Scand J Work Environ Health 2010;36(4):273–288

Economic evaluations of occupational health interventions from a corporate perspective – a systematic review of methodological quality

by Kimi Uegaki, PhD,[1, 2] Martine C de Bruijne, MD, PhD,[2] Ludeke Lambeek, MSc,[2, 3] Johannes R Anema, MD, PhD,[2, 3, 4] Allard J van der Beek, PhD,[2, 3] Willem van Mechelen, MD, PhD,[2, 3] Maurits W van Tulder, PhD [1, 5]

Uegaki K, de Bruijne MC, Lambeek L, Anema JR, van der Beek AJ, van Mechelen W, van Tulder MW. Economic evaluations of occupational health interventions from a corporate perspective – a systematic review of methodological quality. *Scand J Work Environ Health.* 2010;36(4):273–288.

Objective Using a standardized quality criteria list, we appraised the methodological quality of economic evaluations of occupational safety and health (OSH) interventions conducted from a corporate perspective.

Methods The primary literature search was conducted in Medline and Embase. Supplemental searches were conducted in the Cochrane NHS Economic Evaluation Database, the National Institute for Occupational Safety and Health (NIOSH) database, the Ryerson International Labour, Occupational Safety and Health Index, scans of reference lists, and researchers' own literature database. Independently, two researchers selected articles based on title, keywords, and abstract, and if needed, fulltext. Disagreements were resolved by a consensus procedure. Articles were selected based on seven criteria addressing study population, type of intervention, comparative intervention, outcome, costs, language, and perspective. Two reviewers independently judged methodological quality using the Consensus on Health Economic Criteria (CHEC-list), a 19-item standardized quality criteria list. Disagreements in judgment were also resolved by consensus. Data were analyzed descriptively.

Results A total of 34 studies were included. Of these, only 44% of the studies met more than 50% of the quality criteria. Of the 19 quality criteria, 8 were met by 50% or more of the studies. The 11 least-fulfilled criteria related to: (i) performance of a sensitivity analysis, (ii) selection of perspective, (iii) description of study population, (iv) discussion of generalizability, (v) description of competing alternatives, (vi) presentation of the research question, (vii) measurement of outcomes, (viii) measurement of costs, (ix) valuation of costs, (x) declaration of researchers' independence, and (xi) discussion of ethical and distributional issues.

Conclusions Apart from a few exceptions, the overall methodological quality of the economic evaluations of OSH interventions from a corporate perspective was poor. As such, there is a risk of biased results. The quality of future evaluations needs to be improved to increase the validity of their conclusions and recommendations.

Key terms review; work; occupational safety and health; OSH; OSH interventions; economics.

There are two driving forces that have led to a growing interest in economic evaluations of occupational safety and health (OSH) interventions (1–3). First, health problems among the working population have a significant and far-reaching socioeconomic impact. Second, resources for OSH interventions are scarce, which necessitates that choices are made. In practice, occupational health professionals – along with workers, worker representatives and company managers – are required to make choices on a daily basis. Economic evaluations are systematic comparisons of two or more health technologies, services, or programs in terms of both costs and consequences. This simultaneous comparison of the costs and consequences provides insight into which intervention is worth doing over another. A societal perspective is traditionally recommended as the framework

1 Health Technology Assessment Unit, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands.
2 Department of Public and Occupational Health, EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands.
3 Body@Work, Research Center Physical Activity, Work and Health, TNO-VU University Medical Center, Amsterdam, the Netherlands.
4 Research Center for Insurance Medicine, AMC-UWV-VU University Medical Center, Amsterdam, the Netherlands.
5 Institute of Health Sciences, Faculty of Earth and Life Sciences, VU University, Amsterdam, the Netherlands.

Correspondence to: Professor MW van Tulder, EMGO Institute for Health and Care Research, VU University, De Boelelaan 1085, Room U-452, 1081 HV, Amsterdam, the Netherlands. [E-mail: maurits.van.tulder@falw.vu.nl]

for economic evaluations as it takes, in principle, all costs and consequences in account (4). For specific stakeholders, however, not all costs and consequences may be relevant, limiting the interpretability of results from a societal perspective for decisions at a local level. Therefore, to inform decisions at a local level, taking the perspective of a specific stakeholder, such as a company, may be warranted (5–7). With regards to OSH interventions, although success requires collaboration between occupational health professionals, workers, worker representatives, and company managers, the final decision about funding programs or services within companies rests with (top) management.

