the control condition and is therefore said to dominate the control condition. In the NW-Q, the opposite is true and the new intervention is dominated by the control condition. If a new intervention is more effective and more costly (NE-Q: North East Quadrant) or less effective and less costly (SW-Q: South West Quadrant), the decision whether or not to adopt it depends on the so-called "willingness-to-pay" (λ). That is, the maximum amount of money decision-makers are willing to pay for an additional unit of effect (1). To illustrate, a hypothesized λ is depicted as the diagonal line in Figure 2 and divides the CE-plane into a cost-effective and a non-cost-effective halve. ICERs located to the right of this line are considered acceptable, whereas ICERs located to the left are considered inacceptable (14;54;55). The more decision-makers are willing to pay for an additional unit of effect, the steeper the slope of this line (14).

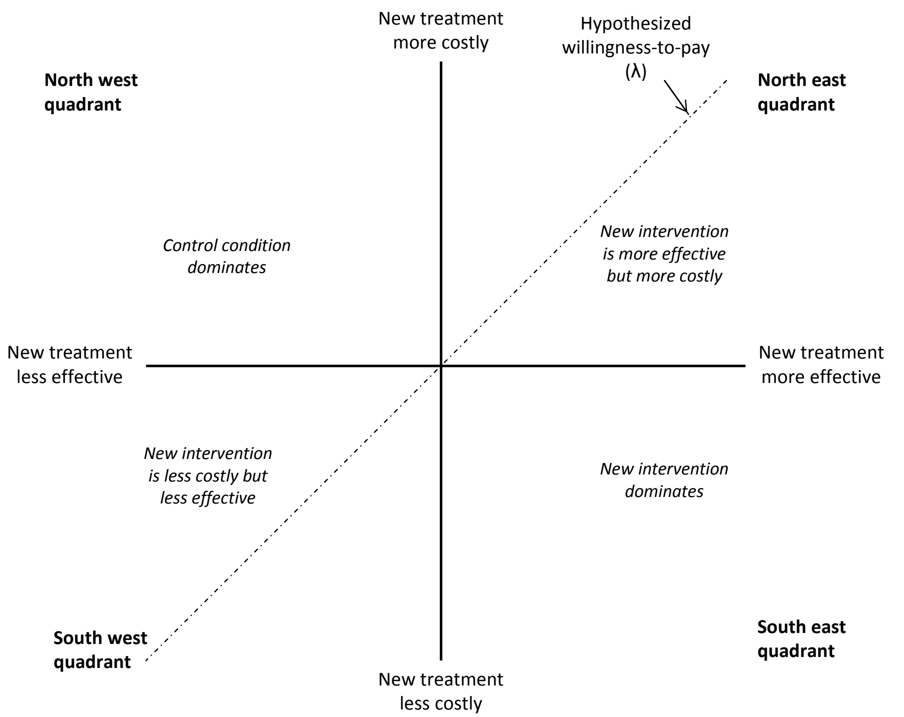

Figure 2: Cost-effectiveness plane

With participant-level data, it is natural to consider representing the uncertainty surrounding ICERs using 95%CIs. However, as a ratio measure, estimating 95%CIs around ICERs is not straightforward and, more importantly, 95%CIs around ICERs suffer from the same interpretation problem as ICERs (55). Therefore, alternative methods have been proposed to estimate the uncertainty surrounding ICERs. Current guidelines recommend using the bootstrap method described in Box 2. In this case, both incremental costs and effects are calculated per bootstrap sample. The uncertainty surrounding an ICER can then be graphically illustrated by plotting these bootstrapped incremental cost-effect pairs (CE-pairs) on a CE-plane. As indicated by the example trial's CE-plane provided in Figure 3, CE-pairs commonly cover more than one quadrant.

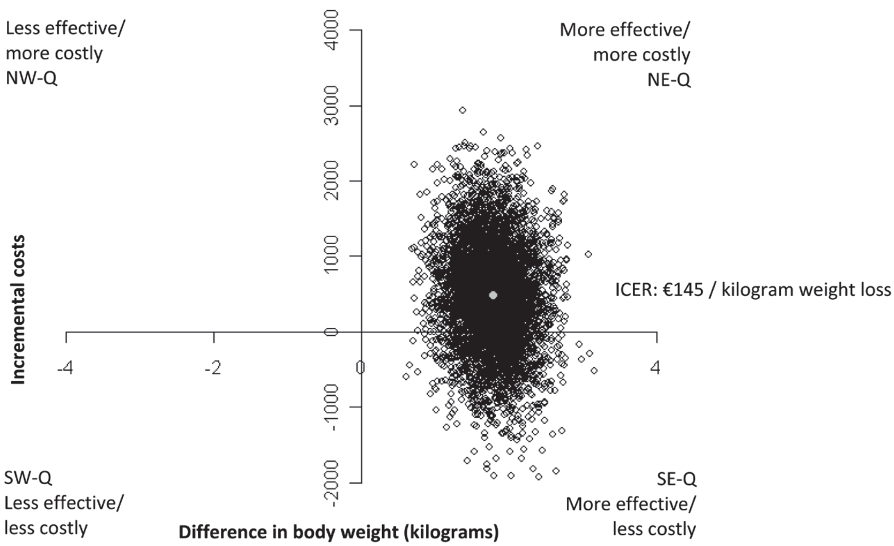

Figure 3: Cost-effectiveness plane for a lifestyle intervention for construction workers at risk for cardiovascular disease compared to usual practice (38)
Abbreviation: NW-Q: North West Quadrant, NE-Q: North East Quadrant, SW-Q: South West Quadrant, SE-Q: South East Quadrant, ICER: Incremental Cost-Effectiveness Ratio

Although CE-planes give a good impression of the uncertainty surrounding the ICER, they do not provide a summary measure of the joint uncertainty of costs and effects (56). Therefore, cost-effectiveness acceptability curves (CEACs) were introduced that provide insight into the probability that a new intervention is cost-effective compared to the control condition. This probability can be estimated by determining what proportion of CE-pairs is located in the cost-effective half of the CE-plane (i.e. to the right of the previously mentioned line with the slope equal to λ)(Figure 2). Since it is generally unknown what decision-makers are willing to pay for an additional unit of effect, λ is varied between its natural bounds (range: 0 to ∞) and the probability that the new intervention is cost-effective compared to the control condition is estimated for a range of λs. These values can then be plotted on CEACs that show the probability of cost-effectiveness (Y-axis) for various λs (X-axis) (55-57). To illustrate, the CEAC of the example trial is provided in Figure 4.

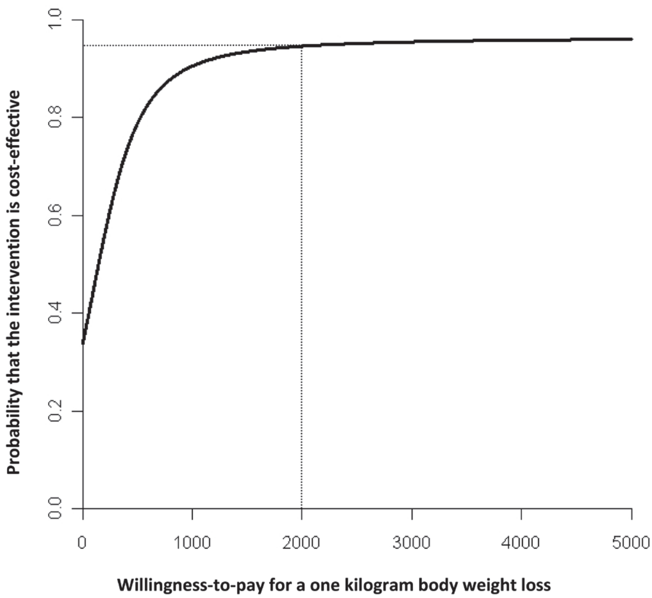

Figure 4: Cost-effectiveness acceptability curve for a lifestyle intervention for construction workers at risk for cardiovascular disease compared to usual practice (38)
Note: This cost-effectiveness acceptability curve corresponds with the cost-effectiveness plane in Figure 3 and indicates the probability of the intervention being cost-effectiveness for different values of willingness-to-pay per kilogram body weight loss.

This CEAC indicates that if decision-makers are not willing to pay anything to obtain an additional kilogram body weight loss (i.e. $\lambda=0$), there is a 0.33 probability that the new intervention is cost-effective compared to the control condition. If decision-makers are willing to pay €2000 (i.e. $\lambda=2000$), this probability is 0.95. When interpreting CEACs, two approaches can be used by decision-makers. If their willingness to pay is known, they have to judge whether the probability of cost-effectiveness at this ceiling ratio is acceptable. If their willingness to pay is unknown, they should consider whether the ceiling ratio at an acceptable probability of cost-effectiveness is acceptable to them. The latter might depend on the scale of the outcome measure and the prevalence of the condition under study.

Cost-benefit analysis

In health economics and business administration, various measures exist for comparing costs and benefits. Of them, the Net Benefits (NB), Benefit Cost Ratio (BCR), and Return-On-Investment (ROI) are the most frequently used measures in occupational health research and can be estimated using the following equations (6):

NB = Benefits – Costs
BCR = Benefits / Costs
*ROI = (Benefits – Costs) / Costs [*100]*

where *Costs* are defined as intervention costs and *Benefits* as the difference in monetized outcomes between the intervention and control group (e.g. difference in productivity costs). Benefits are estimated by subtracting the mean expenses incurred by the intervention group participants from those of the control group. Hereby, positive benefits indicate reduced spending. The NB indicates the amount of money gained after costs are recovered (i.e. net loss or net savings). The BCR indicates the amount of money returned per monetary unit invested. The ROI indicates the percentage of profit per monetary unit invested (58;59). Interventions can be regarded as cost saving if the following criteria are met: NB>0, BCR>1, and ROI>0%. To illustrate, a description of the calculation and interpretation of the example trial's cost-benefit estimates is provided in Box 4.

Box 4: Calculation and interpretation of the cost-benefit estimates of a lifestyle intervention for construction workers at risk for cardiovascular disease in comparison to usual practice (38)
Mean intervention costs per participant were €605. During follow-up, average absenteeism costs per participant were €3302 in the intervention group and €3604 in the control group. Thus, the absenteeism benefits per participant were €302 (€3604 - €3302). Using this information, cost-benefit estimates can be calculated;

NB: €302 - €605 = €-303
BCR: €302 / €605 = 0.50
ROI: ((€302 / €605)/ €302)*100 = -50%

These cost-benefit estimates indicate that the intervention resulted in a net loss to the employer of €303. Also, per Euro invested, the employer gained €0.50 and suffered a loss of 50%. Thus, the intervention cannot be regarded as cost-beneficial in terms of absenteeism costs.

Cost-benefit estimates, and BCRs and ROIs in particular, are typically presented without an indication of their uncertainty. If uncertainty is substantial and this is not taken into account, wrong conclusions could be drawn. Therefore, we recommend the use of the previously described bootstrap method (Box 2) to estimate the uncertainty surrounding cost-benefit estimates. In this case, the NB, BCR, and/or ROI are calculated per bootstrap sample (i.e. bootstrapped NBs, BCRs, and ROIs). Subsequently, 95%CIs can be estimated using the bias corrected and accelerated method (51;53). Even though BCRs and ROIs are ratio measures, estimating their 95%CIs is straightforward as the denominator (i.e. intervention costs) is typically positive. Many occupational health decision-makers, however, may lack the necessary statistical background to interpret 95%CIs (11). A possible way to deal with this issue is to determine what proportion of bootstrapped NBs, BCRs, and/or ROIs indicate cost savings (i.e. "the probability of financial return"). Occupational health decision-makers can subsequently use this information to consider whether the established probability of financial return is acceptable to them.

When reporting CBA results, economists and policy makers prefer the NB, whereas the BCR and ROI are more familiar to business managers. As such, it is recommendable to report at least two of them (i.e. NB and BCR/ROI), so that the results can be easily interpreted by all stakeholders. Another advantage of this approach is that it makes the results easily comparable with those of other studies, because different metrics are used in the literature to estimate whether OHS interventions generate cost savings (6).

Sensitivity analysis

Economic evaluations are often conducted in the context of incomplete information and uncertainty, which necessitates the use of proxy measures, and invariably, the need to make assumption about the methods and unit prices used for valuing resource use, the methods used for dealing with incomplete data, and the way in which adjustments are made for differential timing (4;8). Therefore, sensitivity analyses should be undertaken to assess how study results would change for different key assumptions and parameter values (i.e. the robustness of study results) (17;60). The ranges of values tested, and arguments for selecting these ranges, must be clearly described (10;17). Various kinds of sensitivity analyses exist. One-way sensitivity analyses assess the impact of changes to a single parameter at a time, while multiple parameters are varied simultaneously in multi-way sensitivity analyses (61).

DISCUSSION

Resources for occupational health are scarce. This makes it necessary for decision-makers to have information on the relative efficiency of OHS interventions in order to allocate available resources to their best use. As such, economic evaluations of OHS interventions are becoming increasingly important, many of which are conducted alongside effectiveness trials. Trial-based economic evaluations provide a unique opportunity to reliably estimate the resource implications of OHS interventions at low incremental cost (10;14). However, it is critical that high quality trial-based economic evaluations are performed when this information is used to inform allocation decisions.

Designing a high quality trial-based economic evaluation requires close collaboration between occupational health specialists, individuals executing the trial, and health economists (14). Careful considerations must be made regarding the perspective, the analytic time frame, the identification, measurement, and valuation of resource use and outcomes, as well as the methods used for calculating sample sizes, comparing costs and consequences, and handling missing data and uncertainty. The latter is of particular importance, as few economic evaluations in occupational health report on the uncertainty surrounding their incremental cost-consequence estimates (4-7;15). Failing to estimate values under uncertainty makes it impossible to determine the certainty of results and could thus lead to inappropriate decision-making. To quantify precision, non-parametric bootstrapping can be used as a statistical technique for dealing with the right skewed nature of cost data (1;7). An overview of our core recommendations for trial-based economic evaluations in occupational health can be found in Appendix 1.

Trial-based economic evaluations may also have shortcomings, including limited sample sizes, limited comparators, and truncated time horizons (14). To deal with the latter, researchers might consider extrapolating economic evaluation results beyond the follow-up of a trial by using decision analytic modeling, in which expected costs and consequences between alternatives are compared by synthesizing information from multiple sources (e.g. scientific literature, study results) (1;13;14). For more detailed information about decision analytic modeling we refer to other publications; (14;62). Also, even though we recommend a pragmatic (cluster-)RCT design for economic evaluations, we are aware that randomization itself may not always be feasible and/or desired in the occupational health setting. In those cases, well executed non-randomized studies may provide valuable information, but it is critical that efforts be made to control for selection bias (e.g. by using propensity score matching) (63;64).

When interpreting economic evaluations of OHS interventions, it is important to bear in mind that their results may not be directly applicable to other countries and jurisdictions due to differences in healthcare, social security systems, and other factors. Verbeek et al. (2010) demonstrated that economic evaluation results can be generalized from one country to another. However, to enable the necessary calculations, researchers need to provide an extensive description of the intervention, a detailed list of resource use as well as information of the healthcare system in the original study and the allocation of costs to various stakeholders (65).

By simultaneously providing recommendations for good practice in the economic evaluation of OHS interventions and discussing the methods and principles that underlie them, the present paper aimed to help researchers in conducting and reporting high quality trial-based economic evaluations. Such studies are expected to contribute to the development of a sound evidence base on the resource implications of OHS interventions (3;4), which is a necessary prerequisite for evidence-based practices occurring in occupational health (11). The present paper may also be helpful to consumers of this literature with understanding and critically appraising trial-based economic evaluations of OHS interventions, which might help improve the uptake of their results.

Acknowledgements
We would like to thank all authors of the example trial for the provision of their data.

REFERENCES

1. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. New York: Oxford University Press; 2005.
2. Burdorf A. Economic evaluation in occupational health - its goals, challenges, and opportunities. *Scand J Work Environ Health* 2007;33(3):161-164.
3. Tompa E, Verbeek J, van Tulder MW, de Boer A. Developing guidelines for good practice in economic evaluation of occupational health and safety intervention. *Scand J Work Environ Health* 2010;36(4):313-318.
4. Tompa E, Dolinschi R, de Oliveira C. Practice and potential of economic evaluation of workplace-based interventions for occupational health and safety. *J Occup Rehabil* 2006;16(3):367-392.
5. Uegaki K, de Bruijne MC, Lambeek L, Anema JR, van der Beek AJ, van MW, et al. Economic evaluations of occupational health interventions from a corporate perspective - a systematic review of methodological quality. *Scand J Work Environ Health* 2010 Jun;36(4):273-288.
6. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011;12(12):1031-1049.
7. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, et al. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health* 2012;38(5):393-408.
8. Tompa E, Culyer AJ, Dolinschi J. *Economic evaluation of interventions for occupational health and safety: Developing good practice*. New York: Oxford University Press; 2008.
9. Uegaki K, de Bruijne MC, Anema JR, van der Beek AJ, van Tulder MW, Van Mechelen W. Consensus-based findings and recommendations for estimating the costs of health-related productivity loss from a company's perspective. *Scand J Work Environ Health* 2007;33:122-130.
10. Korthals-de Bos I, van Tulder M, van Dieten H, Bouter L. Economic evaluations and randomized trials in spinal disorders: Principles and methods. *Spine* 2004;29(4):442-448.
11. van Dongen JM, Tompa E, Clune LA, Sarnocinska-Hart A, Bongers PM, van Tulder MW, et al. Bridging the gap between the economic evaluation literature and daily practice in occupation health: a qualitative study among decision-makers in the healthcare sector. *Implement Sci* 2013;8(57).
12. Glick HA, Doshi JA, Sonnad SS, Polsky D. *Economic evaluations in clinical trials.* New York, United States: Oxford University Press; 2007.
13. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A, et al. Good research practices forcost-effectiveness analysis alongside clinical trials: The ISPOR RCT-CEA Task Force Report. *Value Health* 2005;8(5):521-533.
14. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ* 2011;342: d1548.
15. Verbeek J, Pulliainen M, Kankaanpaa E. A systematic review of occupational safety and health business cases. *Scand J Work Environ Health* 2009;35(6):403-412.

16. Downey AM, Sharp DJ. Why do managers allocate resources to workplace health promotion programmes in countries with national health coverage? *Health Promot Int* 2007;22(2):102-11.
17. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313:275-283.
18. Goossens M, Rutten-van Mölken M, Vlaeyen J, van der Linden S. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol* 2000;53(7):688-95.
19. Drummond M, Sculpher M. Common methodological flaws in economic evaluations. *Med Care* 2005;43(7 Suppl):5-14.
20. Frick FD. Microcosting Quantity Data Collection Methods. *Med Care* 2009;47(7 Suppl 1):S76-S81.
21. Hakkaart - van Roijen L, Tan SS, Bouwmans CAM. *Handleiding voor kostenonderzoek. Methoden en standaardkostprijzen voor economische evaluaties in de gezondheidszorg. Geactualiseerde versie 2010* ed. College voor zorgverzekeringen: 2010.
22. Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: A critical review. *Soc Sci Med* 2011;72(2):185-192.
23. Krol M, Brouwer WBF, Rutten FFH. Productivity costs in economic evaluations: past, present, future. *PharmacoEconomics* 2013;31(7):537-549.
24. Koopmanschap MA, Rutten FFH. Indirect costs in economic studies: Confronting the confusion. *PharmacoEconomics* 1993;4(6):446-454.
25. Koopmanschap MA, Rutten FFH, van Ineveld BM, Van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ* 1995;14(2):171-189.
26. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D. The world health organization health and work performance questionnaire (HPQ). *J Occup Environ Med* 2003;45:156-174.
27. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE. Using the world health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46:S23-S37.
28. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res* 2005;5(1):23-28.
29. Van Roijen L, Essink-bot ML, Koopmanschap MA, Bonsel G, Rutten FFH. Labor and health status in economic evaluation of health care: The health and labor questionnaire. *Int J Technol Assess Health Care* 1996;12(03):405-415.
30. Wahlqvist P, Carlsson J, Stalhammar NO, Wiklund I. Validity of a Work Productivity and Activity Impairment questionnaire for patients with symptoms of Gastro-Esophageal Reflux Disease (WPAI-GERD) -- Results from a cross-sectional study. *Value Health* 2002;5(2):106-113.
31. Reilly MC, Bracco A, Ricci JF, Santoro J, Stevens T. The validity and accuracy of the Work Productivity and Activity Impairment questionnaire -- irritable bowel syndrome version (WPAI:IBS). *Aliment Pharmacol Ther* 2004;20(4):459-467.
32. Prasad M, Wahlqvist P, Shikiar R, Shih YT. A review of self-report instruments measuring health-related work productivity. *PharmacoEconomics* 2004;22(4):225-244.
33. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal [Internet]. [Cited 2013 August 21]. Available from: http://www.nice.org.uk/media/D45/1E/GuideToMethodsTechnologyAppraisal2013.pdf

34. Hamberg-van Reenen HH, Proper KI, van den Berg M. Worksite mental health interventions: a systematic review of economic evaluations. *Occup Environ Med* 2012;69(11):837-845.

35. Brazier J, Deverill M, Green C, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 2009;3(9):i-iv, 1-164.

36. Uegaki K. Economic evaluation of interventions for occupational health. Thesis (PhD), Vrije Universiteit Amsterdam, 2010.

37. Eysink PED, Hamberg-van Reenen HH, van Gool CH, Hoeymans N, Burdorf A. Meten van verloren arbeidsjaren door ziekte: Disease-Adjusted Working Years (DAWYs): Verkenning van een nieuwe maat. Bilthoven: RIVM; 2010. Report No.: 270244001.

38. Groeneveld IF, van Wier MF, Proper KI, Bosmans JE, Van Mechelen W, van der Beek AJ. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the contruction industry at risk for cardiovascular disease. *J Occup Environ Med* 2011;53(6):610-617.

39. Briggs AH, Gray AM. Power and sample size calculations for stochastic cost-effectiveness analysis. *Med Decis Making* 1998;18(2):S81-S92.

40. Gafni A, Walter SD, Birch S, Sendi P. An opportunity cost approach to sample size calculation in cost-effectiveness analysis. *Health Econ* 2008;17(1):99-107.

41. Gardiner JC, Sirbu CM, Rahbar MH. Update on statistical power and sample size assessments for cost-effectiveness studies. *Expert Rev Pharmacoeconomics Outcomes Res* 2004;4(1):89-98.

42. Al MJ, Van Hout BA, Michel BC, Rutten FFH. Sample size calculation in economic evaluations. *Health Econ* 1998;7(4):327-335.

43. Briggs A. Economic evaluation and clinical trials: size matters. *BMJ* 2000;321(7273):1362-1363.

44. Goossens M, Evers S, Vlaeyen J, Rutten-van Mölken M, van der Linden S. Principles of economic evaluation for interventions of chronic musculoskeletal pain. *European Journal of Pain* 1999;3(4):343-353.

45. Moher D, Hopewell S, Schultz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.

46. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.

47. Noble SM, Hollingworth W, Tilling K. Missing data in trial-based cost-effectiveness analysis: the current state of play. *Health Econ* 2012;21(2):187-200.

48. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Statist Med* 2011;30(4):377-399.

49. Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.

50. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? *BMJ* 2000;320(7243):1197-1200.

51. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. *Statist Med* 2000;19(23):3219-3236.

52. Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: An example from a randomized trial. *Statist Med* 1996;15(13):1447-1458.

53. Kelley K. The effects of non-normal distributions on confidence intervals around the standardized mean difference: Bootstrap and parametric confidence intervals. *Educational and Psychological Measurement* 2005;65(1):51-69.
54. Black WC. The CE plane: A graphic representation of cost-effectiveness. *Med Decis Making* 1990;10(3):212-214.
55. Briggs AH, O'Brien BJ, Blackhouse G. Thinking outside the box: recent advances in the analysis and presentation of uncertainty in cost-effectiveness studies. *Ann Rev Public Health* 2002;23:377-401.
56. Fenwick E, Marshall D, Levy A, Nichol G. Using and interpreting cost-effectiveness acceptability curves: an example using data from a trial of management strategies for atrial fibrillation. *BMC Health Serv Res* 2006;6(1):52.
57. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. *Health Econ* 2004;13(5):405-15.
58. Phillips JJ. Return on investment in training and performance improvement programs. 2nd ed. Burlington: Elsevier; 2003.
59. Stone PW. Return-on-investment models. *Appl Nurs Res* 2005 Aug;18(3):186-189.
60. Giffin SC. Dealing with uncertainty in the economic evaluation of health care technologies. Thesis (PhD), University of York, 2010.
61. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of healthcare technologies: the role of sensitivity analysis. *Health Econ* 1994;3:95-104.
62. Briggs A, Claxton K, Sculpher M. Decision modelling for health economic evaluation. New York: Oxford University Press; 2006.
63. Dehejia RH, Wahba S. Propensity score-matching methods for non-experimental causal studies. *Review of Economics and Statistics* 2002;84(1):151-161.
64. Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys* 2008 Feb 1;22(1):31-72.
65. Verbeek J, Pulliainen M, Kankaanpaa E, Taimela S. Transferring results of occupational safety and health cost-effectiveness studies from one country to another - a case study. *Scand J Work Environ Health* 2010;36(4):305-312.