The methodological quality of an economic evaluation reflects the extent to which biased results are possible and consequently influences how useful the conclusions and recommendations will be for decision makers. In a review of economic evaluations of OSH interventions in the healthcare setting, Niven (8) found that methodological rigor was lacking in nearly all of the identified publications. Tompa et al (9) made a similar conclusion in their literature review of workplace-based interventions. The former review, however, was limited to evaluations performed in the healthcare setting, and the latter to those addressing musculoskeletal health problems. The methodological quality of economic evaluations of OSH interventions from a corporate perspective has not yet been systematically evaluated. In addition, the systematic appraisal of methodological quality has been hampered by a lack of a standardized criteria list (10). Recently, such a criteria list has been developed and published (11). Thus, the objective of our systematic review was to assess the methodological quality of economic evaluations of preventive interventions for workers from a corporate perspective using a standardized criteria list.

Methods

Study design and search strategy

We conducted a systematic review of economic evaluations of OSH interventions from a corporate perspective and based on primary data. We conducted our primary search in Embase.com, an online database that combines the recorded journal entries of Medline from 1966 to the present and Embase from 1974 to the present. The Embase.com search strategy was developed by an experienced search specialist at our institute and covered the period from 1966 to April 2007. The search strategy was structured as follows: (i) set 1 = (economic-evaluation/exp) AND (employee/exp OR employee* OR employer/exp OR employer* OR industrial-worker/exp OR worker* OR work-site OR worksite OR workman-compensation/de OR workplace/de OR workplace OR work-capacity/de); (ii) set 2 = set 1 AND [(cochrane review)/lim OR (controlled clinical trial)/lim OR (systematic review)/lim OR comparative-study/de OR intermethod-comparison/de OR clinical-study/exp OR controlled-study/exp]; (iii) set 3 = set 1 AND (health-program/exp OR health-maintenance-organization/exp OR occupational-health-service/exp OR preventive-health-service/exp); (iv) set 4 = set 1 AND (productivity/exp OR absenteeism/exp OR return-to-work OR sick-leave OR job-performance/de OR work-resumption/de); (v) set 5 = (Measur* OR valuing* OR valuation OR value* OR significance OR analys* OR estimat* OR assess* OR determinat* OR methodology/de OR accuracy/de); (vi) set 6 = (set 3 OR set 4) AND set 5; and (vii) final set of articles = set 2 OR set 6.

Supplemental searches were conducted in the Cochrane NHS Economic Evaluation Database (NHS EED), the National Institute for Occupational Safety and Health (NIOSHTIC-2) database, and the Ryerson International Labour, Occupational Safety and Health Index (RILOSH) database. Additional articles were identified from reference lists and the researchers' own literature databases.

Study selection

Two reviewers independently determined the eligibility of studies on the basis of title, keywords, and abstract. If uncertainty remained, the fulltext was reviewed. Differences in judgment were resolved through a consensus procedure, in which disagreements were openly discussed and a third reviewer was consulted if disagreements remained.

We selected studies based on the following seven criteria: (i) the study population consisted of working-age individuals; (ii) the intervention in question was a workplace or primary care service, technology, or program targeting workers; (iii) the intervention was compared to an alternative; (iv) an outcome that reflected a worker's health-related production capacity was measured and valued in monetary terms; (v) at a minimum, costs of intervention-related resource use were included; (vi) the study was reported in either Dutch or English; and (vii) the economic evaluation was conducted from a corporate perspective. We excluded: studies involving children, the elderly, unemployed or mixed populations; interventions in hospitals or in-patient settings; and editorials, letters, congress abstracts, reviews, and articles that only reported the design of an economic evaluation but not its results. Modeling studies were excluded because these studies have unique methodological issues compared to economic evaluations based on primary data.