Appendix 1: Core recommendations for trial-based economic evaluation in occupational health

Design of an economic evaluation
Types of economic evaluations
Perform various types of economic evaluations to inform all relevant stakeholders; e.g. cost-effectiveness analysis (CEA), cost-benefit analysis (CBA), cost-utility analysis (CUA).
Timing
Consider economic evaluation requirements during an early phase of the design of a trial.
Trial design
If possible, use randomization to allocate participants to study arms (i.e. (cluster-)RCTs).
Trial conditions should resemble daily practice as much as possible.
Perspective
Apply various perspectives to inform all relevant stakeholders.
The applied perspective(s) should be explicitly stated.
Analytic time frame
Ideally, the analytic time frame covers the entire period over which costs and consequences flow from the alternatives under study.
Identification, measurement, and valuation of costs
Collect all resources that may influence the overall costs related to the applied perspective(s).
Appropriate unit prices may vary between perspectives. Researchers should therefore ensure that unit prices reflect the true resource implications to the decision-maker(s) at hand.
Report aggregate costs, disaggregate resource use, and applied unit prices separately.

Analysis of an economic evaluation
Sample size
Ideally, economic outcomes are used in the sample-size calculation of a trial. If this is not possible, use estimation rather than hypothesis testing.
Adjusting for differential timing
Prices drawn from different years should be adjusted for inflation using consumer prices indices and the applied reference year should be explicitly stated.
Costs and consequences should be discounted using discount rates pertaining to the jurisdiction in which the economic evaluation is performed in order to adjust for time preferences of individuals.
Missing data
Use multiple imputation to impute missing values, particularly if ≥5% of data is missing.
Incremental analysis of costs and consequences
Incremental costs and consequences should be reported as differences in arithmetic means.
Use non-parametric bootstrapping to quantify precision of cost data.
Comparing incremental costs and consequences
The preferred method for comparing incremental costs and consequences depends on the kind of economic evaluation; i.e. incremental cost-effectiveness ratio (ICERs) for CEAs/CUAs, and Net Benefits, Benefit Cost Ratios, and/or Return On Investments for CBAs.
To quantify the uncertainty surrounding incremental cost-consequence estimates, use non-parametric bootstrapping techniques.
Use cost-effectiveness planes to graphically illustrate the uncertainty surrounding ICERs and cost-effectiveness acceptability curves to provide a summary measure of the joint uncertainty of costs and effects/utilities. For cost-benefit estimates, use 95% confidence intervals and/or the probability of financial return.
Sensitivity analysis
Perform a sensitivity analysis to test the robustness of results.
The ranges of values tested, and arguments for selecting these ranges, should be described.

9

Bridging the gap between the economic evaluation literature and daily practice in occupational health: a qualitative study among decision-makers in the healthcare sector

Johanna M van Dongen
Emile Tompa
Laurie Clune
Anna Sarnocinska-Hart
Paulien M Bongers
Maurits W van Tulder
Allard J van der Beek
Marieke F van Wier

Implementation Science 2013, 8:57

ABSTRACT

Background: Continued improvements in occupational health can only be ensured if decisions regarding the implementation and continuation of occupational health and safety interventions (OHS interventions) are based on the best available evidence. To ensure that this is the case, scientific evidence should meet the needs of decision-makers. As a first step in bridging the gap between the economic evaluation literature and daily practice in occupational health, this study aimed to provide insight into the occupational health decision-making process and information needs of decision-makers.

Methods: An exploratory qualitative study was conducted with a purposeful sample of occupational health decision-makers in the Ontario healthcare sector. Eighteen in-depth interviews were conducted to explore the process by which occupational health decisions are made and the importance given to the financial implications of OHS interventions. Twenty-five structured telephone interviews were conducted to explore the sources of information used during the decision-making process, and decision-makers' knowledge on economic evaluation methods. In-depth interview data were analyzed according to the constant comparative method. For the structured telephone interviews, summary statistics were prepared.

Results: The occupational health decision-making process generally consists of three stages: initiation stage, establishing the need for an intervention; pre-implementation stage, developing an intervention and its business case in order to receive senior management approval; and implementation and evaluation stage, implementing and evaluating an intervention. During this process, information on the financial implications of OHS interventions was found to be of great importance, especially the employer's costs and benefits. However, scientific evidence was rarely consulted, sound ex-post program evaluations were hardly ever performed, and there seemed to be a need to advance the economic evaluation skill set of decision-makers.

Conclusions: Financial information is particularly important at the front end of implementation decisions, and can be a key deciding factor of whether to go forward with a new OHS intervention. In addition, it appears that current practice in occupational health in the healthcare sector is not solidly grounded in evidence-based decision-making and strategies should be developed to improve this.

BACKGROUND

The extent to which organizations allocate their limited resources towards occupational health and safety interventions (OHS interventions), including both worksite health promotion and health and safety interventions, is driven by some combination of legal, financial, and moral factors (1,2). Among others, information on the costs and consequences of these interventions is therefore likely to be a valuable input into the decision of whether or not to implement or continue them. This is of particular importance in the healthcare sector, where OHS interventions focused on workers may be seen as redirecting resources away from higher priority ones more focused on patient care (3). Furthermore, rising healthcare expenditures, experienced by many developed countries, may pose another limitation to the resources available for OHS interventions in the healthcare sector (4,5).

To aid occupational health decision-makers, different types of economic evaluations are carried out. Cost benefit analyses (CBAs), also known as return-on-investment analyses, are conducted to provide insight into the net financial benefit or financial return by comparing incremental costs to incremental financial benefits of alternatives (6-9). Cost-effectiveness analyses (CEAs) are conducted to provide insight into the incremental costs of an intervention per additional unit of effect gained. In cost-utility analyses (CUAs), the incremental costs of an intervention are compared to its attributable health improvements measured in utilities (e.g., ‘quality adjusted life years’) (6).

During the last two decades, a growing number of articles has been published about the financial implications of OHS interventions (10), but their use and impact on day-to-day decision-making has not been adequately explored. However, as research indicates that results of economic evaluations of healthcare interventions for patients are rarely used among medical decision-makers (11-14), the use of economic evaluations among occupational health decision-makers is likely to be limited as well. Within the framework of evidence-based decision-making, it is essential that lessons learned from research are applied in practice. That is, continued improvements in occupational health can only be established if (implementation) decisions are based on the best available evidence. To ensure that this is the case, scientific evidence should meet the information needs of decision-makers. Specifically, disparities should be minimized between the way in which evidence is developed and presented and the way in which it is understood and used in daily practice (15). In addition, because a lack of expertise in health economics (specifically economic evaluation) was found to be an important barrier to the use of economic evaluations among medical decision-makers (11,13), it is of importance that occupational health decision-makers are equipped with an adequate skill set to interpret and use scientific evidence on the financial implications of OHS interventions.

Until now, studies have been undertaken to gain insight into evidence-based decision-making and possible ways to improve it among occupational health professionals (e.g., physicians, nurses) (16-18) and individual workers (19), but not among occupational health decision-makers. Therefore, as a first step in bridging the gap between the economic evaluation literature on OHS interventions and daily practice, the present study aimed to explore four issues: the process by which occupational health decisions are made; the importance given to the financial implications of OHS interventions; the sources of information used during the decision-making process; and occupational health decision-makers' knowledge about different economic evaluation methods.

METHODS

In-depth interviews with occupational health decision-makers in the Ontario healthcare sector were conducted to explore the process by which occupational health decisions are made and the importance given to the financial implications of OHS interventions. Structured telephone interviews were conducted to explore the sources of information used during the decision-making process and occupational health decision-makers' knowledge on economic evaluation designs. A qualitative approach was chosen, as little is currently known about these topics. Core categories of analytic foci have not yet been identified (20).

The present study was undertaken in collaboration with partners from the following organizations: the Public Services Health and Safety Association, the Ontario Nurses' Association, and the Ontario Hospital Association. At three meetings held over the course of the study, partner representatives provided input and feedback on data collection activities.

Ontario's occupational health and safety and healthcare system

Canada is a federation of ten provinces and three territories. As such, labour legislation and healthcare are provincial and territory level jurisdictions. Therefore, the OHS system (including regulation and insurance) and the healthcare system vary somewhat between provinces/territories, though there are many common features (21). In Ontario, regulatory responsibilities for the inspection and enforcement aspects of OHS lie with the 'Ministry of Labour' (MOL). Workers' compensation is administered by the 'Workplace Safety and Insurance Board' (WSIB), a monopoly, not-for-profit insurance provider that covers approximately 70% of Ontario's workforce. The WSIB is financed by payroll taxes levied on employers, with some variation among industries reflecting their different risk levels and accident experiences (i.e., industry specific rate groups). Within these rate groups, financial incentives are administered for organizations through experience ratings. Organizations with better-than-average safety records receive a rebate, whereas those with a worse safety records are levied a surcharge (22). The WSIB operates on a 'no fault' principle (i.e., compensation is paid no matter who is at fault) and generally covers healthcare costs and lost earnings associated with occupational injury and disease (21,22). Sickness absences that are not attributable to exposures at work are not compensable through workers' compensation, though the universal, publicly-funded healthcare system provides medical services to all Ontario residents for needed care. Employers may provide wage replacement benefits for these types of sickness absences. However, because these programs are not obligatory, only some employers offer them (21). In the light of this study, it is also important to mention that workplaces with 20 or more employees are required by law to have a 'Joint Health and Safety Committee' (JHSC). A JHSC is made up of worker and employer representatives that work together to identify and resolve health and safety problems in their workplace (21).

Ontario's universal, publicly-funded healthcare system is funded through transfer payments from the federal government and general taxes at the provincial level. Most hospitals are not-for-profit organizations that bill the 'Ministry of Health and

Long Term Care' (MOHLTC) for a wide range of medically necessary services (21,23). Long-term care (LTC), on the other hand, is provided by not-for-profit as well as for-profit facilities.

Recruitment and sampling

In order to focus our sampling efforts and to keep the scope of the study manageable, a subset of organizations from the Ontario healthcare sector was selected, namely hospitals and LTC facilities. Participants for the in-depth interviews and structured telephone interviews were selected by means of purposeful sampling. This sampling method enables researchers to use their own judgement in order to select individuals who could provide in-depth information relevant to the research questions. Project partners assisted in identifying such individuals. Additionally, participants were selected by means of snowballing: i.e., participants were asked whether they knew other people who they thought could provide relevant information about the occupational health decision-making process (20). Participants had to be employees of an Ontario-based hospital or LTC facility that were either responsible for the daily occupational health operations or senior staff members. To reduce the risk of biased responses, decision-makers who participated in the in-depth interviews were excluded from participation in the structured telephone interviews. All participants were informed about the study purpose, were reassured of confidentiality, and provided written informed consent. Study details were approved by the University of Toronto's Office of Research Ethics.

In-depth interviews

In-depth interviews took place from June 2011 to August 2011 during an in-person or telephone meeting arranged at a time and location convenient to the participants. Interviews lasted on average 47 minutes (range: 12 to 116 minutes) and were conducted by two or three researchers (ET, AS-H, LC). One researcher was responsible for asking questions, whereas the other(s) took field notes and probed areas requiring more explanation. An interview protocol was used including questions and prompts. First, short questions were asked regarding the employment and workplace characteristics of the interviewee (e.g., job description, years of relevant work experience, facility size). Subsequently, open-ended questions were asked to explore the decision-making process and the importance given to the financial implications of OHS interventions. The first open-ended question was 'How does your organization go about starting and implementing an OHS intervention?' Possible follow-up questions or prompts were 'Can you describe how you evaluate OHS interventions?' 'What type of information helps move a plan forward?' 'How do you prioritize between alternatives?' 'How does cost-benefit/cost-effectiveness fit into your decision-making process?' Throughout the interview, participants were asked to illustrate their answers by giving examples of recent program implementation and/or continuation decisions, including those concerning both small versus large and mandated versus non-mandated OHS interventions. Among others, the participants' examples concerned workplace violence, return to work, participatory ergonomics, and health education programs. Question prompts were slightly revised throughout the data collection process based on the researchers' sense of what additional information would be useful and the participant's position within the organization. The final topic list is provided in Additional file 1. Analytic field notes were written after each interview by one researcher (LC), including thoughts about the dynamics of the encounter and issues that may be relevant at the analytical stage (20). All interviews were recorded and transcribed verbatim. After 15 interviews, the analytic field notes indicated that no new findings emerged (i.e., data saturation). To be sure that data saturation was indeed reached, three additional interviews were conducted. As no new findings emerged from these interviews as well, data collection was terminated after 18 interviews.

Data analysis: in-depth interviews

Data were analyzed using the constant comparative method, in which each item is checked or compared with the rest of the data to inductively establish analytical categories (24,25). First, analytic field notes and transcripts were read to get a general understanding of the concepts under study and to get some insight into the dynamics of the interviews. Using Nvivo version 10 (QSR international, Burlington, USA), transcripts were subsequently open-coded by one researcher (JvD). That is, transcripts were read line by line and relevant passages were selected and coded, often by using the participants' own words. Interview codes included both 'descriptive' (i.e., within the immediate domain of the interview questions) and 'analytic' (i.e., emerging and overarching) themes (20). Throughout the coding process, conscious efforts were made to detect further examples of previously identified themes and, if applicable, to identify new ones (24-27). Subsequently, similar codes were grouped into so-called analytical categories, and the analytical categories' properties were explored as well as the relationships between those categories (25). At various meetings held over the course of the data analysis process, identified codes, identified analytical categories, and interpretations of the data were checked and discussed with the interviewers (AS-H, ET, LC) to enhance the robustness of the findings. In all cases, consensus was reached through discussion.

Structured telephone interviews

Structured telephone interviews were conducted by one researcher (AS-H) from November 2011 to February 2012 and lasted on average 27 minutes (range: 15 to 60 minutes). First, short questions were asked regarding the employment and workplace characteristics of the interviewee (e.g., job description, facility size). Subsequently, participants were asked to what extent external sources of information were consulted when exploring whether a future intervention was worthwhile (i.e., always, sometimes, never), and if so, what types of sources. Also, a list of inputs/costs and outcomes/consequences in economic evaluations of OHS interventions was provided to the participants, and they were asked to what extent these inputs/costs and outcomes/consequences were considered during the decision-making process (i.e., always, sometimes, never). The list of inputs/costs and outcomes/consequences was derived from a previous study of one of the authors (unpublished data). Subsequently, participants were asked whether they were familiar with CBA, CEA, and CUA, and, if so, whether they could define these economic evaluation designs, whether they previously received training in economic evaluation-related topics, and whether they wanted to acquire more knowledge in this field. An overview of the structured interview items pertaining to research questions is provided in Additional file 2. All telephone interviews were recorded. Data analysis: structured telephone interviews By listening to the audiotapes, descriptive statistics were prepared by two researchers (AS-H, JvD). Inputs/costs and outcomes/consequences of economic evaluations were regarded as 'commonly considered' if they were 'always' considered during the decision-making process by more than 50% of the participating healthcare facilities. Definitions of the various economic evaluation designs were scored as 'correct' if they included some variation of the following information: CBA, a comparison of costs and benefits, in which both are expressed in monetary terms; CEA, a comparison of costs and outcomes, in which costs are expressed in monetary terms and outcomes in natural units; and CUA, a comparison of costs and utilities, in which costs are expressed in monetary terms and utilities (e.g., health improvements) in terms of 'quality adjusted life years,' or possibly some variant, such as 'disability adjusted life years' (6). In all other cases, they were scored as 'incorrect.'

RESULTS

In-depth interviews

Participants

Eighteen in-depth interviews were conducted with a total of 19 participants (i.e., one interview was conducted with two participants). Of them, 11 worked at a hospital and eight at a LTC facility. Twelve were female and seven male. Fifteen were responsible for the daily occupational health operations and four were senior staff members (Table 1).

The process by which occupational health decisions are made

In general, the process by which occupational health decisions are made can be subdivided into three stages: initiation stage, pre-implementation stage, and implementation and evaluation stage (Figure 1).

Table 1: Characteristics of the study population

	In-depth interviews	Structured telephone interviews
Participants [n.]	**19**	**28**
LTC [n. (%)]	**8 (42)**	**1 (4)**
Female [n. (%)]	7 (88)	1(100)
Job description [n. (%)]		
OHS operations	6 (75)	1 (100)
Senior staff members	2 (25)	0 (0)
Years of relevant work experience [mean (SD)]	16.6 (7.8)	N.A.
Hospitals [n. (%)]	**11 (58)**	**27 (96)**
Female [n. (%)]	5 (46)	21 (78)
Job description [n. (%)]		
OHS operations	9 (81)	26 (96)
Senior staff members	2 (19)	1 (4)
Years of relevant work experience [mean (SD)]	7.6 (2.8)	N.A.
Interviews [n.]	**18**	**25**
LTC [n. (%)]	**7 (39)**	**1 (4)**
Size [n. (%)]		
<250 employees	3 (43)	0 (0)
250-999 employees	4 (57)	1 (100)
Type [n. (%)]		
Public (not for profit)	4 (57)	1 (100)
Private (for profit)	3 (43)	0 (0)
Hospital [n. (%)]	**11 (61)**	**24 (96)**
Size [n. (%)]		
<250 employees	0 (0)	3 (13)
250-999 employees	3 (27)	6 (25)
1000-1999 employees	1 (9)	5 (21)
2000-9999 employees	5 (46)	7 (29)
>10000 employees	2 (18)	3 (13)
Type [n. (%)]		
Public (not-for-profit)	11 (100)	24 (100)
Private (for-profit)	0 (0)	0 (0)

Abbreviations: n: number; OHS: Occupational Health and Safety; LTC: Long-Term Care facility; N.A.: Not Available

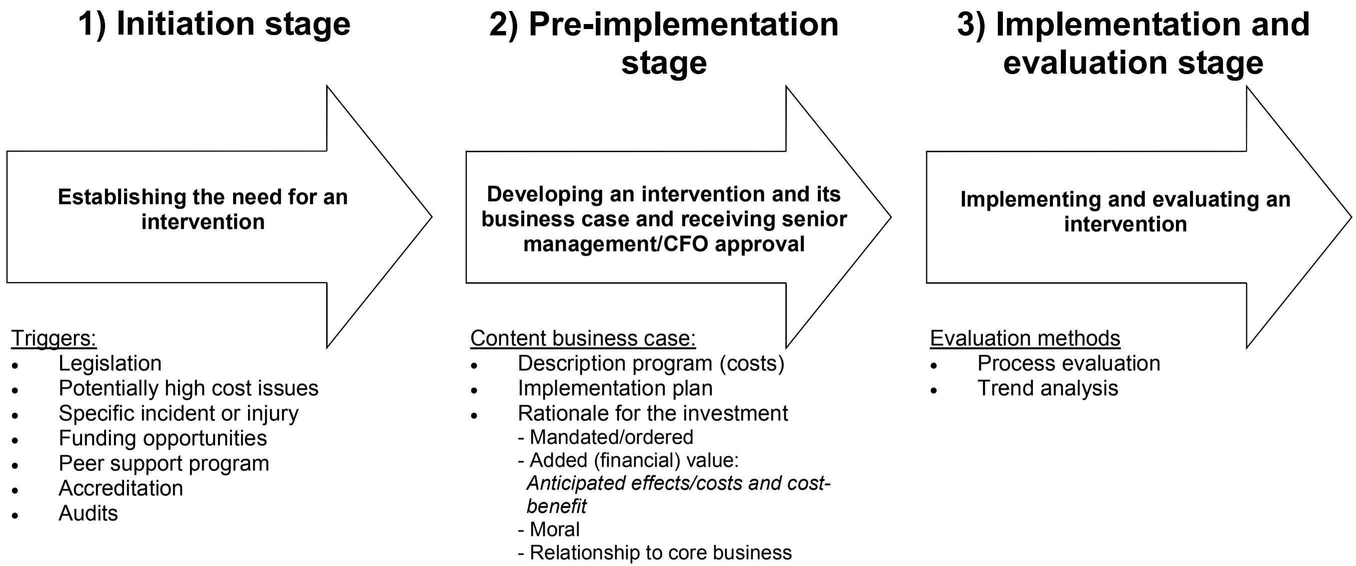

Figure 1: Stages of the occupational health decision making process
Abbreviations: CFO: Chief Financial Officer

Initiation stage

During the first stage of the decision-making process, the need for an intervention is mostly established by employees responsible for the daily occupational health operations and is generally triggered by one or more of the following factors:

- Legislation: Legislation is given top priority and the implementation of many interventions is driven by regulatory requirements. As one participant noted, 'First and foremost, obviously there's the result of legislation … that we have to act upon.' Legislation, however, only relates to health and safety interventions and not to worksite health promotion programs.
- Potentially high cost issues: The need for an intervention may also be triggered by potentially high cost issues within the healthcare facility or sector. Internal statistics, such as incident report trends and sick leave data, are collected in all facilities through a variety of methods such as note taking and various software applications. In reviewing the acquired data, priority is given to finding ways to reduce costs through identifying high risk injury types or high exposure settings. As one participant noted, 'We do collect incident injury data, employee injury data monthly.… We then put it into a quarterly graph and look at possible trends … a lot of initiatives are based on the incident trends.' Some facilities benchmark these statistics against those of similar facilities to help put them into perspective. High cost issues within the facility are also identified by conducting on-site risk assessments and needs assessments among employees. External reports and scientific evidence are consulted by some facilities to identify high cost issues within the healthcare sector.
- Specific incident or injury: After a specific incident or injury, interventions might be requested by the JHSC and/or senior management or ordered by the MOL. MOL orders are the result of so-called 'significant incidents' (i.e., an employee is critically injured or killed at the workplace) that organizations have to respond to after MOL inspection. As one participant noted, '…that [intervention] came about because of an order from the Ministry of Labour. We had a worker that … fractured her arm. It was a critical injury, so we got the order.'

- Funding opportunities: Occasionally, the federal or provincial government provides funding opportunities for OHS interventions. Many facilities apply for such grants, as is indicated by the following quote: 'The government provided funding and, of course, we jumped on it like everybody else.' Facilities have to spend these grants on a specific type of intervention or the reduction of a specific adverse health or safety outcome (e.g., workplace violence). Another way in which facilities make use of funding opportunities is by participating in external research projects.
- Peer support program: In Ontario, healthcare facilities, as well as other types of organizations, may participate in peer-support programs called Safety Group Programs. This is a performance-based rebate program developed by the WSIB. Organizations can join a safety group consisting of their peers to learn more from each other's occupational health experiences. In the program, they are obliged to identify and implement five selected OHS interventions each year. A discount on insurance premiums is given for participating in these groups. Additionally, as one might expect, the successful implementation of OHS interventions may have positive implications for their insurance premiums, given that premiums are experience rated (28).
- Accreditation: The need for an OHS intervention is sometimes identified during the hospital accreditation process. As one participant noted, 'Initiatives come through quality improvement that we deal with through our annual accreditation processes.' While not mandatory, almost all of Ontario's hospitals and LTC facilities opt to go through regular accreditation reviews. The accreditation process is intended to ensure that healthcare facilities are meeting a common set of standards. Accreditation occurs on a three-year cycle and includes the measurement of various performance indicators (e.g., patient safety and quality of care, infection prevention and control, medication management, organizational culture) (29).
- Audits: OHS interventions are sometimes triggered by (upcoming) internal (e.g., by the JHSC) or external audits (e.g., by the MOL inspectorate). External audits may result in orders, which oblige a facility to address particular health hazards within a particular time period. With extreme health hazards, and repeated violations, a financial penalty may be imposed.

Pre-implementation stage

The second stage of the decision-making process, is generally characterized by the development of the intervention as well as its business case in order to receive approval for its implementation from senior management.

Based on the previously identified need(s), interventions are developed by employees responsible for the daily occupational health operations in consultation with various external (e.g., similar facilities, safety group, consultants) and internal (e.g., JHSC) sources of information. Sometimes, a small on-site pilot study is conducted to compare various program options, especially in the case of equipment purchasing decisions. Depending on its size, interventions are either developed by one person or a working group. In most cases, senior management approval is needed before an intervention may be implemented. To convince them of the importance of a specific intervention, a so-called 'business case' is developed. These business cases generally include one or more of the following items: a description of the program and its costs, and sometimes that of alternatives, a program implementation plan, and a rationale for the investment. Various types of rationales emerged from the data (note that these rationales are linked to the triggers of OHS interventions, except for the moral rationale):

- Mandated/ordered: The facility has to implement a certain intervention to comply with legislation, to deal with a specific incident (e.g., after a MOL order), or to meet accreditation standards. As one participant noted, 'They [senior management] always want to know, well, do we have to do it?'.
- Added (financial) value: Implementation of the intervention may produce added value to the facility. In some cases this value is financial. For example, implementation may reduce the incidence of high cost issues (e.g., high cost accidents, sick leave), leading to a worker's compensation insurance rebate, or reduction in replacement staff costs. Implementation may also improve a healthcare facility's reputation which, in turn, may affect its staff recruitment and retention abilities as well as its ability to raise charitable funds.
- Moral: The intervention may be implemented for moral reasons. As one participant noted, 'We actually go through the moral imperative about why it is not appropriate to injure people.'