Applied classification schemes

In order to manage the heterogeneous nature of the data and facilitate data analysis and summarization of the findings, three classification schemes were used. First, we categorized the interventions according to their aim with regard to prevention as follows: (i) primary – to decrease the risk for incurring or developing a health problem; (ii) secondary – to identify individuals at risk for a health problem through screening; (iii) tertiary – to prevent chronicity or limit the consequences of a diagnosed health problem (12).

Secondly, we used the following five labels for study design (13, 14): (i) randomized controlled trial; (ii) controlled before-and-after (ie, studies described as a non-randomized controlled trial or quasi-experimental controlled trial); (iii) uncontrolled before-and-after (ie, designs with a single group pre-test and post-test); (iv) case–control; and (v) historical cohort (ie, designs in which a retrospective analysis of differences between two alternatives was performed based on a review of records).

Thirdly, we labeled the economic evaluations using the conceptual matrix proposed by Drummond et al (4). The type of economic evaluation is determined by the number of alternatives compared, if both costs and consequences are included, and how the consequences are expressed. In a cost–effectiveness analysis (CEA), the consequences are expressed in terms of a unit of effect such as pain, function, or symptom severity. A cost–utility analysis (CUA) is a variant in which the unit of effect is quality-adjusted life years. In a cost–benefit analysis (CBA), all health consequences (ie, benefits due to improved health, future healthcare costs avoided, and increased productive output due to improved health status) are translated into a monetary value using principles of willingness-to-pay (4). In addition, we used the label "financial appraisal" to denote economic evaluations in which the costs and consequences of two or more alternatives are compared, but where the monetary consequences were limited to changes in healthcare use and/or productivity valued using market prices. It should be noted that the financial appraisal label is not found in the conceptual matrix of Drummond. This label was chosen to make a distinction between studies where valuations of health-related productivity are considered conventionally as costs rather than outcomes. When valuations of health-related productivity are considered conventionally as costs, the appropriate label – per Drummond's conceptual matrix – would be "cost analysis".

Methodological quality assessment

The methodological quality of the included studies was evaluated using the Consensus on Health Economic Criteria (CHEC-list), a 19-item assessment tool developed through a Delphi procedure involving 23 international experts in the field of health economics (11). The CHEC-list represents a minimum set of methodological criteria that address internal and external validity aspects of individual economic evaluation studies. Of the quality criteria, 18 relate to internal validity issues regarding study design, conduct, and analysis. Of these, five pertain to study design [eg, the description of study population; description of the interventions that are being compared; the research question; the research design; and the time horizon (ie, follow-up period) of the study]. Eleven criteria relate to the conduct of the economic evaluation [eg, selection and justification of the perspective for the analysis; identification, measurement, and valuation of resource use (ie, costs); identification, measurement, and valuation of outcomes (ie, consequences); adjustment for costs and outcomes occurring in the future (ie, discounting); congruency between the presented data and conclusions; independence of the investigators; and a discussion of ethical and distributional issues]. The remaining two criteria address the issue of analysis, that is, whether an incremental analysis of costs and outcomes and a sensitivity analysis was performed. The former refers to an analysis involving a joint comparison of the difference in costs and difference in outcomes between the two interventions and the latter to a testing of assumptions made in the main analysis. Finally, one criterion addresses external validity, namely, the generalizability of results (11).

The operationalizations of the criteria for the measurement and valuation of outcomes do not, in principle, pertain to health-related work productivity variables, such as sick leave or work presenteeism, for they are considered costs in an economic evaluation. However, given that in economic evaluations conducted from a corporate perspective, the outcome of OSH interventions is commonly expressed in these terms and translated into a monetary value (15), we included health-related work productivity variables in the outcomes judgment. Furthermore, given that almost no health outcomes were considered in the economic evaluations of the included studies, this adaptation offered the opportunity of a more specific evaluation of how health-related work productivity was measured and valued. Consequently, we expanded the existing operationalizations. That is, we judged the use of insurance or workers' compensation databases for measurement negatively because the information is limited to that of approved cases. To judge the valuation of sick leave positively, a clear report of the physical units of sick leave for each group, the cost price for each unit of sick leave, and the source of this cost price were required.

A pair of reviewers independently evaluated the methodology of each study. All reviewers piloted the use of the CHEC-list. For 33 of 34 studies, the first author paired

with one of three co-authors to evaluate the study with one exception (16), where the first author was a co-author of the given study. The reviewers disagreed on 131 of the 646 quality scores (21%) and used the same consensus procedure as in the study selection step to resolve the disagreements. Data from the quality appraisals were quantified per article and per item in terms of percentages of positive ratings. Trends in quality were examined over time periods and categories of health problems. We synthesized the findings descriptively.

Results

Study selection

Our primary search in Embase.com resulted in 1645 hits, and our searches in NHS EED, NIOSHTIC-2 and RILOSH resulted in 166, 477, and 352 hits, respectively. From this total of 2640 hits, 100 duplicates were removed, resulting in 2540 articles to be screened. Of these, 2422 were excluded based on title, keywords and abstract, and the full papers of the remaining 118 articles were assessed. Thirty studies were included based on our selection criteria. Reasons for exclusion were: wrong population or focus (N=477); not a comparative study (N=1777); no work-related outcome (N=176); no resource use-related costs (N=21); language (N=40); and economic evaluation conducted from a perspective other than a corporate one (N=19). Twenty-nine of the included studies were identified in Embase.com, and one in RILOSH. The searches in NHS EED and NIOSHTIC-2 did not result in any additional studies. With an additional four articles being identified ad hoc, a total of 34 studies were included in this review.

General description of the studies

Table 1 presents an overview of the selected studies according to health problem. The largest proportion of the studies (50%) focused on musculoskeletal disorders (16–29). Mental health problems, although a significant cause of work disability, were only addressed in two studies (30, 31). Table 2 provides a description of study populations, the intervention comparisons and

Table 1. Overview of included studies (N=34) per prevention category and according to targeted problem, year, and country.

Type of prevention	Health category	Targeted problem	Reference	Location
Primary prevention (N=22)	Musculoskeletal	Back injury	Shi, 1993 (27)	United States
		Back injury	Mitchell et al, 1994 (25)	United States
		Lacerations	Banco et al, 1997 (17)	United States
		Musculoskeletal injury	Melhorn et al, 1999 (24)	United States
		Musculoskeletal injury	Spiegel et al, 2002 (28)	Canada
		Musculoskeletal injury	Collins et al, 2004 (19)	United States
		Musculoskeletal complaints	Burdorf et al, 2005 (18)	The Netherlands
		Musculoskeletal injury	Engst et al, 2005 (20)	Canada
		Musculoskeletal injury	Nelson et al, 2006 (26)	United States
	Mental health	Stress	Smoot & Gonzales, 1995 (31)	United States
	General health	Health, lifestyle & wellness	Shephard, 1992 (45)	Canada
		Health, lifestyle & wellness	Blaze-Temple & Howat, 1997 (32)	Australia
		Health, lifestyle & wellness	Schultz et al, 2002 (38)	United States
		Health, lifestyle & wellness	Proper et al, 2004 (46)	The Netherlands
		Health, lifestyle & wellness	Aldanaet al, 2005 (42)	United States
		Influenza	Campbell & Rumley, 1997 (35)	United States
		Influenza	Thomson et al, 1999 (34)	Australia
		Influenza	Dille, 1999 (39)	United States
		Influenza	Cohen et al, 2003 (36)	Australia
		Influenza	Morales et al, 2004 (44)	Columbia
		Influenza	Columbo et al, 2006 (33)	Italy
		Influenza	Samad et al, 2006 (40)	Malaysia
Secondary prevention (N=3)	Musculoskeletal	Musculoskeletal injury	Alexander et al, 1977 (47)	USA
		Musculoskeletal injury	Littleton, 2003 (43)	USA
		Musculoskeletal injury	Franzblau et al, 2004 (37)	USA
Tertiary prevention (N=9)	Musculoskeletal	Back pain	Linton & Bradley, 1992 (23)	Sweden
		Back pain	Versloot et al, 1992 (29)	The Netherlands
		Back pain	Hlobil et al, 2007 (16)	The Netherlands
		Musculoskeletal injury	Hochanadel & Conrad, 1993 (21)	USA
		Musculoskeletal pain	Landstad et al, 2002 (22)	Sweden
	Mental health	Major depression	Lo Sasso et al, 2006 (30)	USA
	General health	Health, lifestyle & wellness	Skisak et al, 2006 (41)	USA
		Migraine	Legg et al, 1997 (48)	USA
		Migraine	Lofland et al, 2001 (49)	USA