- Relationship to core business: Implementation of the intervention may improve the core business of the healthcare facility, namely patient care. Participants specifically indicated that they made the connection between OHS interventions and patient care, because their facility receives funds for the provision of patient care activities and not directly for occupational health.

As part of the 'added (financial) value' rationale, an overview of the anticipated effects, benefits, and/or cost-benefit are often presented. Most participants indicated that ex-ante CBAs formed the basis of a business case, but these cases are very high level and stylized in nature. That is, they are not supported by rigorous internal statistics and/or scientific evidence. To illustrate, one participant described the content of a CBA as follows:

> 'Maybe the costs with the WSIB, the modified work etcetera … may have been $60,000 for the year versus the cost of equipping the unit, which would have been maybe $10,000 or $12,000. And so, obviously, we wanna do something like that.'

This finding is also supported by the following comment of a participant with work experience in both the private and public sector:

> 'I know in the private sector when we were doing a cost-benefit analysis on the purchase of a piece of equipment, it was much more quantitative …. here, it seems to be a little more subjective and I don't really understand why that is.'

Cost-comparison analyses of various program options are also performed to identify the least costly alternative, but these analyses are mainly conducted for mandated interventions.
After completion, business cases are taken forward to senior management for approval. In most cases, a final decision is made in consultation with the chief financial officer, especially in the case of expensive interventions. The specific strategies used by the senior management to make and prioritize occupational health decisions are

not transparent. Most operational personnel were unclear about the process, while others described it as subjective. However, the approval process for mandated/ordered interventions and those that require minimal financial investments is less demanding (in terms of information and time required to make the business case) than that of non-mandated and more costly ventures and they are therefore more quickly approved. In LTC facilities, the approval process is not always as complex as described above. For example, when a need for an intervention is established, operational personnel may speak directly to the chief executive officer or director who has the ultimate responsibility for the organization. This is because LTC facilities are generally smaller than hospitals and have a flatter hierarchy.

Implementation and evaluation stage
During the third stage of the decision-making process, an OHS intervention is implemented and evaluated by performing a process evaluation and/or trend analysis. Process evaluations are generally aimed at exploring program execution, and employee satisfaction, compliance, attendance, and/or awareness. Process evaluation data is gathered through surveys, observations, and/or verbal feedback. Trend analyses are conducted to get an indication of the intervention's effectiveness. Therefore, various intervention-related measures, such as accident frequency or sickness absence rates, are collected from company records. Analyses explore whether their frequency decreased after implementation. Some participants, however, doubted the validity of results; either the integrity of their data, or concerns that observed trends were caused by factors other than the intervention. The latter is evident from the following comment:

> 'So overall, we did see a reduction, but it's hard to say whether or not that reduction was because the weather had gotten warmer or because it was just a coincidence. We're not too sure yet.'

Most participants indicated that far from all interventions are subjected to such evaluations and that ex-post CBAs are generally not performed. Their most important explanation for this was that they lacked the resources (time, money, and ability) to do so. As one participant noted, ‘The reason why we don’t do those evaluations on an ongoing basis is because it would cost money to do so.’ Other explanations were: lack of good data, and lack of economic evaluation skills.

The importance given to the financial implications of OHS interventions

Almost all participants indicated that information on the financial implications of OHS interventions is of great importance during the decision-making process, especially their cost-benefit. This is due to the fact that investing in those kinds of interventions literally affects a healthcare facility’s ability to provide patient care, as they have a tight budget (even the for-profit LTC facilities) and all occupational health expenses appear to take away from the patient care budget. Another reason for its importance is that healthcare facilities are mostly publicly funded. As one participant noted, ‘What makes this industry very different is the object. The politics, the perception that, because this is publicly funded…., the need not to waste is greater than on the other side.’ Information on the financial implications of mandated/ordered interventions seems less important. As one participant noted, ‘For our other health and safety programs, really, I would say that the only cost-benefit is that we don’t get fined.’

Structured telephone interviews

Participants

Twenty-five structured telephone interviews were conducted with a total of 28 participants. Of them, 27 worked at a hospital and one at a LTC facility. Twenty-two interviews were conducted with one participant and three with two participants. Twenty-two were female and six male. Twenty-seven were responsible for the daily occupational health operations and one was a senior staff member (Table 1).

The (sources of) information used during the occupation health decision-making process

Sources of information: To explore whether a future intervention is worthwhile, external sources of information were ‘always’ consulted during the decision-making process at 10 facilities (40%), ‘sometimes’ at 13 (52%), and ‘never’ at two (8%). Peer healthcare facilities were the principal external source of information (n = 23; 92%) and were either contacted directly or via a Safety Group Program. At five facilities (20%), participants indicated that they searched for scientific evidence on programs similar to those under consideration for implementation. Other external sources of information were: employers' associations (28%), the government (MOL/MOHLTC) (20%), the WSIB (20%), vendors (8%), law firms (4%), safety specialists (4%), and unions (4%).

Inputs/costs and outcomes/consequences considered during the decision-making process: A broad range of inputs/costs was considered during the decision-making process, though hard cost items (e.g., cost of equipment purchases, equipment installation, employee training) were more commonly considered than softer cost items (e.g., cost of administration, planning, promotion, and evaluation). This was mainly due to the fact that the latter were often considered as part of the regular day-to-day responsibilities of the affected departments (Table 2). A broad range of outcomes/consequences was considered as well. The number of injuries, illnesses, and sickness absences were considered at all facilities. Other commonly considered outcomes/consequences were days lost due to injuries or illnesses, accommodating injured or ill workers, quality of care and patient safety, employer workers' compensation insurance premiums, and meaningful return to work. In contrast, items such as impact on productivity (i.e., presenteeism), attraction and retention (i.e., turnover), worker replacement expenses, and labour relations climate were less commonly considered (Table 2).

Table 2: Inputs/costs and outcomes/consequences considered during the decision-making process

Items	How often are these items considered during the decision making process?		
Inputs (Costs)	**Always [n. (%)]**	**Sometimes [n. (%)]**	**Never [n. (%)]**
Health and safety staff time	11 (44)	10 (40)	4 (16)
Training the worker	15 (60)	10 (40)	0 (0)
Planning, promotion and evaluation	7 (28)	12 (48)	6 (24)
Equipment purchases	23 (92)	2 (8)	0 (0)
Administration	6 (24)	14 (60)	5(20)
Equipment installation	17 (68)	8 (32)	0 (0)
Ongoing equipment repair and maintenance	12 (48)	10 (40)	3 (12)
Professional / consultant fees	18 (72)	5 (20)	2 (8)
Ongoing supplies	14 (56)	10 (40)	1 (4)
Outcomes (Consequences)	**Always [n. (%)]**	**Sometimes [n. (%)]**	**Never [n. (%)]**
Number of injuries, illnesses, sickness absences	25 (100)	0 (0)	0 (0)
Days lost due to injuries, illnesses, and sickness absences	22 (88)	2 (8)	1 (4)
Quality of care and patient safety	16 (64)	7 (28)	2 (8)
Attraction and retention	7 (28)	16 (64)	2 (8)
Accommodating injured or ill workers[1]	14 (56)	10 (40)	1 (4)
Impact on productivity	12 (48)	12 (48)	1 (4)
Worker replacement expenses	10 (40)	11 (44)	4 (16)
Employer workers' compensation insurance premiums	15 (60)	7 (28)	3 (12)
Employer claims management expenses	11 (44)	9 (36)	5 (20)
Overtime payment	8 (32)	12 (48)	5 (20)
Meaningful return to work[2]	14 (58)	8 (33)	2 (8)
Labour relations climate[2]	12 (50)	11 (46)	1 (4)

Abbreviations: n.: number
[1] Provision of accommodated work to injured workers to reduce the duration of work absence.
[2] One participant did not answer this question, as he/she was unsure

Occupational health decision-makers' knowledge of different economic evaluation designs

Most participants (93%) were familiar with the concept of CBA and many (72%) were able to give a correct definition. For them, it meant comparing the costs of implementing an intervention with the financial consequences it was expected to bring:

> 'It's where you factor in all the costs of the intervention … Direct costs associated with whatever it is that you are trying to purchase … On the benefit side you would still put it into dollars, but it would be attributing things like reduced sick time and reduced injury costs. So both sides of the equation and then you would come out with … a positive or negative return on your investment.'

CBAs were undertaken at most facilities (92%), and formed the basis of a business case. These analyses were generally performed from the employer's perspective and not from the worker's or societal perspective. Most participants (71%) indicated that they were familiar with the concept of CEA, but few (11%) were able to give a correct definition. Most of them thought it to be synonymous with on-going monitoring and evaluation and not necessarily a comparison of costs with outcomes measured in natural units:

> 'Cost-effectiveness is looking at how effective an initiative is in terms of … is the outcome what we anticipated it to be.'

> 'Looking at the outcomes to determine whether what you anticipated to be the expected outcome, did you really reach those … But it's a complete guess.'

Others thought it to be synonymous to CBA:

> 'Is the investment of money and time worth the effort and that we will have a return on investment.'

Most participants were not familiar with the concept of CUA; only one indicated that he had heard of it, but was not able give a correct definition. Some participants tried to guess the definition, but most of them thought it to be an evaluation of the utilization (uptake) of an intervention:

> ‘Well utility is utilization, so I guess.... if we spent 25,000 Dollars..... We want to know whether they [the new equipment] are actually being used.’

Few participants (36%) received training in an area related to economic evaluations, such as a business proposal course, certified accountant training for financial planning, and business case sessions on program evaluations. When asked whether they were interested in receiving training, 79% (n = 22) expressed interest, 18% (n = 5) were not interested, and one (4%) was uncertain.

Of those not interested, lack of interest was expressed because they already considered themselves familiar with economic evaluation methods, were already adequately skilled at making informed decisions, or they considered their facility too small for such training to be of added value. Participants who expressed interest in receiving training felt that it would provide them with the skills required to make more informed implementation decisions and to undertake better evaluations themselves.

When asked what topic they wanted to learn more about, most of the participants (77%) indicated that they wanted to acquire more knowledge on CBA and/or writing a business case. Some also indicated that they wanted to acquire more knowledge about CEA and CUA after these terms were briefly explained to them (CEA: 36%, CUA: 36%).

DISCUSSION

As a first step in bridging the gap between the economic evaluation literature and daily practice in health and safety, this study aimed to provide insight into the occupational health decision-making process and information needs of decision-makers in the Ontario healthcare sector. Results showed that this process can be subdivided into three stages: initiation stage, during which the need for an intervention is established; pre-implementation stage, during which an intervention and its business case are developed in order to receive senior management approval; and implementation and evaluation stage, during which an intervention is implemented and evaluated. In line with previous research (1,2), organizations were found to invest in OHS interventions for legal, financial, and moral reasons, and information on their financial implications was found to be of great importance to the decision-making process. Results also indicated that occupational health decisions are currently not being made in an evidence-based manner. That is, scientific evidence on the (financial) implications of OHS interventions was found to be rarely consulted and sound ex-post program evaluations were hardly ever performed (30-32). Also, there seemed to be a need to advance the decision-makers' economic evaluation skill set, as they were either not familiar with economic evaluation methods or had only a modest amount of training in this area. Therefore, strategies should be developed to overcome these issues.

Strengths and limitations

Important strengths of the present study are its explorative and qualitative design. This enabled us to be one of the first to provide detailed insight into the extent to which occupational health decisions are made in an evidence-based manner, as well as to identify the information needs of occupational health decision-makers. By simultaneously exploring both issues, we were able to provide some initial clues to occupational health researchers as to how they might better frame and disseminate their studies to ensure uptake in healthcare organizations as well as organizations in other sectors.

Several methodological limitations deserve attention as well. First, the present study was restricted to a single industry, in a single region of one country. This was done to keep the scope of the study manageable, but likely bears on the generalizability of its results. For example, one might expect that occupational health decisions are made differently in sectors where budgets for occupational health are less tight. Furthermore, occupational health decision-making processes likely vary between jurisdictions (e.g., countries with different OHS and/or healthcare systems), in particular regarding the triggers of OHS interventions. Therefore, future studies should be conducted to explore the extent to which the present findings are generalizable beyond the healthcare sector and beyond Ontario, Canada. Second, due to the qualitative design of the present study, a limited number of interviews were conducted. However, as the healthcare facilities represented by the participants, in aggregate, employ a large number of Ontario healthcare workers, the extent to which this reduced the external validity of the present findings is probably small. Third, data were obtained through interviews, which may have caused 'social desirability bias.' For example, because participants were aware of the fact that they were interviewed by occupational health researchers, they may have overstated their use of scientific evidence as well as the quality of their decision-making process.

Improving evidence-based practice in occupational health

Sackett et al. (2000) identified two separate stages for evidence-based practice. The first stage concerns the generation of scientific evidence and relies heavily on the academic body of a profession. The second stage concerns the use of scientific evidence into daily practice (33,34). To improve the quality of the occupational health decision-making process, both stages should be addressed.

When generating scientific evidence, occupational health researchers should ensure that their products are in line with the information needs of occupational health decision-makers, as it is unrealistic to expect decision-making processes to be redesigned around research priorities (15,35). The present study provided some initial clues as to what these information needs are. For example, process evaluation data and information on the interventions' impact on corporate reputation and business results were found to be of interest to decision-makers. In addition, CBAs performed from the employer's perspective formed the basis of business cases for occupational health. Within these analyses, hard cost items (e.g., equipment costs, employee training costs) were of particular importance and benefits were commonly expressed in terms of reduced injury-, illness-, sickness absence-, and/or workers' compensation-related costs. In line with previous research (2), data on staff retention and productivity were considered relevant but not commonly used. The latter could probably be explained by the fact that these types of benefits are generally viewed as harder to identify and hard to monetize. Researchers, especially those conducting clinical trials, should be encouraged to report on the employer's cost-benefit of OHS interventions as well as their impact on corporate reputation and business results. This, however, does not negate the value of other types of economic evaluations. Various potential program benefits (e.g., job satisfaction, corporate reputation) and health outcomes are hard to monetize and may therefore not be included in a CBA. A possible way to deal with these so-called 'intangible benefits' is to conduct a CEA to estimate the incremental costs per 'intangible benefit' gained (8). In addition, the adoption of the societal perspective may provide insight into the distribution of costs and benefits between various stakeholders and thereby allows for bargaining between them (6). The latter is of particular importance in countries with universal healthcare coverage or dual-payer systems, because employers bear most of the costs of OHS interventions, while the government and/or healthcare system reaps a large part of its benefits (i.e., reduced medical spending) (1).

In daily practice, decisions have to be made within a limited time frame and many decision-makers lack the skills to determine what evidence is most reliable, and what evidence should be considered, under which circumstances (36). It is therefore advisable to provide busy decision-makers with critical summaries of published studies (37). Within the occupational health research field, systematic reviews are increasingly being conducted to critically appraise and summarize the current evidence on the (financial) implications of various OHS interventions. These systematic reviews, however, do not seem to be used in daily practice. This is probably due to the fact that many decision-makers lack the time and skill set required to read and understand these systematic reviews as well. Additionally, most of these reviews are published in scientific journals not well known or inaccessible to occupational health decision-makers. Therefore, it is important to transmit systematic review results to decision-makers in easy-to-use formats (35). This may be accomplished by publishing review fact-sheets in journals and newsletters more familiar to occupational health decision-makers and/or by distributing them through governmental institutes, employers' associations, and workers' compensation insurance boards. In addition, (more) best practice guidelines could be developed in which scientific evidence is summarized, and if unavailable, supplemented by expert opinions (36). To improve evidence-based practice, it is also important to educate decision-makers about economic evaluation methods, as well as the need and importance of integrating scientific evidence into day-to-day occupational health decision-making processes. The former is of particular importance, as many decision-makers were not familiar with various economic evaluation designs, which may not only limit the use of such studies in daily practice, but may also lead to misinterpretations of their results. Occupational health decision-makers may be educated through a variety of formal and informal means, including the development of handbooks and workshops on economic evaluation methods and evidence-based practice, integrating these topics into management and/or occupational health training programs, and involving occupational health decision-makers in the process of commissioning studies (38,39). Participation in scientific studies is namely closely linked with the uptake of their results (37) and may simultaneously lead to an enhanced economic evaluation skill set. Another option would be for researchers to develop hands on program evaluation software applications, so that decision-makers can conduct their own ex-ante or ex-post program evaluations in a relatively non-time consuming way. Additionally, more evidence is needed on the merits of evidence-based decision-making in occupational health, specifically, evidence that demonstrates that it improves organizations' performance. More economic evaluations of OHS interventions are needed to build a solid evidence base in order to support evidence-based practices in occupational health (34).

Implications for future research

Researchers, especially those conducting clinical trials, are recommended to report on the cost-benefit of OHS interventions from the employer's perspective as well as other perspectives. The impact of OHS interventions on operational outcomes and corporate reputation are two important pieces of information of occupational health decision-making. In the healthcare field, patient outcomes are particularly important. In addition, future research should focus on the extent to which the present findings are generalizable to other jurisdictions and on the effectiveness of possible strategies to improve evidence-based decision-making in occupational health.

Conclusion

This exploratory qualitative study on the occupational health decision-making process in healthcare suggests that the process generally consists of three stages; initiation stage: establishing the need for an intervention; pre-implementation stage: developing an intervention and its business case; and implementation and evaluation stage, implementing and evaluating an intervention. Organizations invest in occupational health for legal, financial, and/or moral reasons. Financial information is particularly important at the front end of implementation decisions, and can be a key deciding factor of whether to go forward with a new OHS intervention. In addition, it appears that current practice in occupational health in the healthcare sector is not solidly grounded in evidence-based decision-making and strategies should be developed to improve this.

Acknowledgements

The authors wish to thank all representatives of the Public Services Health and Safety Association, the Ontario Nurses' Association, and the Ontario Hospital Association that provided input and feedback on the data collection activities as well as the participants to the in-depth and structured telephone interviews. Funding for this study was received from the 'Workplace Safety and Insurance Board.' Additional support was provided through a personal travel grant awarded by the EMGO+ Institute for Health and Care Research.

REFERENCES

1. Downey AM, Sharp DJ: Why do managers allocate resources to workplace health promotion programmes in countries with national health coverage? *Health Promot Int* 2007, 22:102–111.
2. Miller P, Haslam C: Why employers spend money on employee health: Interviews with occupational health and safety professionals from British Industry. *Safety Science* 2009, 47:163–169.
3. Niven KJM: A review of the application of health economics to health and safety in healthcare. *Health Policy* 2002, 61:291–304.
4. Stuart N, Adams J: The sustainability of Canada's healthcare system: a framework for advancing the debate. *Healthc Q* 2007, 10:96–103.
5. OECD: Health: spending continues to outpace economic growth in most OECD countries. http://www.oecd.org/document/38/0,3746,en_2157 1361_44315115_482 89894_1_1_1_1,00.html.
6. Drummond M, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL: Methods for the economic evaluation of health care programmes. 3rd edition. New York: Oxford University Press; 2005.
7. Phillips JJ: Return on Investment in Training and Performance Improvement Programs. 2nd edition. Burlington: Elsevier; 2003.
8. van Dongen JM, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, Van Mechelen W, et al: Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011, 12:1031–1049.
9. Verbeek J, Pulliainen M, Kankaanpää E: A systematic review of occupational safety and health business cases. *Scand J Work Environ Health* 2009, 35:403–412.
10. Leigh JP: Expanding research on the economics of occupational health. *Scand J Work Environ Health* 2006, 32:1–4.
11. Christiane H, Graf von der Schulenburg JM: The influence of economic evaluation studies on decision making. A European survey. *Health Policy* 2000, 52:179–192.
12. Eddama O, Coast J: A systematic review of the use of economic evaluation in local decision-making. *Health Policy* 2008, 86:129–141.
13. Eddama O, Coast J: Use of economic evaluation in local health care decision-making in England: A qualitative investigation. *Health Policy* 2009, 89:261–270.
14. Zwart-van Rijkom JEF, Leufkens HGM, Busschbach JJV, Broekmans AW, Rutten FFH: Differences in Attitudes, Knowledge and Use of Economic Evaluations in Decision-Making in The Netherlands: The Dutch Results from the EUROMET Project. *Pharmacoeconomics* 2000, 18(2):149–160.
15. Nutbeam D: Achieving 'best practice' in health promotion: improving the fit between research and practice. *Health Educ Res* 1996, 11:317–326.
16. Hugenholtz N, Nieuwenhuijsen K, Sluiter J, van Dijk F: Do knowledge infrastructure facilities support Evidence-Based Practice in occupational health? An exploratory study across countries among occupational physicians enrolled on Evidence-Based Medicine courses. *BMC Health Serv Res* 2009, 9:18.

17. Schaafsma F, Hulshof C, van Dijk F, Verbeek J: Information demands of occupational health physicians and their attitude towards evidence-based medicine. *Scand J Work Environ Health* 2004, 30:327–330.

18. Schaafsma F, Hulshof C, de Boer A, Hackmann R, Roest N, van Dijk F: Occupational physicians: what are their questions in daily practice? An observation study. *Occup Med (Lond)* 2006, 56:191–198.

19. Rhebergen MDF, Lenderink AF, van Dijk FJH, Hulshof CTJ: Do Dutch workers seek and find information on occupational safety and health? *Am J Ind Med* 2012, 55:250–259.

20. Richie JLJ: Qualitative Research Practice: Guide for Social Science Students and Researchers. London: Sage Publications; 2003.

21. Tompa E, Culyer AJ, Dolinschi J: Economic Evaluation of Interventions for Occupational Health and Safety: Developing Good Practice. New York: Oxford University Press; 2008.

22. WSIB: Funding fairness: A report on Ontario's workplace and safety insurance system. http://www.wsib.on.ca/files/Content/FundingReviewFundingFairnessReport/FundingFairnessReport.pdf.

23. Ministry of Health and Long Term Care. Ministry Plans: Ontario Health Insurance Plan (OHIP). http://www.health.gov.on.ca/en/public/programs/ohip/default.aspx.

24. Glaser BG: The constant comparative method of qualitative analysis. *Soc Probl* 1965, 12:436–445.

25. Pope C, Ziebland S, Mays N: Qualitative research in health care. Analysing qualitative data. *BMJ* 2000, 320(7220):114–116.

26. Hooftman WE, Westerman MJ, van der Beek AJ, Bongers PM, van Mechelen W: What makes men and women with musculoskeletal complaints decide they are too sick to work? *Scand J Work Environ Health* 2008,34:107–112.

27. Dye JF, Schatz IM, Rosenberg RA, Coleman ST: Constant comparison method: a kaleidoscope of data. The Qualitative Report 2000, 4. http://www.nova.edu/ssss/QR/QR4-1/dye.html.

28. The Workplace Safety and Insurance Board (WSIB). http://www.wsib.on.ca/en/community/WSIB.

29. Accreditation Canada. http://www.accreditation.ca/en.

30. Brownson RC, Gurney JG, Land GH: Evidence-based decision making in public health. *J Public Health Manag Pract* 1999, 5:86–97.

31. Sackett DL, Rosenberg WCM, Gray JA, Haynes RB, Richardson WS: Evidence-based medicine: what it is and what it isn't. *BMJ* 1996, 312:71.

32. Franco G: Evidence-based decision making in occupational health. *Occup Med (Lond)* 2005, 55:1–2.

33. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB: Evidence-based Medicine: How to Practice and Teach EBM. 2nd edition. Edinburgh: Churchill Livingstone; 2000.

34. Baba VV, HakemZadeh F: Toward a theory of evidence based decision making. *Management decision* 2012, 50:832–867.

35. Walshe K, Rundall TG: Evidence-based Management: From Theory to Practice in Health Care. *Milbank Q* 2001, 79:429–457.

36. Clancy CM, Cronin K: Evidence-based decision making: global evidence. Local decisions. *Health Aff (Millwood)* 2005, 24:151–162.

37. Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF: Do health-care decision makers find economic evaluations useful? The findings of focus group research in UK health authorities. *Value Health* 2002, 5:71–78.
38. Drummond MF: Economic evaluation and the rational diffusion and use of health technology. *Health Policy* 1987, 7:309–324.
39. Ross J: The use of economic evaluation in health care: Australian decision makers' perceptions. *Health Policy* 1995, 31:103–110.

Additional file 1: Topic list of the in-depth interviews

1) How does your organization go about starting and implementing an OHS intervention? You may think about something you recently did (small versus large & mandated versus non-mandated OHS interventions).

   Prompts: *What is the decision making process? From where do the resources come?*

2) Can you describe how you evaluate OHS interventions?

   Prompts: *How well resourced are you to evaluate such initiatives? Who is responsible? What type of information helps move a plan forward? What information do people draw upon for evaluation? What kind of data and information are available to you for evaluation? What other resources such as funds and occupational health staff are available to you? How do you prioritize between alternatives? How are OHS interventions approved? Who is responsible?*
   What kinds of things do you do after program implementation to monitor, evaluate, and assess whether you reached your targets?

3) How does cost-benefit/cost-effectiveness fit into your decision making process?

   Prompts: *How does your workplace know if an OHS intervention is (financially) worthwhile? What kind of costing/evaluation does your workplace do beforehand? Do you do a business case/cost-benefit analysis? What kinds of outcomes are considered? Where do you get data for this?*

Additional file 2: Topic list of the structured telephone interviews

Sources of information used during the occupational health decision making process

Sources of information

1) Can you list the kinds of information your workplace gathers to know if a future OHS intervention will be worthwhile?
2) Do you get information on outcomes (results) for a future OHS intervention from external sources? (Always; Sometimes; Never)
    a. If so, what types of sources?

Inputs/costs and outcomes/consequences considered during the decision making process

1) I'm going to list a few inputs (costs) of economic evaluations of OHS interventions. I'd like you to tell me if you use this information in an evaluation.

	Inputs / Costs	Ranking total
1	Health and safety staff time	always, sometimes, never
2	Training the worker	always, sometimes, never
3	Planning, promotion and evaluation	always, sometimes, never
4	Equipment purchases	always, sometimes, never
5	Administration	always, sometimes, never
6	Equipment installation	always, sometimes, never
7	Ongoing equipment repair and maintenance	always, sometimes, never
8	Professional / consultant fees	always, sometimes, never
9	Ongoing supplies	always, sometimes, never

2) I'm going to list a few outcomes (consequences) of economic evaluations of OHS interventions. I'd like you to tell me if your organization considers them in an evaluation.

	Outcomes / Consequences	Ranking total
1	Number of injuries, illnesses, sickness absences	always, sometimes, never
2	Days lost due to injuries, illnesses and general sickness	always, sometimes, never
3	Quality of care and patient safety	always, sometimes, never
4	Attraction and retention	always, sometimes, never
5	Accommodating injured or ill workers	always, sometimes, never
6	Impact on productivity	always, sometimes, never
7	Worker replacement expenses	always, sometimes, never
8	Employer workers' compensation insurance premiums	always, sometimes, never
9	Employer claims management expenses	always, sometimes, never
10	Overtime payments	always, sometimes, never
11	Meaningful return to work	always, sometimes, never
12	Labour relations climate	always, sometimes, never

Occupational health decision makers' knowledge of different economic evaluation designs

I'm going to list a few economic evaluation terms you may or may not know. I just want you to tell me if you have heard of them and what the terms mean to you. Not everyone knows the meaning of the terms so just tell me what you think it is.

1) Cost-benefit analysis
   a. What does it mean to you?
   b. Does your workplace perform them? (Always; Sometimes; Never)
2) Cost-effectiveness analysis
   c. What does it mean to you?
   d. Does your workplace perform them? (Always; Sometimes; Never)
3) Cost-utility analysis
   e. What does it mean to you?
   f. Does your workplace perform them? (Always; Sometimes; Never)
4) Have you or anyone at your workplace had any training course, education, or guidance in economic evaluation for OHS interventions? (Yes/No)
   g. What type?
   h. Where did you get it?
   i. How long was it? (hours, days, weeks)

5) Is there anything in particular that you feel you want/need to learn more about to do evaluations of OHS interventions? (Yes/No)

10

General discussion

GENERAL DISCUSSION

The prevalence of modifiable health risks among the population is high, which imposes a large economic burden on society as a whole and on employers in particular. The workplace presents a useful setting to offer behavior change interventions that aim to prevent and/or reduce such risk factors. Amongst others, because a large number of people can be reached, including many who would otherwise be unlikely to engage in preventive health behaviors. Furthermore, Dutch employers themselves may financially benefit from implementing such interventions through reductions in productivity-related costs (1-5).

In practice, numerous occupational health interventions exist, of which only a limited number can be provided with the resources available (6). Therefore, high quality evidence in the form of methodologically sound economic evaluations is needed to demonstrate their value. Nonetheless, this evidence is scarce, which is partly due to the fact that only a few of the studies that consider the effectiveness of worksite health promotion programs take the extra step of considering their resource implications, and the methodological quality of those that do is generally poor. Moreover, the uptake of those that have been performed in daily practice is likely to be limited. Therefore, the aim of this thesis was to contribute to the development of a sound evidence base on the resource implications of worksite health promotion programs as well as to improve the uptake of the results of such studies in daily practice. This was done by summarizing the current literature on the cost-effectiveness and financial return of worksite physical activity and/or nutrition programs (*Chapter 2* and *3*), generating new evidence by performing economic evaluations of various newly developed worksite health promotion programs (*Chapter 4, 5, 6,* and *7*), and developing and providing recommendations for good practice when conducting and disseminating economic evaluations in occupational health (*Chapter 8* and *9*).

This general discussion is divided into five parts. First, the main findings of the systematic reviews, the applied studies, as well as a qualitative study into the information needs of occupational health decision-makers will be summarized and discussed. Second, various considerations will be discussed that warrant further exploration in relation to the methodology of economic evaluations in occupational health. Third, the present findings will be compared to the literature. Fourth, recommendations for practice and research will be presented. The discussion will end with concluding remarks.

Main findings

What is known about the cost-effectiveness and financial return of worksite physical activity and/or nutrition programs?

Chapter 2 and *chapter 3* describe two systematic reviews that summarize and critically appraise the current evidence of the cost-effectiveness and financial return of worksite physical activity and/or nutrition programs, respectively. From the review results described in *chapter 2*, no firm conclusions can be drawn about the cost-effectiveness of such interventions. This was due to the fact that the included studies used a broad range of outcome measures and analytic perspectives, which hampered pooling of their results. Also, most interventions were more costly and more effective in improving various health outcomes (e.g. body weight and cholesterol level reduction), whereas set levels as to how much decision-makers are willing to pay for these improvements are currently lacking. The review in *chapter 3* found that average financial return estimates of worksite physical activity and/or nutrition programs in terms of absenteeism benefits, medical benefits, or both, were positive in non-randomized studies, but negative in randomized controlled trials (RCTs). These results indicate that financial return estimates derived from non-randomized studies should be interpreted with great caution. Economic evaluations alongside RCTs with a low risk of bias, on the other hand, indicate that worksite physical activity and/or nutrition programs may not pay for themselves in terms of absenteeism and/or medical benefits during the first years after implementation. However, as such programs are thought to be associated with additional types of benefits (e.g. presenteeism benefits), which have not been measured in most of the studies included in the review, conclusions about their overall profitability cannot be made.

Methodological quality of economic evaluations of worksite physical activity and/or nutrition programs

In both of the aforementioned systematic reviews (C*hapter 2* and *3*), the methodological quality of the included studies was assessed using a consensus-based methodological quality checklist. Both assessments indicated that the methodological quality of the included studies was generally poor. Examples of quality criteria that were least fulfilled included those related to the description of the study population and alternatives under study, the identification, measurement, and valuation of resource use, as well as the performance of sensitivity analyses and discounting. Also, few studies reported on the uncertainty of their cost-effectiveness and/or financial return estimates. The latter is a critical oversight as failing to estimate values under uncertainty may lead to biased conclusions and could thus result in inappropriate decision-making.

Do the evaluated worksite health promotion programs provide good value?

In *chapter 4* through *chapter 7*, four economic evaluations of various newly developed worksite health promotion programs were presented. Three economic evaluations were conducted alongside RCTs (*Chapter 4, 5,* and *6*), whereas the fourth used a 2X2 factorial design (*Chapter 7*). All interventions were compared to usual practice, both their cost-effectiveness and financial return were evaluated, analyses were performed from both the societal and employer's perspective, and the follow-up duration of all studies was 12 months. The main findings of the studies were:

- <u>Vital@Work study</u>: The worksite vitality intervention for older hospital workers evaluated in *chapter 4* was neither cost-effective from the societal perspective in improving general vitality, work-related vitality, and need for recovery, nor did it result in financial savings for the employer.
- <u>The Mindful VIP study</u>: The mindfulness-based worksite intervention for knowledge workers evaluated in *chapter 5* was neither cost-effective from the societal perspective in improving work engagement and general vitality, nor from that of the employer in improving work engagement, job satisfaction, and work ability. Also, the intervention was not saving costs to the employer.

- The VIP in Construction study: The worksite physical activity and nutrition program for construction workers evaluated in *chapter 6* was not cost-effective from the employer's perspective in improving work-related vitality and job satisfaction. The intervention's cost-effectiveness in improving weight-related outcomes (societal perspective) and musculoskeletal disorders (employer's perspective) depends on the respective decision-makers' willingness-to-pay for these effects. Also, even though financial return estimates were positive, the intervention was not considered cost saving to the employer due to a high level of uncertainty.
- The Be Active & Relax VIP study: Whether the combined social and physical environmental intervention evaluated in *chapter 7* can be regarded as cost-effective in improving need for recovery from both the societal and employer's perspective depends on the respective decision-makers' willingness-to-pay for these effects. The separate interventions were not cost-effective in improving this outcome. Moreover, none of the interventions was cost-effective in improving general vitality (societal perspective) and job satisfaction (employer's perspective), nor did they result in financial savings for the employer.

Information needs of occupational health decision-makers

Chapter 9 presents the results of a qualitative study into the occupational health decision-making process and information needs of occupational health decision-makers in the Ontario healthcare sector. The study indicated that the decision-making process can be generally subdivided into three stages: 1) initiation stage, during which the need for an intervention is established; 2) pre-implementation stage, during which an intervention and its business case are developed in order to receive senior management approval; and 3) implementation and evaluation stage, during which an intervention is implemented and evaluated. Organizations were found to invest in occupational health and safety interventions for legal, financial, and moral reasons. Financial information, especially the employer's costs and benefits, was found to be particularly important at the front end of implementation decisions, and can be a key deciding factor of whether to go forward with a new intervention. Results also indicated that occupational health decisions are currently not being made in an evidence-based manner and that there is a need to advance the decision-makers' economic evaluation skill set. Further research is needed to explore whether these results are generalizable to the Dutch occupational health context. Nonetheless, it seems reasonable to assume that Dutch occupational health decision-makers are also particularly interested in the interventions' costs and benefits to the employer and that implementation decisions are not being made in an evidence-based manner as well.

Methodological considerations

Many of the methodological strengths and limitations of the applied studies have been discussed in *chapter 4* through *chapter 7*. In addition, recommendations for good practice when conducting economic evaluations in the field of occupational health research have been described in *chapter 8*. However, a selection of methodological considerations in relation to the study population, analytic perspective, study power, missing data, the identification, measurement, and valuation of resource use, time horizon, as well as the generalizability of our results warrant further exploration.

Study design

Three studies were conducted alongside an RCT (*Chapter 4, 5,* and *6*), while the Be Active & Relax VIP study used a 2X2 factorial design (*Chapter 7*). All studies used a pragmatic design, meaning that the interventions were evaluated under circumstances that resembled routine practice conditions as much as possible (7).
The pragmatic design of the studies made it possible to evaluate the interventions' (resource) implications under "real world" circumstances. This facilitates the generalizability of our results (i.e. external validity), whereas the randomization of participants improved the studies' internal validity (i.e. the ability to draw true conclusions about causes and effects) (7). The importance of randomization, on the other hand, was underscored by the review presented in *chapter 3*, in which average financial return estimates were found to differ between studies with and without randomization. Nonetheless, many economic evaluations of worksite health promotion programs are currently performed alongside non-randomized studies

(See *chapter 2* and *3*), even though these are particularly prone to selection bias. Selection bias arises when allocation methods other than randomization are used, meaning that the intervention and control group are unlikely to be comparable (8). For example, due to the lack of randomization it is unclear whether program participants were healthier and/or more motivated to change their health behavior(s) to begin with than non-participants. The possible existence of such a priori differences makes it hard to attribute study results to the intervention and to rule out the possibility that they were caused by (baseline) differences between study arms (i.e. confounding caused by selection bias) (3;8). Some people question the applicability of RCT results to daily practice, because the same design aspects that contribute to their high internal validity (e.g. well-defined inclusion and exclusion criteria) may simultaneously hamper the generalizability of their results in an extended population and/or setting (7). However, although other research designs may add to the existing knowledge on worksite health promotion programs, RCTs should be viewed as the "gold standard" for evaluating their (resource) implications untainted by bias (8).
In three studies (*Chapter 4, 5,* and *6*), randomization was performed at the participant-level, whereas group allocation was performed at the department-level in the Be Active & Relax VIP study (*Chapter 7*). The latter was done because the intervention under study operated on the group-level rather than on the individual-level as well as to avoid contamination between study groups (9). Methods for economic evaluations alongside RCTs are relatively well established (10;11), and these methods were used to evaluate the data of such studies. A fundamental issue in clustered studies, however, is that costs and consequences within a cluster may be more similar to each other than costs and consequences from a different cluster. As a consequence, methods that ignore clustering in economic evaluations generally underestimate the statistical uncertainty and are likely to have inaccurate point estimates (9;12). Based on recent research findings on the optimal strategy to account for clustering in economic evaluations (12;13) we therefore used multilevel analyses to assess the cost-effectiveness and financial return of the Be Active & Relax VIP interventions.

Study population

All interventions were aimed at primary prevention. That is, they were directed at all employees of the participating companies, who on average were generally healthy, instead of high-risk individuals (3). This approach attempts to shift the whole distribution of exposure in a particular population in a favorable direction by controlling the determinants of a disease and by lowering the mean level of risks. A drawback of primary prevention is that it offers only small benefits to individuals at the short-term, because their absolute risk for a disease is generally low (14). Consequently, it is relatively hard to motivate them to change their unhealthy behavior(s), and thus to achieve sustained health improvements (14;15). This may partially explain the lack of, or relatively small, effects of the interventions. To produce better effects, a high-risk strategy may be needed, in which prevention efforts are solely aimed at high-risk individuals (e.g. overweight and/or obese construction workers in the case of the VIP in Construction intervention (*Chapter 6)).* Such an approach likely offers a more cost-effective use of limited resources, because it is generally more efficient to concentrate limited time and money where the need, and therefore also the benefits, are likely to be greatest (14). High-risk strategies, however, do not deal with the root of a problem and it is questionable whether employers are willing to discriminate between their employees by providing different worksite health promotion programs to different groups of high-risk individuals. Therefore, a combination of various prevention strategies may ultimately be needed to achieve a meaningful degree of prevention in the workplace (16).

In three of the applied studies (*Chapter 4, 5,* and *7*), participants had relatively good baseline values of the primary outcomes, which further reduced the interventions' ability to accomplish sustained health improvements (i.e. ceiling effect) (Table 1). Selective enrolment of healthy individuals is not uncommon in health promotion programs/studies and is explained by the fact that people with healthy lifestyle behaviors are also the ones who are generally most motivated to pursue and maintain health (17;18). In the VIP in Construction study (*Chapter 5*), on the other hand, a relatively large number of obese construction workers was included (Table 1). This might have resulted from the fact that occupational physicians, who played an import role in the recruitment process, may have been particularly inclined to motivate obese construction workers for study participation. The selective inclusion of either healthy or unhealthy employees likely bears on the generalizability of our findings, and should thus be taken into account when making inferences about the interventions' resource implications in a broader working population/setting.

Table 1: Participants' baseline values of the primary outcomes in relation to their respective norm scores

Study	Primary outcome	Baseline scores participants	Norm scores
VIP in Construction Study	Body weight		
	Normal Weight (BMI ≤ 25 kg m^{-2})	30.1%	34.3%[1]
	Overweight (BMI ≥ 25 kg m^{-2} and	47.4%	48.8%
	BMI < 30 kg m^{-2})	22.4%	16.9%
	Obesity (BMI ≥ 30 kg m^{-2})		
Vital@Work Study	Work-related vitality (Range: 0-6) [Mean (SD)]	4.88 (0.85)	4.01[2]
Mindful VIP study	Work engagement (Range: 0-6) [Mean (SD)]	4.10 (0.89)	3.82[3]
Be Active & Relax VIP study	Need for recovery (Range: 0-100, with lower scores indicating a lower need for recovery) [Mean (SD)]	33.2 (29.3)	38.1[4]

Abbreviations: BMI: Body Mass Index, SD: Standard Deviation, n: number
[1]Percentage of normal weight, overweight, and obese Dutch construction workers (100)
[2] Average work-related vitality score among the general Dutch working population (101)
[3] Average work engagement score among the general Dutch working population (101)
[4] Average need for recovery score among the general Dutch working population (102)

Analytic perspective

All studies applied a so-called two-perspective approach, in which analyses were performed from both the societal and employer's perspective (19;20). In the societal perspective, all costs and consequences are taken into account irrespective of who pays or benefits, whereas only those borne by, or accruing to, employers are included when the employer's perspective is applied. The main advantage of the employer's perspective is that its results are directly interpretable for those who we are trying to aid with our economic evaluations, namely occupational health decision-makers. A disadvantage of this perspective is that it does not provide an indication of whether the "local rationality" of the company is in line with societal optimality (i.e. maximizing

the welfare of society as a whole with the resources available) (19). To deal with this issue, analyses were also performed from the societal perspective, which provides insight into the interventions' net societal effects. Even though occupational health decision-makers themselves may view societal perspective results as externalities, having them ignorant of these results may lead to non-optimal resource allocation decisions at the aggregate level (19;21). Another advantage of the societal perspective is that its disaggregate information on costs and consequences gives a good sense of their distribution across stakeholders, which could provide a starting point for bargaining between them (11). Moreover, the application of the societal perspective improves the transferability of our results to countries with different (occupational) health and welfare systems. For example, U.S. employers who typically bear most of the healthcare costs of their employees, can extract this information from the disaggregate information on costs and consequences from the societal perspective.
Even though it was not the case in the applied studies, it is important to mention that economic evaluations from the societal and employer's perspective may provide conflicting results. For example, worksite health promotion programs whose benefits fall entirely on employees in the form of improved health, but do not have a positive impact on productivity and/or occupational health costs, may be justified in social terms, but may not be in any company's financial interest to implement (22). In case of such a scenario, other stakeholders (e.g. "*the Dutch Ministry of Health, Welfare, and Sports*") may wish to consider giving incentives to companies to ensure that a socially preferred program goes ahead (11;22). If the opposite is true (i.e. a new intervention is cost-effective from the employer's perspective, but not from the societal one), it is of utmost importance that occupational health decision-makers are made aware of the fact that an intervention which benefits their goals is unattractive to other stakeholders and society as a whole in order to discourage them from implementing such an intervention (19).

Study power

All sample sizes were based on detecting relevant differences in health and/or work-related outcomes, and not to detect relevant cost differences. However, as only a small proportion of participants incur high costs and costs are naturally bound by zero, cost data have the tendency to follow a rightly skewed distribution. As a consequence, economic evaluations generally require much larger sample sizes than their corresponding effect analyses in order to achieve sufficient power to detect relevant cost differences (23;24). Thus, all of the applied studies are likely to be underpowered. This is a common problem in trial-based economic evaluations and is often due to various factors. First, many economic evaluations are "piggybacked" onto effectiveness trials, and power calculations are therefore typically performed before the economic evaluation requirements are considered (23). Second, a large number of parameters has to be specified in order to perform sample size calculations for economic endpoints, many of which are hard to forecast a priori (25). Third, and most importantly, if studies would be sufficiently powered to detect relevant cost differences, they typically become infeasible with extremely large sample sizes and very high research expenses (20;26).

If studies are likely to be underpowered, it is recommended to use estimation and/or decision uncertainty rather than hypothesis testing (11;23). Therefore, economists typically focus on estimating cost and effect differences and assessing the probability of an intervention being cost-effective (i.e. "*How confident are we that an intervention is cost-effective?*"), rather than testing a particular hypothesis (e.g. *"Are the cost-effectiveness outcomes statistically significant?"*) (26;27). In line with this recommendation, confidence intervals around cost and effect differences as well as financial return estimates were presented, and the interventions' probabilities of cost-effectiveness were explored at different ceiling ratios (i.e. the maximum amount of money decision-makers are willing to pay for an additional unit of effect) (11). Although confidence intervals around financial return estimates are relatively straightforward to interpret for researchers, many occupational health decision-makers lack the required economic and/or statistical background (See *chapter 9*). Therefore, the concept of the "probability of financial return" was introduced in *chapter 8*. This probability provides an indication of the likelihood that, given the data, a new intervention is cost saving. Occupation health decision-makers can subsequently use this information to consider whether the established probability of financial return is acceptable to them.

Missing data

All studies had some missing data, ranging from 12% to 41% on the effect measures and from 29% to 62% on the cost measures. Missing data are often inevitable in trials due to participant drop-out and/or non-response (11;28). In economic evaluations, the problem of missing data is even more pronounced, because cost data are generally the sum of numerous components and relatively short recall periods (and thus more measurement points) are needed to reliably estimate them (26;29). When data are missing, the key challenge is to maximize usage of available data while minimizing the bias introduced by the elements that are missing (28). Simply eliminating participants who have missing data (i.e. complete-case analysis) is inefficient, as it ignores available data of incomplete cases and produces a reduced-sized dataset of complete-cases, and thus a loss of power (11;28). On top of that, complete-case analyses may be biased when systematic differences exist between the missing and observed values (28;30). In all studies, multiple imputation was therefore used to fill in missing values. Multiple imputation is currently preferred over so-called naive methods (e.g. last-observation carried forward), because it accounts for the uncertainty associated with filling in the missing values (30;31).

Within a study, results derived using multiple imputation may differ from those of a complete-case analysis. To a greater or lesser extent, this was also the case in the applied studies. For example, excluding participants with incomplete data in the Vital@Work study (*Chapter 4*) resulted in positive financial return estimates, whereas the reverse was the case when multiple imputation was applied. On the basis of the aforementioned reasons, we always considered the results derived from the multiple imputed datasets to be more reliable than those of the complete-case analysis. Nonetheless, it is important to bear in mind that multiple imputation is based on the assumption that data are missing at random (MAR; missing data has a relation to observed factors and not to unobserved factors), an assumption that may not necessarily hold true but cannot be tested. Therefore, having a complete dataset is always preferred and every endeavor should be made in future studies to reduce the amount of missing data. Amongst others, this may be accomplished by minimizing the length of the questionnaires, using incentives, systematically contacting participants when their responses are missing, unclear, and/or incorrect, and using modern data collection technologies, such as online questionnaires and mobile apps, to reduce the burden of the data collection process. When doing so, it is advisable to use a so-called mixed approach, since a strategy that may limit non-response among one type of participant, may not be effective for another (32).

Identification of resource use

As has been explained earlier, relevant resource use categories for inclusion in an economic evaluation depend on its analytic perspective. Other factors that might determine their relevance are, amongst others, the country or jurisdiction in which the study is undertaken, the nature of the alternatives being compared, and the relative order of magnitude of the resource use categories (11). From the societal perspective, resource use from the healthcare, alternative care, and occupational health sector, as well as that of employees, and changes in paid productivity were included. The latter were expressed in terms of changes in lost production due to sickness absence (i.e. absenteeism) as well as reduced performance while at work (i.e. presenteeism). The inclusion of presenteeism costs in economic evaluations is a much debated topic, particularly because a sound methodological framework for their assessment is currently lacking (21;29;33). After some consideration, we decided to include this resource use category in all studies, because presenteeism seems to account for the largest component of paid productivity changes and efforts to improve health were found to have a more immediate effect on presenteeism than on absenteeism (29;33-35). Resource use of family members and changes in unpaid productivity, on the other hand, were not included, as our economic evaluation results were expected to be unaffected by them (11). When the employer's perspective was applied, analyses were restricted to resource use from the occupational health sector and changes in paid productivity.

Measurement of resource use

Resource use data can be collected through a variety of means, including the use of insurance records, company databases, questionnaires, and prospective cost diaries. Of them, more objective measurement strategies are favoured over those that rely on participant self-report, because they minimize the possible influence of recall bias (i.e. bias due to inaccurate and/or incomplete recollections of events) (37). Unfortunately, however, objective measurements were not always feasible and/or preferred in the applied studies.

Questionnaires were used in all studies to assess healthcare utilization, because collecting health insurance claim data of participants was practically infeasible and would not have provided all required information. To illustrate, Dutch employees can buy insurance packages from over 30 different insurance companies, most insurance companies offer various levels of supplementary insurance packages, and people can buy basic and supplementary insurance packages from different insurance companies (36). Even if all insurance companies would have been willing to provide data, which is highly unlikely, healthcare claim data would not have been comparable between employees, because the treatments covered (and claimed) differ between them. Furthermore, health insurance records often lack detailed resource use information and information on the healthcare services borne by employees themselves (e.g. co-payments, over-the-counter medication) are typically not included (37).

As it was not feasible to objectively measure on-the-job productivity, presenteeism data were collected using questionnaires as well. For this purpose, the *"World Health Organization – Health and Work Performance Questionnaire"* (WHP-HPQ) was used, which has shown good concordance with archival performance data (38;39). It should be noted, however, that numerous instruments exist for assessing presenteeism and that their estimates may vary widely. This suggests a lack of comparability among instruments, but it is still unclear which instrument provides the best estimates (29). We opted for the WHO-HPQ, because it is the most frequently used instrument in economic evaluations of similar interventions, and thus increases the comparability of our results (See *chapter 3*).

Questionnaires were also used for assessing sickness absence in the Vital@Work study (*Chapter 4*), whereas sickness absence data were extracted from company records in all other studies (*Chapter 5, 6,* and *7*). Research indicates that absenteeism estimates may differ extensively between both methods, and that the accuracy of self-reported sickness absence estimates strongly decreases with an increasing recall period (40;41). Given the available evidence on the optimal recall period for absenteeism, Zang et al. (2011) recommended the application of a 3-month recall period in order to balance loss in precision and the increase in research costs and participant burden (29). As this recall period was used in the Vital@Work study as well, we do not expect that its results are severely distorted by recall bias. Nonetheless, as most employers systematically track employee sickness absence and sickness absence data are relatively easy to collect when conducting studies at the workplace, future economic evaluations of worksite health promotion programs are recommended to use company records whenever possible.

As indicated above, questionnaires may be prone to recall bias. However, as it seems highly unlikely that the extent of impairment in recall systematically differed between study groups, we do not expect that our reliance on them severely biased our results (42). When having to rely on participant self-report, the possible influence of recall bias may be reduced by reducing a questionnaire's recall period (e.g. 3 months for absenteeism and healthcare utilization data (29;43) and 2 weeks for presenteeism data (29)) or by using a more accurate data collection method, such as a prospective cost diary. Provided that participants truly complete such diaries in a prospective way, they are thought to result in a minimum recall error and therefore in a better and more complete reporting of resource use (37).

Valuation of resource use

One of the most important challenges when valuing resource use is the identification of the "best" price weight for translating units of resource use into monetary values. Such price weights should be based on the true opportunity cost of a good or service (i.e. the amount of money that is not available for its best alternative use), and should be reflective of the analytic perspective (11;22). Our ideas about the "best" price weights, as well as the most appropriate methods for valuing resource use, have slightly evolved over the course of this thesis and will be discussed below.

In the Vital@Work study (*Chapter 4*), intervention costs were estimated using a so-called bottom-up micro-costing approach for both the societal and employer's perspective. This means that we estimated the cost of the Vital@Work intervention by collecting detailed data regarding the resources consumed as well as their unit prices (11). In doing so, we aimed to best reflect the true cost of the intervention, meaning that profit margins and transfer payments were excluded as much as possible. In the subsequent studies, however, we deviated from this approach in that bottom-up micro-costing was solely used for the societal perspective, whereas market prices were used when the employer's perspective was applied. This was done because we are of the opinion that market prices better reflect an intervention's true value at the company level (i.e. the amount of money that is not available to the company for its best alternative use).

Healthcare utilization, which was only included when the societal perspective was applied, was valued using standard price weights whenever possible. Such standard price weights are preferred over market prices, because market prices are an inaccurate reflection of its societal opportunity cost if a perfect market does not exist for a healthcare service. For example, if a healthcare provider has a local monopoly, its charges are often an overestimation of their true (societal) value because monopolists have the power to set their own price (11;36). Healthcare provider fees may not be an accurate reflection of the time and relative skill level that is needed for different procedures. Moreover, drug prices are often set in negotiations between the government and pharmaceutical companies, where the pharmaceutical company's commitment to research and the provision of employment might be taken into account, as well as the costs of discovery, production, and distribution of the drug in question (11).

In all studies, sports costs were based on the participants' self-reported expenses on sports membership fees and sports equipment. We considered this gross-costing approach to be appropriate, because the impact of changes in sports costs on the resulting cost-effectiveness and/or financial return estimates was expected to be low (11).

Occupational health costs were only considered in the VIP in Construction (*Chapter 6*) and Mindful VIP study (*Chapter 5*). In the VIP in Construction study, they solely included employer-provided gym membership subsidies, and were valued using data derived from financial department staff. In the Mindful VIP study, on the other hand, occupation health costs consisted of a broad range of occupational health services and in-company health promotion activities of the participating companies. In line with our methods for estimating intervention costs, micro-costed price weights were used for the societal perspective, whereas marked prices were used when the employer's perspective was applied.

In three studies, changes in productivity were valued using gross salaries of participants (*Chapter 4, 5,* and *7*). In the VIP in Construction study (*Chapter 6*), on the other hand, we had to use the average salary of construction workers, because the participating construction company did not provide permission to collect participant salary data. Even though the use of age- and gender-specific price weights may have improved the generalizability of our results (44), we decided to rely on participant salary data instead in order to account for the fact that the magnitude of production losses is likely to be greater among employees with higher incomes.
Another important issue when valuing changes in productivity is the method used for estimating absenteeism costs. In the first two economic evaluations (*Chapter 4* and *5*), the “Friction Cost Approach” (FCA) was used for both the societal and employer’s perspective. The FCA is recommended by the “*Dutch Manual of Costing*” and assumes that production losses are confined to the time-span companies need to replace a sick worker by a formerly unemployed person to restore the company’s initial production level (i.e. friction period, which is estimated to be 23 weeks in the Netherlands) (21;44;45). In the subsequent studies (*Chapter 6* and *7*), we deviated from this approach in that the FCA was only used for the societal perspective, whereas the “Human Capital Cost approach” (HCA) was used when the employer’s perspective was applied. This was done because Dutch employers are obliged to pay at least 70% of the salary of sick employees for a period of two years, and most of them top up the wage payments from 70% to 100% during the first year of sickness absence (46). Thus, although the initial production level of a Dutch company may be restored after the friction period, employers still bear the additional cost of having to pay the salary of the sick worker.
It should be noted that it is unclear how accurate our productivity-related cost estimates are. First, we may have underestimated the actual productivity-related costs, because the applied methods do not account for the negative effect of absenteeism and presenteeism on co-workers in team-dependent production. The productive output of a full team may namely be jeopardized by one member’s reduced labour input, and this may be especially relevant when substitutes are less productive and/or unavailable (i.e. “The multiplier effect”). Until now, some attempts have been made in the U.S. to construct “job-dependent multipliers” that account

for the (average) effect on co-worker absenteeism and presenteeism in specific job types (21;47). However, future studies are needed to establish the validity of these multipliers and to investigate their transferability across countries and/or jurisdictions (21). Conversely, we may have overestimated the actual productivity-related costs, because productivity losses may partly be compensated during normal working hours (41;48). For example, work that is normally performed by the sick employee in question may be completed by colleagues or made up by the sick employee itself after return to work (20). Currently, it is unknown what the best method is for correcting for such possible compensations of productivity losses. As such, compensation adjustments are typically uncommon in economic evaluations. Even though we may have already included some form of correction for such compensations by factoring in the 0.8 elasticity factor when using the FCA, whether this elasticity factor indeed represents compensation during normal working hours is currently unknown. Therefore, as various studies indicate that over half of the lost work is compensated during normal working hours, further research in this area is warranted (41;48;49).

Time horizon

All studies applied a follow-up of one year. As many of the (health) benefits of preventive interventions, such as ours, are thought to occur in the future, this follow-up is probably insufficient to capture all costs and consequences flowing from the interventions under study (11). Decision analytic modeling may be used to bridge the gap between what has been observed in the applied studies and what would be expected to happen over a longer time horizon (11). The validity of such modeling studies, however, strongly relies on the quality of the information used for constructing the model (11;22). Amongst others, there is a risk of overstating the benefits, especially if there is the possibility of decreased intervention effectiveness over time (22). Evidence indicates that the latter is often the case in health promotion studies (50), and this phenomenon was also observed in some of the applied studies. Therefore, when trying to extrapolate the present findings, various scenarios for the sustainability of the effects should be used. One should bear in mind, however, that it is highly unlikely that the longer-term cost-effectiveness and/or financial return of the evaluated interventions would be much more favorable than those observed in the applied studies, because most of them did not result in statistically significant (health) improvements at one-year follow-up. Furthermore, it is questionable whether employers would wish to implement interventions that only generate financial savings after an extensive number of years. Employees typically switch employers a couple of times during their working life, and many of the benefits are therefore likely to accrue to future employers and/or the public (i.e. "The free rider problem") (51;52).

Generalizability of results

Some factors influencing the generalizability of our findings have been mentioned earlier, including the pragmatic design of the applied studies as well as the selective enrollment of healthy and unhealthy individuals. Furthermore, most studies were performed within a single company and the worksite health promotion programs themselves were specifically tailored to the needs of stakeholders involved. As a consequence, it is unknown to what extent the results may be generalized to other companies, work settings or the general working population. Nonetheless, we at least assume that they are generalizable to other companies with similar employee populations, with similar health issues. Also, the companies' participation in the current health promotion trials may be reflective of their degree of problem recognition, and thus their current workplace culture, available policies for improving employee health, the health status, sickness absence, and work performance of their employees, as well as their motivation to improve the current situation. As such, the participating companies may represent an optimal setting and any of our effect, cost-effectiveness, and/or financial return estimates could thus be an overestimation (20). The generalizability of our findings to other countries may be limited by differences in (occupational) healthcare and social security systems (53). In the Netherlands, for example, most healthcare costs are borne by health insurance companies and the government, whereas in countries with employer-provided health insurance (e.g. the United States (U.S.)) they typically accrue to the employer. As such, our employer's perspective findings are mainly of interest to countries with similar healthcare systems. Another factor that should be noted is that healthcare expenditure levels may differ extensively between countries. For example, per capita spending on healthcare in the U.S. is double that of most European countries, leaving more room for improvements in healthcare costs (54). The generalizability of our productivity-related cost estimates, on the other hand, may be hampered by the fact that income rates, friction periods, and sickness absence behaviors may differ between countries as well. For example, it is reasonable to expect that Dutch employees are more inclined to report sick than, for example, U.S. employees, because Dutch employees generally get paid during sickness absence, while many U.S. employees are not (22;46). Other factors that may contribute to different resource use patterns include differences in the organization of (occupational) healthcare as well as the incidence of the health risk factors in question (20;55).

The easiest way to transfer economic evaluation results from one country to another would be to recalculate the monetary value of resource use for the target country where the results are to be applied and then recalculate the cost-effectiveness and/or financial return estimates of interest (53). This approach, however, is probably too simple, as additional adjustments are likely to be needed to account for differences in healthcare utilization patterns and sickness absence behaviors (56). Therefore, Manca and Willan (2006) proposed an algorithm based on the availability of data for choosing the appropriate analytic strategy for adapting economic evaluation results from one country to another. If the country of interest has participated in a multinational clinical trial in which data on costs and consequences have been collected, the preferred strategy would be to analyze the individual patient data of the trial. Such studies, however, have neither been performed in the present thesis nor in the field of worksite health promotion research. Therefore, a modeling approach would be required according to Manca and Willan (2006), in which as much cost and consequence data as possible are used from the jurisdiction of interest (56;57).

Comparison with the literature

During the last 15 years, Pelletier published a series of reviews of the clinical and cost-effectiveness of comprehensive worksite health promotion and disease management programs (58-64). The most recent review in this series reported fairly consistent positive effects on employee health and costs, as well as improvements in the number and quality of studies (59). However, effect sizes were generally small and only seven of the 27 most recently performed studies used an RCT design. Based on these results, he concluded that there was "*guarded cautious optimism*" about the clinical and/or cost-effectiveness of comprehensive health promotion and disease management programs, a conclusion that is not necessarily supported by the findings of the review presented in *chapter 2*.

Up until now, various reviews have been conducted on the financial return of worksite health promotion programs in general. For example, a 1999 review of early worksite health promotion studies, mostly conducted in the 1980s and early 1990s, found financial return estimates in terms of healthcare benefits, productivity-related benefits, or both, to range from $1.4 to $3.1 per Dollar invested in the program (65). In 2001, Aldana performed a comprehensive review of the financial return of worksite health promotion programs. Seven of the included studies reported both costs and healthcare benefits, with an average financial return of $3.5 per Dollar spent. Only three of the included studies reported financial return estimates in terms of absenteeism benefits, which ranged from $2.5 to $10.1 (66). In a more recent review, Baicker et al. (2010) found that healthcare and absenteeism costs fell by $3.3 and $2.7 per Dollar invested in the program, respectively (67). Moreover, based on a review of 62 worksite health promotion studies conducted during the last three decades, Chapman (2012) reported that participants to worksite health promotion programs had 25.1% lower absenteeism costs and 24.5% lower healthcare costs than non-participants. Twenty-five of the included studies reported financial return estimates in terms of various types of benefits, with an average of $5.6 per Dollar spent (68). All of these reviews included a broad range of worksite health promotion programs (e.g. smoking cessation, stress reduction, physical activity, and/or nutrition programs) and most of them were evaluated using non-randomized studies, of which many even lacked a comparison group. Moreover, even though all review authors reported that the quality of the included studies was less than optimal, none of them assessed their methodological quality using a consensus-based checklist, nor did they explore the possible difference in results between non-randomized studies and RCTs. The latter, however, was explored in a recent review of U.S. worksite health promotion studies published after 2000 (69). The authors found that only one of the seven studies showing cost savings utilized an RCT. In line with the review presented in *chapter 3*, they concluded that strong evidence of cost savings is currently lacking due to the general use of weak evaluation designs, and thus the possible distortion of results by selection bias.

Implications for practice

In order to prevent spending already scarce resources on ineffective and/or inefficient strategies, worksite health promotion program implementation and continuation decisions should be made in an evidence-based manner. That is, methodologically sound scientific evidence on their (financial) implications should be consulted before program implementation and sound ex-post program evaluations ought to be performed to inform continuation decisions.

Even though multiple reviews showed favorable, albeit small, effects of worksite health promotion programs on various health-related outcomes (70-75), the present thesis indicated that (strong) evidence of their cost-effectiveness and financial return is currently lacking. The latter is in contrast to the findings of most of the aforementioned reviews (65-68), which generally concluded that wider adoption of worksite health promotion programs could prove beneficial for company budgets. These reviews, however, mainly included non-randomized studies with a high risk of bias, while the review presented in *chapter 3* found financial return estimates to systematically differ between studies with and without randomization (i.e. positive in non-randomized studies and negative in RCTs). Moreover, none of the interventions evaluated in the present thesis were found to generate cost savings to the employer. Therefore, widespread implementation of existing worksite health promotion programs in an effort to generate cost savings is not recommended. It should be noted, however, that some of the evaluated interventions (*Chapter 6* and *7*) may be considered cost-effective if decision-makers are willing to invest a certain amount of money to improve employee health (52). Whether the latter is the case, however, is currently unknown.

Our recommendation is in contrast to the current widespread advertisement and implementation of worksite health promotion programs. Many program vendors advertise them by implying that they are an evidence-based strategy for reducing healthcare and/or productivity-related costs. Advertisement statements such as *"With careful planning, efficient and effective wellness programs offered to employees and their families can shrink both the waistline and the bottom line"* are common (76), but not supported by methodologically sound evidence (i.e. high quality (cluster-)RCT-based economic evaluations). Nonetheless, a recent industry survey indicated that about 50 percent of Dutch employers invest in preventive strategies at the workplace, including stress management and lifestyle interventions (77). Although some of these employers may implement such programs purely to improve employee health, controlling costs seems to be their most important motivation (See also *Chapter 9*) (52;78). Moreover, more than half (52%) of the U.S. employers that offered worksite health promotion programs in 2012 were found to believe that they were effective in reducing the company's health care costs (79). As such, the present findings indicate that an innovative and dynamic industry appears to have outpaced the underlying evidence (69;80).

Next to the fact that worksite health promotion programs are generally thought to result in financial savings, they are also expected to result in various intangible corporate benefits that cannot be considered in a return on investment analysis. Examples of such intangible corporate benefits are improved job satisfaction, employee morale, and in-role performance (i.e. behavior required by formal job descriptions) (81;82). Moreover, worksite health promotion programs are hypothesized to strengthen a company's ability to attract new talent in a competitive market place, because healthy lifestyle benefits may entice younger employees. Among existing employees, on the other hand, worksite health promotion programs are thought to improve overall perceptions of the company, engender a greater sense of commitment and trust, and thus improve employee retention (83). However, the hypothesis regarding the positive effect of worksite health promotion programs on job satisfaction is not supported by the findings presented in *chapter 5* through *chapter 7*, and strong evidence of their favorable impact on other types of intangible corporate benefits is currently lacking.

Implications for research

Future research efforts in the field of worksite health promotion should be directed towards two important gaps in knowledge. First, the relatively small effects and lack of evidence of cost savings associated with existing worksite health promotion programs, does not negate the value of improving employee health. Therefore, more research is needed to explore what attributes of worksite health promotion programs are most important and how such interventions should be optimally designed (67). Second, researchers should help ensure that worksite health promotion program implementation and/or continuation decisions are made in an evidence-based manner, because a lack thereof may result in inappropriate decision-making and thus a waste of scarce resources.

Future directions of worksite health promotion programs

The absence of, or relatively small, effects of the evaluated interventions as well as their lack of cost savings is in line with the findings of other high-quality studies on primary prevention strategies in the workplace (84-88). This raises the question of whether primary prevention programs are indeed the "optimal" strategy for improving employee health and costs. The adoption of a high-risk approach may be more likely to be cost-effective and/or cost saving, as it is generally more efficient to concentrate limited resources where the need, and therefore also the benefits, are likely to be greatest (14). As such, future worksite health promotion programs are recommended to shift their focus from primary prevention for all employees towards prevention programs that are aimed at high-risk individuals (89). A possible way to do this is by offering more comprehensive worksite health promotion programs, in which all employees are screened for various health risks, after which only those with high-risks are referred to the necessary prevention and/or treatment programs. Amongst others, such comprehensive worksite health promotion programs may be aimed at tobacco cessation, physical activity promotion, stress management, weight management, and nutritional guidance (83). In addition, most of the evaluated interventions were mainly targeted at individual determinants of behavior (e.g. through health education and communication) (see *chapter 4, 5*, and *6*), whereas interventions targeted at both individual and environmental determinants are expected to be more effective in achieving (health) behavior change (90;91). Therefore, future worksite health promotion programs are recommended to include both individual and environmental modifications. Examples of environmental modifications are healthy canteen food and physical activity promoting adaptations to the workplace, such as standing conference tables and the introduction of exercise balls. Moreover, a necessary prerequisite for any successful worksite health promotion program is a high level of participation, because "*nothing happens until [people] participate*" (3). Research indicates that participation levels are often far from optimal in worksite health promotion programs (92), and this was also the case in the applied studies. Possible means to improve program participation include the use of incentives, the provision of a variety of program modalities (e.g. coaching, health information), the use of multi-component programs, as well the integration of health promotion into the company's culture (3;70;92). Furthermore, as many worksite health promotion programs are associated with decreased effectiveness over time, future interventions are recommended to include follow-up contacts and/ or booster sessions after their completion in order to better maintain their initial results. The cost-effectiveness and/or financial return of such "optimally" designed interventions should subsequently be established by performing (cluster-)RCT-based economic evaluations.

Improving evidence-based practice in the worksite health promotion field

Two important factors currently hinder worksite health promotion program implementation and/or continuation decisions from being made in an evidence-based manner, namely the poor methodological quality of most economic evaluations of worksite health promotion programs (*Chapter 2* and *3*) and the lack of uptake of their results (*Chapter 9*). To prevent inappropriate decision-making, researchers should ensure that both issues are addressed.

Improving the methodological quality of economic evaluations of worksite health promotion programs

Recommendations for improving the quality of economic evaluations of worksite health promotion programs have been extensively provided and discussed in *chapter*

8 as well as in the methodological considerations section of this chapter. Our most important recommendations include:

- Future economic evaluations should be conducted alongside (cluster-)RCTs to minimize the possible influence of selection bias.
- Future economic evaluations should be performed from both the employer's and societal perspective. This approach ensures that the results are directly interpretable for occupational health decision-makers and provides an indication of whether the "local rationality" of the company is in line with societal optimality.
- Future economic evaluations should assess the uncertainty surrounding their cost-consequence estimates, as failing to evaluate values under uncertainty may lead to biased conclusions and may thus result in inappropriate decision-making.
- Ideally, future economic evaluations base their sample sizes on economic endpoints. If this is not possible, researchers should use estimation and/or decision uncertainty rather than hypothesis testing (i.e. providing confidence intervals and assessing the probability of cost-effectiveness and/or financial return).
- Future economic evaluations should use multiple imputation for handling missing data, as study results may be biased when systematic differences exist between missing and observed values.
- Future economic evaluations should use price weights for valuing resource use that represent their true opportunity cost to the decision-maker at hand.

Moreover, methodological issues that warrant further inquiry include the methods for economic evaluations of clustered data, the measurement and valuation of changes in on-the-job productivity, the conceptualization of multipliers and compensation mechanisms in the valuation of changes in paid productivity, as well as the transferability of economic evaluation results across countries and jurisdictions.

Improving the uptake of economic evaluation results

In order to improve the uptake of economic evaluation results, researchers should ensure that their products are in line with the information needs of occupational health decision-makers (93;94). The qualitative study included in *chapter 9* provided some initial clues as to what these information needs are at the company level. Namely, return-on-investment analyses performed from the employer's perspective were found to form the basis of business cases for worksite health promotion programs. Within these analyses, hard cost items (e.g. equipment costs, employee training costs) were of particular importance and reduced sickness absence-related costs were viewed as one of the most important benefits. Furthermore, decisions typically have to be made within a limited time frame and many decision-makers lack the skill set required to determine what economic evaluation results are most reliable, and what information should be considered, under which circumstances (95). Therefore, it is advisable to provide them with easy-to-use critical summaries of published studies (96). In the Netherlands, such critical summaries may be distributed through (applied) research institutes and/or employers' associations, or published in easily accessible journals, newsletters, or websites. Improving the economic evaluation skill set of occupational health decision-makers may be accomplished by educating them through a variety of means, including the development of handbooks and workshops on economic evaluation methods, integrating these topics into management, occupational health, and/or worksite health promotion training programs, and involving occupational health decision-makers in the process of commissioning studies (95;97;98). Participation in scientific studies is namely closely linked to the uptake of their results and may simultaneously lead to an improved economic evaluation skill set (96).

To further advance the development of a solid evidence base on the resource implications of worksite health promotion programs and to facilitate the uptake of their results, it is recommendable to develop a set of consensus-based guidelines for good practice when conducting and reporting economic evaluations of interventions in the workplace. In order to be successful, such guidelines must be based on sound economic principles and meet the needs of all stakeholders (99). As such, they are ideally developed through a close cooperation between economists, occupational health researchers, workplace parties, policy-makers, and all other possibly relevant stakeholders.

Concluding remarks

The present thesis indicated that (strong) evidence of the cost-effectiveness and/or financial return of worksite health promotion programs is currently lacking. Therefore, widespread implementation of such interventions in an effort to generate cost savings is not recommended, while some of them may be considered cost-effective if decision-makers are willing to invest a certain amount of money to improve employee health. Whether the latter is the case, however, is currently unknown.

The lack of evidence of cost savings associated with existing worksite health promotion programs, does not negate the value of improving employee health. Therefore, more research is needed to explore what attributes of worksite health promotion programs are most important and how such interventions should be optimally designed. Amongst others, existing worksite health promotion programs may be improved by using a so-called high-risk approach, including environmental modifications, incorporating strategies to improve program participation, and including follow-up contacts and booster sessions after their completion in order to better maintain their initial effects. The cost-effectiveness and/or financial return of such “optimally” designed interventions should subsequently be established by performing (cluster-) RCT-based economic evaluations. Furthermore, the methodological quality of economic evaluations of worksite health promotion programs is generally poor, as is the uptake of their results in daily practice. To prevent inappropriate decision-making, researchers should ensure that both issues are addressed and recommendations have been provided in this thesis as to how this may be established.

REFERENCES

1. Glasgow RE, McCaul KD, Fisher KJ. Participation in worksite health promotion: A critique of the literature and recommendations for future practice. *Health Educ Q* 1993;20(3):391-408.
2. McDaid D. The economics of mental health in the workplace: what dowe know and where do we go? *Epidemiol Psichiatr Soc* 2007;16:294-298.
3. Goetzel RZ, Ozminkowski RJ. The health and cost benefits of work site health-promotion programs. *Annu Rev Public Health* 2008;29(1):303-323.
4. Goetzel RZ, Juday TR. What's the ROI? A systematic review on return of investment (ROI) studies of corporate health and productivity management initiatives. *AWPH's Worksite Health* 1999;(6):12-21.
5. Bull SS, Gillette C, Glasgow RE, Estabrooks P. Work site health promotion research: to what extent can we generalize the results and what is needed to translate research to practice? *Health Educ Behav* 2003;30(5):537-549.
6. Burdorf A. Economic evaluation in occupational health-its goals, challenges, and opportunities. *Scan J Work Environ Health* 2007;33(3):161-164.
7. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues Clin Neurosci* 2011;13(2):217-224.
8. Higgins JPT, Altman DG. Chapter 8: Assessing risk of bias in included studies. In: Higgins J.P.T., Green S., editors. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. (updated March 2011) ed. The Cochrane Collaboration; 2011.
9. Gomes M, Grieve R, Nixon R, Edmunds WJ. Statistical methods for cost-effectiveness analyses that use data from cluster randomized trials: A systematic review and checklist for critical appraisal. *Med Decis Making* 2012;32(1):209-220.
10. Willan AR, Briggs AH, Hoch JS. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health Econ* 2004;13(5):461-475.
11. Drummond MF, Sculpher M.J., Torrance G.W., O'Brien B.J., Stoddart G.L. *Methods for the Economic Evaluation of Health Care Programmes. 3rd ed*. Oxford University Press: New York, 2005.
12. Grieve R, Nixon R, Thompson SG. Bayesian hierachical models for cost-effectiveness analyses that use data from cluster randomized trials. *Med Decis Making* 2010;30(2):163-175.
13. Gomes M, Grieve R, Nixon R, Ng ES, Carpenter J, Thompson SG. Methods for covariate adjustments in cost-effectiveness analysis that use cluster randomised trials. *Health Econ* 2012;21(9):1101-1118.
14. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 2001;30(3):427-432.
15. Gatchel R, Kishino N. *Conceptual Approaches to Occupational Health and Wellness*: An Overview. In: Gatchel RJ, Schultz IZ, editors. Handbook of Occupational Health and Wellness. Springer US; 2012. p. 3-21.
16. Verweij LM. Occupational Health Guideline for Preventing Weight gain among Employees. Thesis (PhD), VU University Amsterdam, 2012.
17. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. *J Med Internet Res* 2010;12:e69.

18. van Dongen J, van Poppel M, Milder I, van Oers H, Brug J. Exploring the reach and program use of hello world, an email-based health promotion program for pregnant women in the Netherlands. *BMC Research Notes* 2012;5(1):514.
19. Brouwer WB, van Exel NJ, Baltussen RM, Rutten FF. A dollar is a dollar is a dollar--or is it? *Value Health* 2006;9(5):341-347.
20. Uegaki K. Economic evaluation of interventions for occupational health. Thesis (PhD), Vrije Universiteit Amsterdam, 2010.
21. Krol M, Brouwer W, Rutten F. Productivity costs in economic evaluations: past, present, future. *Pharmacoeconomics* 2013;31(7):537-549.
22. Tompa E, Culyer AJ, Dolinschi R. *Economic evaluations of interventions for occupational health and safety. Developing good practice.* Oxford University Press: New York, 2008.
23. Briggs A. Economic evaluation and clinical trials: size matters. The need for greater power in cost analyses poses an ethical dilemma. *BMJ* 2000;321(7273):1362-1363.
24. Bosmans JE. Cost-effectiveness of treatment of depression in primary care. Thesis (PhD). Vrije Universiteit Amsterdam, 2007.
25. Briggs AH, Gray AM. Power and Sample Size Calculations for Stochastic Cost-Effectiveness Analysis. *Med Decis Making* 1998;18(2):S81-S92.
26. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. BMJ 2011;342:d1548.
27. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic evaluations in clinical trials. Oxford University Press: New York, United States, 2007.
28. Annemans L, Ollendorf LD. The Missing Link: Managing Missing Data in Economic Evaluations*. The Official News & Technical Journal Of The International Society For Pharmacoeconomics And Outcomes Research* 2007.
29. Zhang W, Bansback N, Anis AH. Measuring and valuing productivity loss due to poor health: A critical review. *Soc Sci Med* 2011;72(2):185-192.
30. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice*. Statist Med* 2011;30(4):377-399.
31. Sterne JAC., White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009;338:b2393.
32. van de Kerckhove W, Krenze T, Mohadjer L. Approaches to a Nonresponse Bias Analysis in an Adult Literacy Survey. *Proceedings of the Survey Research Methods Section, American Statistical Association* 2006;3790-3795.
33. Søgaard R, Sørensen J, Linde L, Hetland ML. The significance of presenteeism for the value of lost production: the case of rheumatoid arthritis. *Clinicoecon Outcomes Res* 2010;2:105-112.
34. Caverley N, Cunningham JB, MacGregor JN. Sickness Presenteeism, Sickness Absenteeism, and Health Following Restructuring in a Public Service Organization. *Journal of Management Studies* 2007;44(2):304-319.
35. Burton WN, Conti DJ, Chen CY, Schultz AB, Edington DW. The Economic Burden of Lost Productivity Due to Migraine Headache: A Specific Worksite Analysis. J Occup Environ Med 2002;44(6):523-529.
36. *Economie van de Gezondheidszorg*. Reed Business education: Amsterdam, 2012.

37. Goossens M, Rutten-van Mölken M, Vlaeyen J, van der Linden S. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol* 2000;53(7):688-695.

38. Kessler R, Barber C, Beck A, Berglund P, Cleary P, McKenas D. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med* 2003;45:156-174.

39. Kessler RC, Ames M, Hymel PA, Loeppke R, McKenas DK, Richling DE. Using the world health organization health and work performance questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med* 2004;46:S23-S37.

40. van Poppel MNM, de Vet HCW, Koes BW, Smid T, Bouter LM. Measuring sick leave: a comparison of self-reported data on sick leave and data from company records. *Occupl Med (Lond)* 2002;52(8):485-490.

41. Severens JL, Mulder J, Laheij RJ, Verbeek AL. Precision and accuracy in measuring absence from work as a basis for calculating productivity costs in The Netherlands. *Soc Sci Med* 2000;51(2):243-9.

42. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol* 1990;43(1):87-91.

43. Bhandari A, Wagner T. Self-reported utilization of health care services: improving measurement and accuracy. *Med Care Res Rev* 2006;63(2):217-235.

44. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. *Handleiding Voor Kostenonderzoek*. Methoden en Standaardkostprijzen Voor Economische Evaluaties in de Gezondheidszorg. Diemen, the Netherlands: College Voor Zorgverzekeringen; 2010.

45. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. *J Health Econ.* 1995;14:171–189.

46. OECD. Sickness and Disability Schemes in the Netherlands: Country memo as a background paper for the OECD Disability Review. http://www.oecd.org/social/soc/41429917.pdf

47. Nicholson S, Pauly MV, Polsky D, Sharda C, Szrek H, Berger ML. Measuring the effects of work loss on productivity with team production. *Health Econ* 2006;15(2):111-123.

48. Krol M, Brouwer WBF, Severens JL, Kaper J, Evers SMAA. Productivity cost calculations in health economic evaluations: Correcting for compensation mechanisms and multiplier effects. *Soc Sci Med* 2012;75(11):1981-1988.

49. Jacob-Tacken KHM, Koopmanschap MA, Meerding WJ, Severens JL. Correcting for compensating mechanisms related to productivity costs in economic evaluations of health care programmes. *Health Econ* 2005;14(5):435-443.

50. Barte JCM, Ter Bogt NCW, Bogers RP, Teixeira PJ, Blissmer B, Mori TA, et al. Maintenance of weight loss after lifestyle interventions for overweight and obesity, a systematic review. *Obes Rev* 2010;11(12):899-906.

51. Herring B. Suboptimal provision of preventive healthcare due to expected enrollee turnover among private insurers. *Health Econ* 2010;19(4):438-448.

52. Horwitz JR, Kelly BD, DiNardo JE. Wellness Incentives In The Workplace: Cost Savings Through Cost Shifting To Unhealthy Workers. *Health Aff (Millwood)* 2013;32(3):468-476.

53. Verbeek J, Pulliainen M, Kankaanpaa E, Taimela S. Transferring results of occupational safety and health cost-effectiveness studies from one country to another - a case study. *Scand J Work Environ Health* 2010;36(4):305-312.

54. Reinhardt UE, Hussey PS, Anderson GF. U.S. Health Care Spending In An International Context. *Health Aff (Millwood)* 2004;23(3):10-25.
55. Sculpher M, Pang F, Manca A, Drummond M, Golder S, Urdahl H. Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technol Assess* 2004;8(49):206.
56. Drummond M, Barbieri M, Cook J, Glick HA, Lis J, Malik F, et al. Transferability of economic evaluations across jurisdictions: ISPOR good research practices task force report. *Value Health* 2009;12(4):409-418.
57. Manca A, Willan AR. Lost in translation: accounting for between-country differences in the analysis of multinational cost-effectiveness data. *Pharmacoeconomics* 2006;24:1101-1119.
58. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VI 2000-2004. *J Occup Environ Med* 2005;47(10):1051-1058.
59. Pelletier KR. A Review and Analysis of the Clinical and Cost-effectiveness Studies of Comprehensive Health Promotion and Disease Management Programs at the Worksite: Update VIII 2008 to 2010. *J Occup Environ Med* 2011;53(11):1310-1331.
60. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VII 2004-2008. *J Occup Environ Med* 2009;51(7):822-837.
61. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1991-1993 update. *Am J Health Promot* 1993;8(1):50-62.
62. Pelletier KR. A review and analysis of the health and cost-effective outcome studies of comprehensive health promotion and disease prevention programs at the worksite: 1993-1995 update. *Am J Health Promot* 1996;10(5):380-388.
63. Pelletier KR. A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1995-1998 update (IV). *Am J Health Promot* 1999;13(6):333-345.
64. Pelletier KR. A review and analysis of the clinical- and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: 1998-2000 update. *Am J Health Promot* 2001;16(2):107-116.
65. Goetzel RZ, Juday TR, Ozminkowski RJ. What's the ROI? A systematic review on return of investment (ROI) studies of corporate health and productivity management initiatives. *AWPH's Worksite Health* 1999;6:12-21.
66. Aldana SG. Financial impact of health promotion programs: a comprehensive review of the literature. *Am J Health Promot* 2001;15(5):296-320.
67. Baicker K, Cutler D, Song Z. Workplace wellness programs can generate savings. *Health Aff (Millwood)* 2010;29(2):304-11.
68. Chapman LS. Meta-Evaluation of Worksite Health Promotion Economic Return Studies: 2012 Update. *Am J Health Promot* 2012;26(4):TAHP-1-TAHP-12.
69. Osilla KC, van Busum K, Schnyer C, Larkin JW, Eibner C, Mattke C. Systematic review of the impact of worksite wellness programs. *Am J Manag Care* 2012;18(2):e68-e81.
70. Soler RE, Leeks KD, Razi S, Hopkins DP, Griffith M, Aten A, et al. A systematic review of selected interventions for worksite health promotion: The assessment of health risks with feedback. *Am J Prev Med* 2010;38(2, Supplement 1):S237-S262.

71. Anderson LM, Quinn TA, Glanz K, Ramirez G, Kahwati LC, Johnson DB, et al. The effectiveness of worksite nutrition and physical activity interventions for controlling employee overweight and obesity: a systematic review. *Am J Prev Med* 2009;37(4):340-357.

72. 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. *Clin J Sport Med* 2003;13(2):106-117.

73. Engbers LH, van Poppel MNM, Chin AP, van Mechelen W. Worksite Health Promotion Programs with Environmental Changes: A Systematic Review. *Am J Prev Med* 2005 Jul;29(1):61-70.

74. Groeneveld IF, Proper KI, van der Beek AJ, Hildebrandt VH, van Mechelen W. Lifestyle-focused interventions at the workplace to reduce the risk of cardiovascular disease--a systematic review. *Scand J Work Environ Health* 2010;36(3):202-215.

75. Verweij LM, Coffeng J, Van Mechelen W, Proper KI. Meta-analyses of workplace physical activity and dietary behaviour interventions on weight outcomes. *Obes Rev* 2011;12(6):406-429.

76. Nordin M. Corporate Wellness Programs Are a Secure Investment (Part 4). 2013. http://www.huffingtonpost.com/melissa-nordin/corporate-wellness-progra_b_2862830.html

77. Hooftman W, van der Klauw M, Klein Hesselink J, Jongen M, Kraan K, Wevers C, et al. *Arbobalans 2011.* Hoofddorp: TNO; 2011.

78. Downey AM, Sharp DJ. Why do managers allocate resources to workplace health promotion programmes in countries with national health coverage? *Health Prom Int* 2007;22(2):102-111.

79. KFF/HRET. Employer health benefits: 2012 Annual Survey. 2012. Menlo Park, Calif: Chicago, Ill, Kaiser Family Foundation/Health Research and Education Trust.

80. Mattke S, Seid M, Ma S. Evidence for the effect of disease management: is $1 billion a year a good investment? *Am J Manag Care* 2007;13(12):670-676.

81. Riketta M. Attitudinal organizational commitment and job performance: a meta-analysis. *J Organiz Behav* 2002 1;23(3):257-266.

82. Anderson DR, Sexner SA, Gold DB. Conceptual Framework, Critical Questions, and Practical Challenges in Conducting Research on the Financial Impact of Worksite Health Promotion. *Am J Health Promot* 2001;15(5):281-288.

83. Shaw W, Reme S, Boot CRL. Health and Wellness Promotion in the Workplace. In: Gatchel RJ, Schultz IZ, editors. Handbook of Occupational Health and Wellness. Springer US; 2012. p. 365-332.

84. Robroek SJW, Polinder S, Bredt FJ, Burdorf A. Cost-effectiveness of a long-term Internet-delivered worksite health promotion programme on physical activity and nutrition: a cluster randomized controlled trial. *Health Educ Res* 2012;27(3):399-410.

85. Driessen M, Bosmans J, Proper K, Anema J, Bongers P, Van der Beek A. The economic evaluation of a Participatory Ergonomics programme to prevent low back and neck pain. *Work: A Journal of Prevention, Assessment and Rehabilitation* 2012;41(0):2315-2320.

86. Meenan RT, Vogt TM, Williams AE, Stevens VJ, Albright CL, Nigg C. Economic evaluation of a worksite obesity prevention and intervention trial among hotel workers in Hawaii. *J Occup Environ Med* 2010;52(Suppl 1):S8-S13.

87. Groeneveld IF, van Wier MF, Proper K, Bosmans JE, Van Mechelen W, van der Beek A. Cost-effectiveness and cost-benefit of a lifestyle intervention for workers in the construction industry at risk for cardiovascular disease. *J Occup Environ Med* 2011;56:610–617.

88. van Wier MF, Verweij LM, Proper KI, Hulshof CTJ, van Tulder MW, van Mechelen W. Economic evaluation of an occupational health guideline for preventing overweight among employees. *J Occup Environ Med* 2013;55(9):1000-1109.

89. Oude Hengel OH. Sustainable employability of Construction Workers. Thesis (PhD). VU University Amsterdam, 2013.

90. McLaren L, Hawe P. Ecological perspectives in health research. *J Epidemiol Community Health* 2005;59:6-14.

91. Richard L, Gauvin L, Raine K. Ecological Models Revisited: Their Uses and Evolution in Health Promotion Over Two Decades*. Annu Rev Public Health* 2011;32(1):307-326.

92. Robroek S, van Lenthe F, van Empelen P, Burdorf A. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act* 2009;6:26.

93. Nutbeam D. Achieving ‘best practice’ in health promotion: improving the fit between research and practice. *Health Educ Res* 1996;11(3):317-326.

94. Walshe K, Rundall TG. Evidence-based Management: From Theory to Practice in Health Care. *Milbank Q* 2001;79(3):429-457.

95. Clancy CM, Cronin K. Evidence-Based Decision Making: Global Evidence, Local Decisions. *Health Aff (Millwood)* 2005;24(1):151-162.

96. Hoffmann C, Stoykova BA, Nixon J, Glanville JM, Misso K, Drummond MF. Do health-care decision makers find economic evaluations useful? The findings of focus group research in UK health authorities. *Value Health* 2002 Mar;5(2):71-78.

97. Drummond M, Brown R, Fendrick AM, Fullerton P, Neumann P, Taylor R, et al. Use of Pharmacoeconomics Information: Report of the ISPOR Task Force on Use of Pharmacoeconomic/Health Economic Information in Health-Care Decision Making. *Value Health* 2003;6(4):407-416.

98. Ross J. The use of economic evaluation in health care: Australian decision makers’ perceptions. *Health Policy* 1995;31(2):103-110.

99. Tompa E, Verbeek J, van Tulder MW, de Boer A. Developing guidelines for good practice in economic evaluation of occupational health and safety intervention. *Scand J Work Environ Health* 2010;36(4):313-318.

100. Arbouw. Bedrijfstaksatlas 2012. http://www.arbouw.nl/pdf/tools/bedrijfstak-atlas-2012.pdf

101. Schaufeli WB, Bakker AB. *Utrecht Work Engagement Scale*. Utrecht, the Netherlands: Occupational Health Psychology, Unit Utrecht University; 2003.

102. Jansen NWH, Kant IJ, van den Brandt PA. Need for Recovery in the Working Population: Description and Associations With Fatigue and Psychological distress. Int J Behavl Med 2001;9(4):322-340.

Summary

Samenvatting

About the author

List of publications

Dankwoord

SUMMARY

Background

The prevalence of modifiable health risks, such as an insufficient level of physical activity, unhealthy dietary habits, and a low level of relaxation, among the population is high. This imposes a large economic burden on society as a whole and on employers in particular. The workplace presents a useful setting for offering behavior change interventions aimed at preventing and/or reducing the prevalence of such risk factors (i.e. worksite health promotion programs). In practice, however, numerous kinds of worksite health promotion programs exist, of which only a restricted number can be provided with the available resources. Therefore, high quality evidence is needed to demonstrate their value. This evidence can be provided by performing methodologically sound economic evaluations of worksite health promotion programs, in which both the costs and consequences of alternatives are compared. Unfortunately, however, such studies are scarce. This is due to the fact that only a few of the studies that consider the effectiveness of worksite health promotion programs take the extra step of considering their resource implications. On top of that, the methodological quality of those that have been performed is generally poor, as is the uptake of their results in daily practice (*Chapter 1*). Therefore, this thesis aimed to contribute to the development of a sound evidence base on the resource implications of worksite health promotion programs as well as to improve uptake of this evidence in daily practice. This was done by summarizing the current literature on the cost-effectiveness and financial return of worksite physical activity and/or nutrition programs, generating new evidence by performing economic evaluations of various newly developed worksite health promotion programs, as well as developing and providing recommendations for good practice when conducting and disseminating economic evaluations in occupational health.

Part 1: Systematic reviews

Chapter 2 described a systematic review on the cost-effectiveness of worksite physical activity and/or nutrition programs. A literature search was performed in EMBASE, MEDLINE, SportDiscus, PsycInfo, NIOSHTIC-2, NHSEED, HTA, and Econlit

for studies published up to January 14th, 2011. Additionally, articles were searched by reviewing references, searching authors' databases, and contacting authors of included studies. Ten studies were found to be eligible for inclusion, of which four evaluated worksite nutrition programs (seven programs) and six worksite physical activity and nutrition programs (eleven programs). A risk of bias assessment indicated that the methodological quality of the included studies was generally poor. From various perspectives, all worksite nutrition as well as worksite physical activity and nutrition programs (N=6) were more costly and more effective in reducing body weight compared to usual care. When only intervention costs were considered, most worksite nutrition (N=4/5) and worksite physical activity and nutrition programs (N=5/6) were more costly and more effective in reducing cholesterol level and cardiovascular disease risks. Currently, however, there are no set levels as to how much various kinds of decision-makers are willing to pay per unit of improvement in these outcomes. One of the included studies also evaluated the cost-utility of two different delivery modes of a worksite physical activity and nutrition program (i.e. telephone-based and internet-based), and provided mixed results. That is, when comparing its results with various pre-established thresholds regarding the amount of money decision-makers are willing to pay per QALY gained, the internet-based intervention could be regarded as cost-effective ($1,698/QALY gained), whereas the phone-based intervention ($311,523/QALY gained) could not. Thus, based on the current literature, strong conclusions about the cost-effectiveness of worksite physical activity and/or nutrition programs could not be made and there seemed to be an urgent need to improve the methodological quality of such studies.

Chapter 3 described a systematic review on the financial return of worksite physical activity and/or nutrition programs. In order to identify relevant studies, a literature search was performed in eight electronic databases (EMBASE, MEDLINE, SportDiscus, PsycInfo, NIOSHTIC-2, NHSEED, HTA, and Econlit), references of relevant review articles as well as authors' own databases were searched, and authors of included studies were contacted. Eventually, 18 studies were included in the review, of which four were performed alongside a randomized controlled trial (RCT), 13 alongside a non-randomized study (NRS), and one was a modeling study. For all included studies, three metrics were (re-)calculated, including the Net Benefits, Benefit Cost Ratio (BCR), and Return On Investment (ROI). These metrics were averaged and a subgroup analysis was performed to compare them between study designs (i.e. RCT versus NRS). Additionally, a risk of bias assessment was performed to assess the methodological quality of the studies. The results showed that average financial return estimates in terms of absenteeism benefits (NRS: ROI 325%, BCR 4.25; RCT: ROI -49%, BCR 0.51), medical benefits (NRS: ROI 95%, BCR 1.95; RCT: ROI -112%, BCR -0.12), or both (NRS: ROI 387%, BCR 4.87; RCT: ROI -92%, BCR 0.08) were positive in NRSs, but negative in RCTs. Moreover, the methodological quality of the included studies was generally poor, and even poorer in NRSs than in RCTs. These results indicate that financial return estimates of NRSs are likely to be distorted by selection bias (i.e. study results are caused by (baseline) differences between study arms, rather than by the intervention itself). Financial return estimates derived from NRSs should therefore be interpreted with great caution. RCTs with a lower risk of bias, on the other hand, indicated that worksite physical activity and/or nutrition programs may not pay for themselves in terms of reduced absenteeism costs, medical costs, or both. However, since worksite physical activity and/or nutrition programs are thought to be associated with additional types of benefits (e.g. reduced on-the-job productivity costs, also known as presenteeism benefits), conclusions about their overall profitability could not be made. Therefore, it is advisable to perform more RCT-based ROI analyses that include a consensus-based set of financial benefits.

Part 2: Applied studies

In order to generate new evidence, four economic evaluations were performed. All of them evaluated a newly developed worksite health promotion program in comparison with the existing health promotion activities of the participating companies (i.e. usual care / usual practice). In all studies, missing values were multiply imputed and uncertainty was assessed using bootstrapping techniques.

Chapter 4 presented the economic evaluation results of the Vital@Work intervention, a worksite health promotion program aimed at improving physical activity, nutrition, and relaxation, as a potentially effective tool to keep older workers vital and healthy, and thereby contributing to prolonged employability. The objective was to perform a cost-effectiveness analysis (CEA) in terms of general vitality, work-related vitality, and need for recovery (NFR) from the societal perspective and a ROI analysis from that of the employer. Within this study, a total of 730 older hospital workers (≥ 45 years) were randomized to an intervention (n = 367) or control group (n = 363). Effect data were collected at baseline, 6-, and 12-month follow-up. Cost data were collected on a 3-monthly basis using questionnaires. The cost of the Vital@Work intervention was found to be €149 per participant. After 12 months, no statistically significant between-group differences were found for all other cost and effect measures. A joint comparison of costs and effects revealed that a substantial amount of money had to be paid by society to reach a reasonable probability of cost-effectiveness for all outcomes (e.g. ± €3,500 per 1-point improvement in general vitality (range: 0-100) for a 0.9 probability). Moreover, per Euro invested in the program, the employer was found to lose €2.21. Based on these results, it was concluded that the Vital@Work intervention was neither cost-effective from the societal perspective, nor cost-saving from that of the employer. Thus, the economic evaluation provided no evidence to support its implementation.

Chapter 5 described the economic evaluation of the Mindful VIP intervention, a mindfulness-based worksite intervention aimed at improving work engagement among knowledge workers. This economic evaluation aimed to evaluate the intervention's cost-effectiveness in comparison with usual practice from both the societal and employer's perspective. Moreover, a ROI analysis was performed to explore the intervention's impact on the company's bottom line. A total of 257 employees of two Dutch governmental research institutes were randomized to the intervention (n = 129) or control group (n = 128). Data on work engagement, general vitality, job satisfaction, and work ability were collected at baseline, six, and 12 months. Salary and absence data were collected from company records. Data on all other cost measures were collected using 3- or 6-monthly questionnaires. The cost of the Mindful VIP intervention was found to be €171 per participant from the societal perspective (estimated using bottom-up micro-costing) and €464 from that of the employer (based on market prices). After 12 months, a statistically significant but not clinically relevant adverse effect on work engagement was found (-0.19; 95%CI -0.38 to -0.01; i.e. a decrease of 0.19 on a scale from 0 to 6). There were no statistically significant differences between study groups in job satisfaction (-0.02; 95%CI -0.22

to 0.17), general vitality (-3.0; 95%CI -6.1 to 0.1), work ability (-0.34; 95%CI -0.84 to 0.17), and total costs (societal: 1,814; 95%CI -800 to 4,588, employer: 2,038; 95%CI -548 to 4,752). Moreover, the intervention's maximum probability of cost-effective was low for all outcomes (≤ 0.25) and the intervention did not result in a positive financial return to the employer. Based on these results, the Mindful VIP intervention could neither be considered cost-effective from both the societal and employer's perspective, nor cost-saving from that of the employer. Thus, this study provided no evidence to support its implementation.

Chapter 6 presented the economic evaluation results of the VIP in Construction intervention, a worksite health promotion program aimed at improving physical activity and nutrition among construction workers. The study aimed to explore the intervention's cost-effectiveness in comparison with usual practice from the societal and employer's perspective, as well as its financial return to the employer. Within this study, 314 construction workers were randomized to the intervention (n = 162) or control group (n = 152). Data on body weight, waist circumference, musculoskeletal disorders (MSD), work-related vitality, and job satisfaction were collected at baseline, six, and 12 months. Sickness absence data were collected from company records. Data on all other cost measures were collected using 3-monthly questionnaires. From the societal perspective, the cost of the VIP in Construction intervention was found to be €178 per construction worker (bottom-up micro-costed). From the employer's perspective, these costs were €287 (market prices). At 12-month follow-up, no statistically significant cost and effect differences were found between groups. Results also indicated that the intervention's probabilities of cost-effectiveness for body weight, waist circumference, and MSD gradually increased with an increasing willingness-to-pay to 0.84 (willingness-to-pay = €21,000/kg), 0.77 (willingness-to-pay = €18,000/cm), and 0.84 (willingness-to-pay = €42,000/person prevented from having a MSD), respectively. The intervention's maximum probabilities of cost-effectiveness for work-related vitality and job satisfaction were low at all ceiling ratios (≤ 0.54) and financial return estimates were positive, but their confidence intervals were rather wide and none of them was statistically significant. Based on these results it was concluded that the intervention's cost-effectiveness in improving weight-related outcomes and MSD depends on the societal and employer's willingness to pay for these effects as well as the probability of cost-effectiveness that they consider acceptable. From the employer´s perspective, the intervention was not cost-effective in improving work-related vitality and job satisfaction. Also, due to a high level of uncertainty, it could not be concluded that the intervention was cost-beneficial to the employer.

Chapter 7 described the economic evaluation of the Be Active & Relax intervention. The objective was to evaluate the cost-effectiveness and financial return of a combined social and physical environmental intervention in office employees in comparison with usual practice, and of both intervention conditions separately. Moreover, the probabilities of the intervention conditions being cost-effective in comparison with each other were explored. This study used a 2X2 factorial design, in which 412 employees were allocated at the department level to the combined intervention (n = 92), social environmental intervention (n = 118), physical environmental intervention (n = 96), or control group (n = 106). Data on NFR, general vitality, and job satisfaction were collected at baseline, 6-, and 12-month follow-up. Salary and sickness absence data were collected from company records. Data on all other cost measures were collected using 3-monthly questionnaires. Using linear multilevel analyses, CEAs were performed from the societal (NFR and general vitality) and employer's perspective (NFR and job satisfaction), and ROI analyses from that of the employer. After 12 months, combined intervention group participants statistically significantly improved their NFR compared with the control group (-8.4; 95%CI -14.6 to -2.2). Their total employer's costs, however, were statistically significantly higher than those of the control group (3,102; 95%CI 598 to 5,969). All other between-group differences in costs and effects were not statistically significant. For NFR, the combined intervention became the preferred option at willingness-to-pay values of €170 (societal perspective) and €300 (employer's perspective) per point improvement, after which its probability of cost-effectiveness gradually increased to 0.85. For general vitality and job satisfaction, the maximum probabilities of the interventions being cost-effective in comparison with each other were low at all ceiling ratios (≤ 0.55), as were their probabilities of financial return (≤ 0.41). Depending on the societal and employer's willingness to pay and the probability of cost-effectiveness that they consider acceptable, the combined intervention may be considered cost-effective in improving NFR. However, both separate interventions were not cost-effective in improving this outcome. Furthermore, all interventions were neither cost-effective in improving general vitality (societal perspective) and job satisfaction (employer's perspective), nor cost-saving to the employer.

Part 3: Improving evidence-based practice

As previous research indicates that the methodological quality of economic evaluations in occupational health is generally poor, *chapter 8* aimed to help occupational health researchers conduct high quality trial-based economic evaluations. This was done by discussing the theory and methodology that underlie them and by providing recommendations for good practice regarding their design, analysis, and reporting. Amongst others, it was recommended to consider the economic evaluation requirements at the earliest stage possible and to perform such evaluations alongside studies with a randomized design. Within these studies, careful considerations must be made regarding the perspective, the analytic time frame, the identification, measurement, and valuation of resource use and outcomes, as well as the methods used for calculating sample sizes, comparing costs and consequences, and handling missing data and uncertainty. The latter is of particular importance, as few economic evaluations in occupational health report on the uncertainty surrounding their incremental cost-consequence estimates, whereas failing to estimate values under uncertainty makes it impossible to determine the certainty of results and could thus lead to inappropriate decision-making.

As a first step in bridging the gap between the economic evaluation literature and daily practice in occupational health, *chapter 9* aimed to explore four issues; 1) the process by which occupational health decisions are made, 2) the importance given to the financial implications of occupational health and safety (OHS) interventions, 3) the sources of information used during the decision-making process, and 4) the occupational health decision-makers' knowledge about different economic evaluation methods. This was done by performing 18 in-depth and 25 structured interviews with occupational health decision-makers in the healthcare sector of Ontario, Canada. The analyses showed that the occupational health decision-making process could generally be subdivided into three stages: 1) *initiation stage*, during which the need for an intervention is established, 2) *pre-implementation stage*, during which and intervention as well as its business case are developed in order to achieve senior management approval, and 3) *implementation and evaluation stage*, during which an intervention is implemented and evaluated. In line with previous research, organizations were found to invest in OHS interventions for legal, financial, and moral reasons. Moreover, information on the financial implications of such interventions was found to be of great importance to the decision-making process, particularly the employer's costs and benefits. Results also indicated that occupational health decisions are currently not being made in an evidence-based manner. That is, scientific evidence on the (financial) implications of OHS interventions was found to be rarely consulted and sound ex-post program evaluations were hardly ever performed. Furthermore, there seemed to be a need to advance the decision-makers' economic evaluation skill set. Possible strategies to overcome these issues may include the generation of scientific evidence that is in line with the needs of occupational health decision-makers (e.g. ROI analyses performed from the employer's perspective), providing busy decision-makers with critical summaries of published studies, transmitting (economic evaluation) results in easy-to-use formats, and educating occupational health decision-makers in economic evaluation methods.

Discussion

In *chapter 10*, the main findings were discussed and interpreted, and recommendations for research and practice were presented. In conclusion, the present thesis indicated that (strong) evidence of the cost-effectiveness and/or financial return of worksite health promotion programs is currently lacking. Therefore, widespread implementation of such interventions in an effort to generate cost savings is not recommended, while some of them may be considered cost-effective if decision-makers are willing to invest a certain amount of money to improve employee health. Whether the latter is the case is currently unknown. The lack of evidence of cost savings associated with existing worksite health promotion programs, however, does not negate the value of improving employee health. Therefore, future research should explore what attributes of worksite health promotion programs are most important and how such interventions should be optimally designed. The cost-effectiveness and/or financial return of such "optimally" designed interventions should subsequently be established by performing (cluster-)RCT-based economic evaluations. Furthermore, the methodological quality of economic evaluations of worksite health promotion programs was found to be generally poor, as was the uptake of their results in daily practice. To prevent inappropriate decision-making, (occupational health) researchers as well as other relevant stakeholders should ensure that both issues are addressed.

SAMENVATTING

Achtergrond

Veel mensen hebben ongezonde leefgewoonten. Zo beweegt en ontspant men tegenwoordig gemiddeld gezien te weinig en hebben veel mensen een ongezond voedingspatroon. Dergelijke ongezonde gewoonten hebben negatieve economische gevolgen voor zowel de samenleving als werkgevers. De werksetting biedt een goede omgeving voor interventies die mensen ondersteunen bij het maken van gezonde keuzes, zogenaamde leefstijlprogramma's op de werkplek. In de praktijk bestaan er echter veel verschillende typen leefstijlprogramma's, die niet allemaal geïmplementeerd kunnen worden met de beschikbare middelen. Daarom is wetenschappelijk bewijs nodig dat inzicht geeft in de financiële gevolgen van dergelijke interventies. Dit bewijs kan geleverd worden met behulp van hoge kwaliteit economische evaluaties, welke zowel de kosten als de (gezondheids-)effecten van verschillende interventieopties met elkaar vergelijken. Op dit moment zijn zulke studies echter schaars. Daarnaast worden de resultaten van eerder uitgevoerde economische evaluaties in de dagelijke (bedrijfsgezondheidszorg) praktijk nauwelijks gebruikt (*Hoofdstuk 1*). Dit proefschrift had daarom tot doel de huidige literatuur met betrekking tot de kosteneffectiviteit en de financiële opbrengsten van beweeg- en/of voedingsinterventies op de werkplek samen te vatten, nieuw bewijs te genereren door economische evaluaties uit te voeren van vier recent ontwikkelde leefstijlinterventies voor diverse groepen werknemers, en aanbevelingen te genereren voor het verrichten en verspreiden van economische evaluaties binnen de bedrijfsgezondheidszorg.

Deel 1: Systematische literatuuronderzoeken

Hoofdstuk 2 beschrijft een systematisch literatuuronderzoek naar de kosteneffectiviteit van beweeg- en/of voedingsinterventies op de werkplek. Om relevante studies te vinden, zijn acht bronnen van wetenschappelijke literatuur doorzocht (EMBASE, MEDLINE, SportDiscus, PsycInfo, NIOSHTIC-2, NHSEED, HTA, en Econlit). Tevens zijn referentielijsten van relevante literatuuronderzoeken nagekeken, hebben wij onze eigen literatuur databases doorzocht, en zijn auteurs van geïncludeerde studies aangeschreven om eventuele nog niet gepubliceerde artikelen te vinden. Uiteindelijk zijn 10 relevante studies geïdentificeerd, waarvan vier een voedingsinterventie evalueerden (zeven interventies) en zes een beweeg- en voedingsinterventie (elf interventies). De methodologische kwaliteit van deze studies bleek over het algemeen slecht te zijn. Vanuit diverse perspectieven bleken alle beweeg- en voedingsinterventies (N=6) duurder en effectiever te zijn ten opzichte van gebruikelijke zorg in het verminderen van lichaamsgewicht. Daarnaast bleken zowel de meeste voedingsinterventies (N=4/5) als de meeste beweeg- en voedingsinterventies (N=5/6) duurder en effectiever te zijn in het verminderen van cholesterol level en het aantal cardiovasculaire risicofactoren van de deelnemers. Voor deze uitkomsten is het echter onbekend hoeveel diverse beslissers bereid zijn te betalen per extra eenheid in effect. Een van de geïncludeerde studies had ook de kostenutiliteit van twee typen beweeg- en voedingsinterventies onderzocht (een telefoonprogramma en een internetprogramma). De resultaten van beide interventies liepen sterk uiteen. Als wij hun investering die gedaan moet worden om één voor kwaliteit van leven gecorrigeerd levensjaar te winnen (QALY) vergelijken met enkele in de literatuur gebruikte drempelwaarden, kan het internetprogramma als kosteneffectief worden beschouwd ($1.698 / gewonnen QALY), maar het telefoonprogramma niet ($311.523 / gewonnen QALY). Op basis van deze resultaten hebben wij geen eenduidige conclusie kunnen trekken over de kosteneffectiviteit van beweeg- en/of voedingsinterventies op de werkplek. Daarnaast is het van groot belang dat de methodologische kwaliteit van dergelijke studies verbeterd wordt, vooral het in kaart brengen van de onzekerheid rondom de uitkomsten.

Hoofdstuk 3 beschrijft een systematisch literatuuronderzoek naar de financiële opbrengsten van beweeg- en/of voedingsinterventies op de werkplek. Om relevante studies te vinden zijn wederom acht bronnen van wetenschappelijke literatuur doorzocht (EMBASE, MEDLINE, SportDiscus, PsycInfo, NIOSHTIC-2, NHSEED, HTA, en Econlit). Daarnaast hebben wij zowel referentielijsten van relevante literatuuronderzoeken als onze eigen literatuur databases nagekeken en zijn auteurs van geïncludeerde studies aangeschreven om eventuele nog niet gepubliceerde artikelen te vinden. Uiteindelijk werden 18 relevante studies geïdentificeerd, waaronder 13 niet-gerandomiseerde studies (NRSs), vier gerandomiseerde studies

(RCTs) en één modelleringstudie. Twee onderzoekers hebben onafhankelijk van elkaar de methodologische kwaliteit van deze studies beoordeeld. Voor alle studies werden de gemiddelde Netto Baten (NB), Baten Kosten Ratio (BKR), en Return On Investment (ROI) (her-)berekend. Ook is een subgroepanalyse uitgevoerd om deze uitkomsten te vergelijken tussen verschillende typen studie designs (RCT versus NRS). De resultaten van het onderzoek lieten zien dat de gemiddelde ROI uitkomsten in termen van verzuimbaten [(NRS, ROI:325%; BKR:4,25)(RCT, ROI:-49%; BKR:0,51)], medische baten [(NRS, ROI:95%; BKR:1,95)(RCT, ROI:-112%; BKR:-0,12)], en zowel medische als verzuimbaten [(NRS, ROI:387%; BKR:4,87)(RCT, ROI:-92%; BKR:0,08)] positief waren in NRSs, maar negatief in RCTs. Ook was de methodologische kwaliteit van de studies over het algemeen slecht en zelfs slechter in NRSs dan in RCTs. Concluderend kan gesteld worden dat de resultaten van NRSs mogelijk vertekend zijn door selectiebias; i.e. een vertekening van de onderzoeksresultaten als gevolg van het feit dat onderzoeksgroepen (bij aanvang) niet gelijk aan elkaar waren. Aan de andere kant laten RCTs zien dat beweeg- en/of voedingsinterventies op de werkplek mogelijk geen winst opleveren als gevolg van verminderde medische, verzuim- en zowel medische als verzuimkosten. Op basis van de huidige literatuur kunnen echter geen conclusies getrokken worden over hun algehele winstgevendheid, omdat dergelijke interventies ook met andere typen baten geassocieerd zijn (bijvoorbeeld baten als gevolg van een verbeterde productiviteit op het werk).

Deel 2: toegepaste studies

Om nieuw bewijs te genereren zijn in dit proefschrift vier economische evaluaties verricht. In deze economische evaluaties zijn zowel de kosteneffectiviteit als de financiële opbrengsten van enkele recent ontwikkelde leefstijlprogramma's op de werkplek onderzocht in vergelijk met het gebruikelijke aanbod. In alle studies zijn missende waarden multipel geïmputeerd en is de onzekerheid rondom de kostenuitkomsten in kaart gebracht met behulp van zogenaamde bootstrapping technieken.

Hoofdstuk 4 beschrijft de economische evaluatie van de Vital@Work interventie, een leefstijlprogramma voor oudere werknemers dat erop gericht is om de vitaliteit van deze groep werknemers te bevorderen. In deze studie is zowel een kosteneffectiviteitsanalyse (KEA) verricht vanuit maatschappelijk perspectief als een "return on investment" analyse (ROI analyse) vanuit het perspectief van de werkgever. De onderzoeksgroep bestond uit 730 oudere werknemers (≥ 45 jaar) die werkzaam waren bij twee Nederlandse academische ziekenhuizen. Na de nulmeting werden alle werknemers op basis van kans verdeeld over een interventie- (n = 367) en een controlegroep (n = 363). Ten behoeve van de KEA werden tijdens de nulmeting en na zes en 9 maanden diverse uitkomsten gemeten (algemene vitaliteit, werkgerelateerde vitaliteit en herstelbehoefte). Daarnaast is kostendata verzameld met behulp van 3-maandelijkse vragenlijsten. Uit de analyses bleek dat de Vital@Work interventie €149 per werknemer kostte. Er werden geen statistisch significante kosten- en effectverschillen gevonden tussen de interventie- en controlegroep. Ook bleken maatschappelijke beslissers bereid te moeten zijn om een relatief hoog bedrag te betalen per extra eenheid effect om een hoge kans op kosteneffectiviteit te bewerkstelligen (e.g. ± €3.500 per punt verbetering in algemene vitaliteit voor een kans van 0,90). Tevens bleek de interventie per geïnvesteerde Euro voor de werkgever tot een verlies van €2,21 te leiden. Op basis van deze resultaten is geconcludeerd dat de Vital@Work interventie niet kosteneffectief is vanuit maatschappelijk perspectief en dat invoering ervan geen financiële winst oplevert voor de werkgever. Op basis van deze resultaten werd grootschalige implementatie van de Vital@Work interventie dan ook afgeraden.

De resultaten van de economische evaluatie van de Mindful VIP interventie zijn gepresenteerd in *hoofdstuk 5*. Deze interventie had tot doel de bevlogenheid van medewerkers te bevorderen en omvatte een mindfulness training, e-coaching, en het aanbieden van werkfruit en lunchwandelroutes. Het doel van deze studie was om de kosteneffectiviteit van de interventie te bepalen vanuit zowel het maatschappelijke als het bedrijfsperspectief. Daarnaast zijn de eventuele financiële opbrengsten van de interventie voor de werkgever in kaart gebracht. Bij aanvang van de studie deden 257 werknemers mee, welke op basis van kans over een interventie- (n = 129) en een controlegroep (n = 128) zijn verdeeld. Voor de KEA, zijn tijdens de nulmeting en na zes en 9 maanden diverse uitkomsten gemeten (bevlogenheid, algemene vitaliteit, werktevredenheid en werkvermogen). Informatie over het verzuim en het salaris van de werknemers werd verzameld vanuit de bedrijfsregistratie. Informatie over alle overige kosten werd verzameld met behulp van 3- of 6-maandelijkse vragenlijsten. De interventie bleek vanuit het maatschappelijke perspectief €171 per werknemer te kosten (gebaseerd op een bottum-up kostprijsbepaling) en vanuit het bedrijfsperspectief €464 (gebaseerd op facturen). Na 12 maanden werd een klein statistisch significant negatief effect gevonden op bevlogenheid (-0,19; 95%CI -0,38 tot -0,01, dat wil zeggen een daling van 0,19 op een schaal van 0 tot 6). Er werden geen statistisch significante verschillen gevonden tussen de onderzoeksgroepen in werktevredenheid (-0,02; 95%CI -0,22 tot 0,17), algemene vitaliteit (-3,0; 95%CI -6,1 tot 0,1), werkvermogen (-0,34; 95%CI -0,84 tot 0,17), en totale kosten (maatschappelijk perspectief: 1.814; 95%CI -800 tot 4.588, bedrijfsperspectief: 2.038; 95%CI -548 tot 4.752). Ook was de maximale kans op kosteneffectiviteit voor alle uitkomsten laag (≤ 0,25) en leverde de interventie geen financiële winst op voor de werkgever. Op basis van deze resultaten werd grootschalige invoering van de Mindful VIP interventie dan ook afgeraden.

Hoofdstuk 6 beschrijft de economische evaluatie van de VIP in Construction interventie, een beweeg- en voedingsinterventie voor bouwvakkers. Het doel van deze studie was om de kosteneffectiviteit van deze interventie te bepalen vanuit zowel het maatschappelijke als het bedrijfsperspectief. Daarnaast is een ROI analyse uitgevoerd om de eventuele financiële opbrengsten voor de werkgever in kaart te brengen. Aan de start van de studie deden 314 bouwvakkers mee, welke op basis van kans zijn verdeeld over een interventiegroep (n = 162) en een controlegroep (n = 152). Voor de KEA, zijn tijdens de nulmeting en na zes en 9 maanden diverse effectmaten gemeten (lichaamsgewicht, middelomtrek, klachten aan het bewegingsapparaat, werkgerelateerde vitaliteit en werktevredenheid). Verzuimdata werd verzameld vanuit de bedrijfsregistratie. Alle overige kostendata werd verzameld met behulp van 3-maandelijkse vragenlijsten. Vanuit het maatschappelijke perspectief bleek de interventie €178 per bouwvakker te kosten (bottom-up kostprijsbepaling) en vanuit het bedrijfsperspectief €287 (facturen). Na 12 maanden werden geen statistisch significante verschillen gevonden in alle kosten en effectmaten. Voor de uitkomsten lichaamsgewicht, middelomtrek en klachten aan het bewegingsapparaat bleek de kans op kosteneffectiviteit geleidelijk toe te nemen met een toenemende betalingsbereidheid tot respectievelijk 0,84 (betalingsbereidheid: €12.000/kilogram), 0,77 (betalingsbereidheid: €18.000/kilogram) en 0,84 (betalingsbereidheid: €42.000/ per persoon minder met klachten aan het bewegingsapparaat). Voor werkgerelateerde vitaliteit en werktevredenheid was de maximale kans op kosteneffectiviteit echter laag (≤ 0,54). Ook bleek de interventie gemiddeld genomen per geïnvesteerde Euro, €1,48 op te leveren voor de werkgever. De onzekerheid rondom deze schatting was echter erg groot. Op basis van deze resultaten is geconcludeerd dat de kosteneffectiviteit van de VIP in Construction interventie voor lichaamsgewicht, middelomtrek, en klachten aan het bewegingsapparaat afhangt van zowel de betalingsbereidheid van beslissers als de kans op kosteneffectiviteit die zij acceptabel achten. De interventie was echter niet kosteneffectief voor werkgerelateerde vitaliteit en werktevredenheid en deze leek ook geen financiële winst op te leveren voor de werkgever.

De resultaten van de economische evaluatie van de Be Active & Relax interventie zijn gepresenteerd in *hoofdstuk 7*. De Be Active & Relax interventie was een gecombineerd programma gericht op het verbeteren van zowel de sociale als de fysieke omgeving van werknemers. In deze studie zijn de kosteneffectiviteit en de eventuele financiële opbrengsten van dit gecombineerde programma onderzocht, en dat van beide losse onderdelen (een sociaal omgevingsprogramma en een fysiek omgevingsprogramma). Daarnaast is de kans dat de diverse interventieopties kosteneffectief waren ten opzichte van elkaar onderzocht. Deze studie gebruikte een zogenaamd 2X2 factorieel design, waarin 412 medewerkers van een verzekeringsbedrijf op afdelingsniveau zijn verdeeld over een gecombineerde omgevingsgroep (n = 92), een sociale omgevingsgroep (n = 118), een fysieke omgevingsgroep (n = 96) en een controlegroep (n = 106). Voor de KEA, zijn tijdens de nulmeting en na zes en 9 maanden diverse effectmaten gemeten (herstelbehoefte, algemene vitaliteit en werktevredenheid). Verzuim- en salarisgegevens van de deelnemers werden verzameld vanuit de bedrijfsregistratie. Alle overige kostendata zijn verzameld met behulp van 3-maandelijkse vragenlijsten. Gebruikmakend van lineaire multilevel analyses zijn KEAs verricht vanuit zowel het maatschappelijke (herstelbehoefte en algemene vitaliteit) als het bedrijfsperspectief (herstelbehoefte en werktevredenheid), en een ROI analyse vanuit het perspectief van de werkgever. In de gecombineerde groep werd na 12 maanden een significante verbetering in herstelbehoefte gevonden ten opzichte van de controlegroep (-8,4; 95%CI -14,6 tot

-2,2). De totale werkgeverskosten waren binnen deze groep echter wel statistisch significant hoger dan in de controlegroep (3,102; 95%CI 598 tot 5,969). Voor alle overige kosten en effectmaten werden geen statistisch significante verschillen gevonden. Voor herstelbehoefte, bleek het gecombineerde programma de voorkeursoptie te worden bij een betalingsbereidheid van €170 (maatschappelijk perspectief) en €300 (bedrijfsperspectief) per punt verbetering, waarna de kans op kosteneffectiviteit geleidelijk toenam met een toenemende betalingsbereidheid tot 0,85. Voor algemene vitaliteit en werktevredenheid was de maximale kans op kosteneffectiviteit voor alle interventieopties relatief laag (≤ 0,55). Dit was tevens het geval voor hun maximale kans op financiële opbrengsten (≤ 0,41). Geconcludeerd werd dat de kosteneffectiviteit van het gecombineerde programma voor herstelbehoefte afhangt van de betalingsbereidheid van beslissers en de kans op kosteneffectiviteit die zij acceptabel achten. De losse programma's kunnen daarentegen niet als kosteneffectief worden geschouwd in termen van deze uitkomst. Tevens bleek geen van de programma's kosteneffectief te zijn voor algemene vitaliteit (maatschappelijk perspectief) en werktevredenheid (bedrijfsperspectief) en leverde geen enkel programma financiële winst op voor de werkgever.

Deel 3: Verbeteren van "Evidence-based practice"

Omdat de methodologische kwaliteit van economische evaluaties binnen de bedrijfsgezondheidszorg doorgaans slecht is, zijn in *hoofdstuk 8* aanbevelingen gedaan met betrekking tot het opzetten, analyseren en rapporteren van dergelijke studies. Aanbevolen werd om de randvoorwaarden van een economische evaluatie in een zo vroeg mogelijk stadium mee te nemen in de opzet van een trial en om dergelijke studies uit te voeren met behulp van een gerandomiseerd design. Daarnaast moeten doordachte keuzes gemaakt worden met betrekking tot het perspectief van de studie, de duur van de dataverzameling, het identificeren, meten en waarderen van kosten en effecten, de gebruikte methodes voor het berekenen van het benodigde aantal deelnemers en het omgaan met missende waarden een onzekerheid. Met name het laatste is van belang, omdat het niet in kaart brengen van de onzekerheid rondom economische evaluatie resultaten kan leiden tot verkeerde investeringsbeslissingen.

Om het gebruik van economische evaluaties binnen de bedrijfsgezondheidszorg te verbeteren zijn in *hoofdstuk 9* vier verschillende zaken onderzocht; 1) het besluitvormingsproces rondom het invoeren van interventies op de werkplek, 2) het belang van informatie over de financiële gevolgen van dergelijke interventies, 3) de informatiebronnen die gebruikt worden tijdens het besluitvormingsproces, en 4) de economische kennis van de beslissers. Bovenstaande zaken zijn onderzocht met behulp van 18 diepte en 25 gestructureerde interviews met beslissers die werkzaam waren in de gezondheidszorgsector van Ontario, Canada. Uit de analyses bleek dat het besluitvormingsproces over het algemeen onderverdeeld kan worden in 3 stadia: 1) *initiatie stadium*, tijdens welke de vraag naar een bepaalde interventie werd vastgesteld, 2) *pre-implementatie stadium*, tijdens welke de interventie en de bijbehorende “business case” werden ontwikkeld om toestemming voor implementatie te krijgen van het senior management, 3) *implementatie en evaluatie stadium*, tijdens welke de interventie ingevoerd en geëvalueerd werd. In overeenstemming met voorgaand onderzoek, bleek uit de analyses dat interventies op de werkplek over het algemeen ingevoerd worden vanuit wettelijke, morele en financiële overwegingen. Informatie over de financiële gevolgen van dergelijke interventies bleek een grote invloed te hebben op het besluitvormingsproces, met name informatie over hun kosten en baten voor de werkgever. De resultaten lieten tevens zien dat keuzes met betrekking tot het al dan niet invoeren van interventies op de werkplek bijna nooit gebaseerd zijn op wetenschappelijk bewijs en dat hun (kosten) effectiviteit en financiële opbrengsten na afloop nauwelijks worden geëvalueerd. Ook bleken beslissers weinig economische kennis te hebben. Bovenstaande tekortkomingen kunnen mogelijk aangepakt worden door wetenschappelijk bewijs te genereren dat in overeenstemming is met de informatiebehoeften van relevante beslissers (ROI analyses vanuit het bedrijfsperspectief), samenvattingen te maken van wetenschappelijke bewijs en beslissers te onderwijzen in economische evaluatietechnieken.

Discussie

In *hoofdstuk 10* zijn de belangrijkste bevingen en de methodologisch sterke en zwakke punten van dit proefschrift besproken. Daarnaast zijn aanbevelingen gedaan voor zowel de praktijk als verder onderzoek. Op basis van dit proefschrift kan geconcludeerd worden dat (sterk) bewijs voor het kosteneffectief zijn van leefstijlprogramma's op de werkplek momenteel ontbreekt. Deels komt dit doordat het voor veel uitkomstenmaten onbekend is hoeveel beslissers bereid zijn te betalen per extra eenheid effect. Daarnaast lijkt het erop dat leefstijlinterventies op de werkplek in hun huidige vorm geen winst opleveren voor de werkgever. Het afwezig zijn van dergelijk bewijs betekend echter niet dat gezondheidsbevordering op de werkplek onbelangrijk is. Het is daarom raadzaam om verder onderzoek te doen naar wat de meest effectieve onderdelen van leefstijlinterventies op de werkplek zijn en hoe dergelijke programma's zo optimaal mogelijk ontworpen kunnen worden. De kosteneffectiviteit en financiële opbrengsten van dergelijke "optimale" interventies moeten vervolgens onderzocht worden met behulp van gerandomiseerde studies. Daarnaast bleek zowel de methodologische kwaliteit als het gebruik van economische evaluaties binnen de bedrijfsgezondheidszorg slecht te zijn. Om het maken van verkeerde investeringsbeslissingen te voorkomen is het van belang dat onderzoekers, mogelijk in samenwerking met andere belanghebbenden, beide problemen aanpakken.

ABOUT THE AUTHOR

Dimitri Valentijn Photography

Johanna Maria (Hanneke) was born on December 30th 1982 in Haarlem, the Netherlands. After completing secondary school at the Eerste Christelijk Lyceum in Haarlem in 2001, she started studying at the VU University in Amsterdam. In 2008, she received a Master degree in Health Sciences with a specialization in 'Policy and Organization of Healthcare' and a Master degree in Human Movement Sciences with a specialization in 'Psychomotor Therapy'.

Shortly after her study, she worked as a junior researcher at the Department of Public and Occupational Health of the EMGO+ Institute for Health and Care Research and the VU Medical Center. There she worked on a process evaluation of 'Hello World', an email-based health promotion program for pregnant women in the Netherlands. In 2009, she started her PhD project at the Department of Health Sciences of the EMGO+ Institute and the VU University. During her PhD, she attended several courses within the Postgraduate Epidemiology Program at the VU Medical Center as well as the Centre for Health Economics at the University of York, York, United Kingdom. During the summer of 2012, she visited the Institute for Work and Health in Toronto to work on a qualitative study into the occupational health decision-making process. On a regular basis, she was a teaching assistant for several economic evaluation and health economics courses. In 2013, Hanneke started working as a postdoctoral researcher at the Department of Health Sciences of the EMGO+ Institute and the VU University.

LIST OF PUBLICATIONS

Full Papers

1) **JM van Dongen**, KI Proper, MF van Wier, AJ van der Beek, PM Bongers, W van Mechelen, MW van Tulder. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev* 2011; 12: 1031–1049.

2) **JM van Dongen**, KI Proper, MF van Wier, AJ van der Beek, PM Bongers, W van Mechelen, MW van Tulder. A systematic review of the cost-effectiveness of worksite physical activity and/or nutrition programs. *Scand J Work Environ Health* 2012;38(5):393-408.

3) L Viester, EALM Verhagen, KI Proper, **JM van Dongen**, PM Bongers, AJ van der Beek. VIP in construction: systematic development and evaluation of a multifaceted health programme aiming to improve physical activity levels and dietary patterns among construction workers. *BMC Public Health* 2012;12:89

4) **van Dongen JM**, van Poppel MN, Milder IE, van Oers HA, Brug J. Exploring the reach and program use of Hello World, an email-based health promotion program for pregnant women in the Netherlands. *BMC Research Notes* 2012;5:514

5) van Wier MF, **van Dongen JM**, van Tulder MW. Worksite physical activity and nutrition programs: beneficial to our health and wallet? *Ned Tijdschr Geneeskd* 2013;157(6):A4963 [In Dutch]

6) Miller PS, Biddle EA, **van Dongen JM**, van Tulder MW, Tompa E, Shemilt I. Economic incentives to enhance safety behaviours in workers for preventing occupational injuries (protocol). *Cochrane Database of Systematic Reviews* 2013;4.

7) **Van Dongen JM,** Strijk JE, Proper KI, van Wier MF, van Mechelen W, van Tulder MW, van der Beek AJ. A cost-effectiveness and return-on-investment analysis of a worksite vitality intervention among older hospital workers: results of a randomized controlled trial. *J Occup Environ Med* 2013;55(3):337-346.

8) **Van Dongen JM**, Tompa E, Clune E, Sarnocinska-Hart A, Bongers PM, van Tulder MW, van der Beek AJ, van Wier MF. Bridging the gap between the economic evaluation literature and daily practice in occupational health: a qualitative study among decision-makers in the healthcare sector. *Implement Sci* 2013;3;8:57

9) Oude Hengel KM, Bosmans JE, van **Dongen JM**, Bongers PM, van der Beek AJ, Blatter BM. Prevention program at construction worksites aimed at improving health and work ability is cost-saving to the employer: Results of an RCT. *Am J Ind Med* 2014;57(1);56-68

10) **van Dongen JM**, van Wier MF, Tompa E, Bongers PM, van der Beek AJ, van Tulder MW, Bosmans JE. Trial-based economic evaluations in occupational health: principles, methods, and recommendations. *J Occup Environ Med* 2014. Accepted for publication

11) **van Dongen JM**, van Berkel J, Boot CRL, Bosmans JE, Proper KI, Bongers PM, van der Beek AJ, van Tulder MW, van Wier MF. Cost-effective and financial return of a mindfulness-based worksite intervention aimed at improving work engagement: results of a randomized controlled trial. Submitted

12) Makkes S, **van Dongen JM**, Renders CM, van der Baan-Slootweg OH, Seidell J, Bosmans JE. Cost-effectiveness of two intensive inpatient treatments for severely obese children and adolescents: results of a randomized controlled trial. Submitted

13) **van Dongen JM**, Viester L, van Wier MF, Bosmans JE, Verhagen EALM, van Tulder MW, Bongers PM, van der Beek AJ. Cost-effectiveness and return-on-investment of a worksite intervention aimed at improving physical activity and nutrition among construction workers. To be submitted

14) **van Dongen JM**, Coffeng J, van Wier MF, Hendriksen IJM, Boot CRL, van Mechelen W, Bongers PM, van der Beek AJ, Bosmans JE, van Tulder MW. Cost-effectiveness and return-on-investment analysis of a combined social and physical environmental intervention in office employees. To be submitted

Book chapters

1) **van Dongen JM**, Proper KI, van Wier MF, van der Beek AJ, Bongers PM, van Mechelen W, van Tulder MW. Financiële opbrengsten van beweeg- en/of voedingsinterventies op de werkplek: een systematisch literatuuronderzoek. In: Hildebrandt VH, Bernaards CM, Stubbe JH. Trendrapport bewegen en gezondheid 2010/2011. Leiden: TNO, 2013.

DANKWOORD

Het einde is in zicht! Nu mijn proefschrift bijna klaar is en ik me, met gezonde spanning, kan gaan voorbereiden op mijn verdediging ben ik toe aan het schrijven van mijn dankwoord. Met heel veel plezier denk ik terug aan de afgelopen 4 jaar, waarin veel mensen mij op welke manier dan ook geholpen hebben met het afronden van dit proefschrift. Voordat ik iemand vergeet wil ik op deze plek daarom alvast iedereen heel erg bedanken!!

Marieke, Maurits, Allard en **Paulien**, bedankt voor het vertrouwen dat jullie in mij hebben gehad en alle wetenschappelijke kennis die ik, grotendeels door jullie, in de afgelopen jaren heb opgedaan. **Marieke**, in het begin moest ik misschien nog een beetje wennen aan de kritische blik die je doorgaans op mijn artikelen wierp, maar in de loop van mijn promotietraject ben ik je daardoor juist enorm gaan waarderen. Je oog voor detail hebben mijn artikelen namelijk een stuk beter gemaakt! Daarnaast heb ik je ervaren als een erg leuke, meegaande en gezellige copromotor, met als hoogtepunt denk ik toch wel ons bezoek aan, "*Wat ontzettend leuk dat we hier zijn"*, Sinaia en York! **Maurits**, eerlijkheid gebiedt te zeggen dat ik zonder jou als stagebegeleider mogelijk iets anders was gaan doen dan promoveren! Van het feit dat ik mede door jou ben gaan solliciteren op een onderzoeksfunctie heb ik nooit spijt gehad (wel als ik het niet had gedaan denk ik)! Heel erg bedankt voor daarvoor!! Tijdens mijn promotieonderzoek heb ik je ervaren als een hele fijne promotor, waarbij ik altijd even kon binnenlopen met een vraag en met wie ik altijd kon sparren over de richting van onze artikelen. Ik ben daarom ook heel blij dat we onze samenwerking in mijn huidige functie kunnen voortzetten!! **Allard**, ik heb het altijd erg leuk gevonden om jou als promotor te hebben. Misschien wel vooral doordat we het over de inhoud en/of de richting van de studies bij aanvang niet altijd eens waren. Hierdoor daagde je me uit om kritisch over mijn eigen ideeën en keuzes na te denken en deze altijd te onderbouwen met goede argumenten. Na leuke en leerzame discussies zijn we er vervolgens gelukkig altijd uitgekomen. Daarnaast heb ik erg veel respect voor het feit dat jij bijna altijd als eerste mijn stukken had gelezen, inclusief zorgvuldig commentaar. Dit ondanks je vele promovendi! **Paulien**, zonder jou had dit proefschrift er waarschijnlijk heel anders uitgezien. Tijdens mijn promotietraject heb je er namelijk constant voor gezorgd dat ik het perspectief van de werkgever niet uit het oog verloor. Dank daarvoor! Daarnaast waardeer ik het erg dat je, ondanks je drukke agenda, bij zoveel mogelijk overleggen aanwezig was!

Leden van de leescommissie, prof.dr. A Burdorf, prof.dr. W.B.F. Brouwer, prof.dr. S.M.A.A. Evers, prof.dr. C.T.J. Hulshof en prof.dr.ir. A.J. Schuit, hartelijk dank voor de tijd en energie die jullie hebben gestoken in het lezen en beoordelen van mijn proefschrift. Ik kijk uit naar jullie vragen tijdens mijn verdediging.

Graag wil ik ook alle andere co-auteurs van mijn artikelen bedanken; **Willem, Karin, Jorien, Jantien, Laura, Evert, Ingrid, Cécile, Judith, Emile, Laurie,** en **Anna.** Bedankt voor het vertrouwen in mijn analyses van de door jullie ontwikkelde interventies en jullie feedback op de artikelen. In het bijzonder wil ik hier **Karin** en **Judith** nog even iets uitgebreider noemen. **Karin**, ik denk dat voor iedere promovendus zijn of haar eerste artikel de grootste uitdaging is. Je hebt mij hier enorm mee geholpen en me af en toe echt door de, soms iets wat saaie, onderdelen van de systematische reviews heen gesleept. Ik heb het ontzettend leuk en leerzaam gevonden om met jou samen te werken! **Judith**, wat ben ik blij dat je met vier van mijn artikelen hebt meegeschreven en meegedacht. Met name de ontwikkeling van de STATA Do Files en het schrijven van het methodologische artikel heb ik super leuk en leerzaam gevonden. Hiermee heb je mijn proefschrift echt naar een hoger niveau weten te tillen! Daarnaast vind ik je een hele fijne collega en ben ik blij dat we in de toekomst nog veel vaker economische evaluaties kunnen schrijven en onderwijs kunnen geven. **Emile**, *thanks a lot for giving me the opportunity to visit you at "Institute for Work and Health" and for being able to analyze your qualitative data. I absolutely loved my stays in Toronto as well as working with you on both the qualitative study and the economic evaluation methods paper.* **Anna**, *thanks for all your help during the transcription process of the interviews as well as all your other work on the qualitative study.* **Laurie**, *thanks a lot for helping me out during the qualitative data analysis process. Without your help, I probably wouldn't have been able to pull it off!*

Promoveren doe je nooit alleen en dit geldt zeker als je deel uitmaakt van een groter (VIP) project. **Laura**, **Jantien**, en **Jennifer**, heel erg bedankt voor de fijne samenwerking aan de economische evaluaties en het feit dat ik altijd bij jullie binnen kon lopen met vragen! Daarnaast wil ik **Ernst**, als projectleider van het VIP project, heel hartelijk bedanken voor al zijn steun, interesse en begeleiding. Je hebt op een hele prettige manier de voortgang van de diverse projecten nauwlettend in de gaten gehouden!! Ook wil ik graag alle **deelnemers** en **overige (onderzoeks-)medewerkers** van de (VIP) projecten bedanken! Zonder jullie inzet had ik namelijk niks te analyseren gehad!

Speciale dank ben ik hier ook verschuldigd aan **Mia** en **Claudia**. Als er twee mensen zijn die mij hebben geleerd hoe ik een wetenschappelijk artikel moet schrijven dan zijn jullie het wel! Ik vind het heel bijzonder dat jullie zoveel tijd hebben weten vrij te maken om met jullie stagiaires achter de computer te zitten om te kijken hoe we ons stuk steeds een beetje beter konden maken! **Mireille, Ivon, Hans B,** en **Hans van O,** heel erg bedankt voor jullie begeleiding en fijne samenwerking tijdens mijn eerste werkzaamheden als onderzoeker.

Tijdens mijn promotietraject heb ik op veel verschillende plekken gezeten, wat als voordeel heeft gehad dat ik veel leuke collega's heb leren kennen. **Qaisar**, *thanks for being such a nice roomy during the start of my PhD!* Ook wil ik alle overige '(ex-) bewoners' van **T6** bedanken voor de gezellige tijd die ik daar heb gehad! **Teddy**, **Susan**, **Esther**, **Maartje**, **Wilma**, **Irma**, **Liesbeth** en **Marije**, bedankt voor alle gezelligheid op **T5** en ik ben blij dat ik dat ik nog steeds bij (een deel van) jullie mag zitten ☺! Naast mijn werkplek bij Gezondheidswetenschappen zat ik ook bij Sociale Geneeskunde. Wat een feest van herkenning, allemaal promovendi die onderzoek deden binnen de bedrijfsgezondszorg. Beste **(ex-)H/G0-ers**, bedankt voor alle gezelligheid, morele steun en gezellige etentjes! Dan wil ik nog een paar (oud)collega's in het bijzonder noemen. **France**, ik heb het altijd erg leuk en gezellig gevonden om met jou vanaf het eerste jaar van mijn promotietraject enkele gezondheidseconomie vakken te geven! Gelukkig kunnen we dit in de toekomst blijven doen. **Raymond**, bedankt voor het vertrouwen dat je in me had toen je me aannam voor mij huidige functie! Op het moment dat ik dit dankwoord schrijf zie ik hoe Niki Terpstra Parijs-Roubaix wint.

Hopelijk zorgt dit ervoor dat ik dit jaar wel een keer boven je eindig in de EMGO+ Tourpoule. **Karen OH**, wat was je een gezellige kamergenoot en wat was het leuk om je te mogen helpen met je economische evaluatie! **Karen B**, vanaf mijn eerste dag bij EMGO+ zat ik bij jou op de kamer!! Ik heb echt enorm leuke jaren gehad met jou als kamergenoot en daarnaast heb ik erg genoten van al onze etentjes en stapavondjes. Hopelijk gaan we dit in de toekomst veel vaker doen! **Caroline**, wat ben ik blij dat je naast collega ook zo'n goede vriendin bent! Jij hebt altijd een luisterend oor!! Ook vond ik onze vakantie naar Londen helemaal top, iets wat zeker voor herhaling vatbaar is!!! **Jennifer**, wat hebben wij een geweldige roadtrip gehad na ons congres in Los Angeles. Wat was ik trots toen je ons in het aarde donker over een bergpas naar Yosemite National Park hebt gereden!!! Daarnaast spijt het mij nog steeds dat ik dacht dat ik wel een leuke wandelroute door San Francisco wist zonder op de kaart te kijken, een route die uiteindelijk toch bijna 20 kilometer bleek te zijn ☺! **Jorientjuh**, wat fijn dat ik voor mijn eerste economische evaluatie jouw interventie mocht evalueren! Ik heb de samenwerking met jou als super prettig ervaren en ben daarom blij dat we dat nu nog steeds doen in diverse stageprojecten. Daarnaast ben ik heel blij met jou als vriendinnetje! Hopelijk gaan we snel weer eten, parkhangen, sporten en/of racefietsen!

Marieke en **Pam**, lieve paranimfen! Wat fijn dat jullie 13 juni naast mij staan. **Pammie**, vanaf de eerste dag van de introductieweek van Bewegingswetenschappen zijn wij vriendinnen. Tijdens onze studie trokken we zelfs zoveel met elkaar op dat docenten hun berichten aan mij via jou doorgaven! Ik ben hierdoor altijd met heel veel plezier naar de VU gegaan. Daarnaast kan ik nog steeds nagenieten van onze vakantie in Malta, alle feestjes en natuurlijk Thunderdome 2007 en 2012! Hopelijk gaan we dit in de toekomst nog vaak doen en natuurlijk veel pizza's eten ☺! **Marieke**, lief zusje! Wat ben ik ontzettend trots op alles wat je hebt bereikt in je leven!!! Ik weet zeker dat je voor je huidige studie ook met vlag en wimpel zult slagen! Daarnaast ben jij degene die me vaak rustig krijgt als ik een beetje "stressed out" ben en hopelijk lukt dit je vrijdag de 13^{e} ook ☺!

Er is natuurlijk ook een leven naast je promotie. **Anna**, **Hannah**, **Kirsten**, **Anne**, **Selma**, **Cindy**, **Yvonne B**, **Yvonne R**, **Ilja**, **Carlijn**, **Ivo**, **Vincent**, en **Rens** wat is het fijn om op te groeien met zo'n groep vrienden om je heen. Wat kijk ik met veel plezier terug op onze atletiekwedstijden, swalmenkampen, overige vakanties, feestjes, etentjes, kerstdiners, verjaardagen en bowlingwedstrijden, waarvan sommige ook voor de nodige afleiding van mijn proefschrift hebben gezorgd ☺! **Carlijn** en **Cindy**, poepiescheetjes, bedankt voor het feit dat jullie er altijd voor me zijn en jullie onafgebroken interesse in mijn proefschrift! Ik zie nu al uit naar de volgende film en frituuravond!!! **Linda**, **Caroline**, **Minke**, **Jolle** en **Susan**, bedankt voor de broodnodige gezellige afleiding van mijn/onze proefschrift(en)! Gelukkig zit ons volgende weekendje alweer planning! **Linnie** en **Caatjuh**, wat waren jullie geweldige huisgenoten! Door jullie heb ik een top studententijd gehad en ik ben blij dat we elkaar nog steeds vaak zien!! Bedankt voor alles lieffies! **Jolle**, als er iemand een top vriendin is ben jij het wel! Jij staat altijd voor mij klaar en hebt altijd een luisterend oor, al jaren!! Bedankt voor alles!! **Alex**, twee jaar hebben we bij elkaar in huis gewoond! Dat je via woningnet zo'n leuk vriendinnetje kunt vinden had ik nooit gedacht! Met name onze vakantie naar Marokko heb ik als hilarisch ervaren ☺! Kom je snel weer Indisch bij mij eten? **Marielle** en **Floor** wat fijn dat we na Bewegingswetenschappen altijd contact hebben gehouden! **Floor**, wat was het super om je op te mogen zoeken in Bosten en **Marielle** wat was het leuk in Bretagne. We spreken elkaar misschien iets minder door onze drukke agenda's, maar via WhatsApp houden we elkaar gelukkig op de hoogte! **Krista**, **Anouk** en **Marleen**, misschien wel mijn oudste vriendinnetjes ☺, bedankt voor alle leuke weekendjes die we hebben gehad!

Lieve **papa** en **mama**, ik had me geen betere en lievere ouders kunnen wensen! Bedankt voor het vertrouwen dat jullie altijd in mij hebben gehad en voor alle bemoedigende woorden als er even iets tegen zat! Daarnaast wil ik jullie heel erg bedanken voor het feit dat jullie mij hebben bijgebracht dat het niet uitmaakt wat je doet, als het je maar gelukkig maakt en je er maar alles uithaalt wat erin zit. Dat laatste heb ik met dit proefschrift ook geprobeerd!

Dimi, wat ben ik blij dat ik je ”gevonden” heb! Bedankt voor al je vertrouwen, liefde en steun, met name tijdens de laatste loodjes van dit proefschrift. En als ik het even niet zie zitten (bijvoorbeeld met het ontwerpen van de cover ☺), ben ik blij dat jij altijd een oplossing hebt! Lieffie, je maakt me heel erg gelukkig!

TNO

EMGO+
Institute for Health and Care Research

BODY@WORK