outcomes. Working populations from diverse industry sectors were represented: healthcare (19, 20, 22, 23, 26, 28, 31–34), manufacturing (24, 35–38), energy (21, 39–41), education (42, 43), finance (44, 45), government (27, 46), transport (16, 29), construction (18), defense (25), retail (17), telecommunications (47), and multiple sectors (30, 48, 49). With the exception of influenza, no single intervention was consistently evaluated in the other subgroups of studies addressing the same health problem. Also, while all studies included an outcome reflecting health-related work productivity, few studies included a health effect [such as musculoskeletal pain, discomfort, or fatigue (18, 20, 23, 46)], function (23), or mental health complaints, such as anxiety, helplessness, or depression (23).

Methodological quality assessment

The main characteristics of the economic evaluations and the percentages of quality criteria fulfilled by each of the studies are presented in table 3. Only 15 of the 34 studies (44%) met more than 50% of the quality criteria (16, 18, 30, 33, 35–37, 39, 40, 44, 46, 49), and of these, 3 met greater than 75% of the criteria (30, 33, 36).

Six of the economic evaluations were based on data from randomized controlled trials (16, 17, 27, 30, 36, 46), 16 from a controlled before-and-after design (20, 22, 24, 29, 31–33, 35, 38, 40–45, 47), seven from a uncontrolled before-and-after design (18, 19, 23, 26, 28, 48, 49), four from case–controls (25, 34, 37, 39), and one from a historical cohort (21).

All the studies conducted a financial appraisal while two (46, 49) carried out additionally a CEA. No included study performed a CUA. With respect to expressing the efficiency of a given intervention compared to its alternative, 65% of the financial appraisals (16, 17, 21–23, 26, 27, 29–31, 33, 34, 36–40, 43, 44, 46–48) reported the difference between monetary benefits and program costs as net savings or benefits, 32% (21, 24, 28, 32, 33, 35, 38, 42, 45, 48, 49) provided a benefit-to-cost ratio, 21% (18, 19, 27, 30, 38, 41, 45) reported the return-on-investment (ROI), 9% (20, 22, 26) calculated a payback period, and 6% (26, 28) noted an internal rate of return. Note that the total percentage is greater than 100% because ten studies reported two expressions each (21, 22, 26–28, 30, 33, 38, 45, 48).Twenty-eight of the studies reported cost savings or monetary benefits in favor of the intervention (16, 17, 19–22, 24, 26–31, 33–36, 38–46, 48, 49); three reported negative savings (25, 37, 47); two reported both negative and positive monetary benefits (18, 23); and one reported both a cost-neutral and positive situation (32). With the exception of three studies, none of the studies conducted a statistical analysis of the differences in costs or monetary benefits, or the joint cost–effect estimate [Hlobil et al (16) and Proper et al (46) applied a non-parametric bootstrapping technique, and Landstad et al (22) integrated a regression model into their analysis of subgroups.]

Table 4 presents the percentage of studies meeting each quality criterion. Of these, 8 criteria were met by 50% or more of the studies: (i) appropriate economic study design (item 4, 97%); (ii) identification of important and relevant outcomes measures (item 10, 85%); (iii) performance of an incremental analysis of costs and outcomes of alternatives performed (item 13, 79%); (iv) identification of important and relevant costs (item 7, 74%); (v) selection of an appropriate time horizon (item 5, 62%); (vi) valuation of outcomes and congruency between conclusions and reported data (items 12 and 16, 53%); and (viii) discounting of all future costs and outcomes (items 14, 50%).

The ten most prevalent methodological shortcomings were: (i) performance of a sensitivity analysis and discussion of the generalizability of findings (items 15 and 17, both 41%); (ii) selection of an appropriate perspective (item 6, 38%); (iii) clear description of the study population (item 1, 35%); (iv) clear description of competing alternatives (item 2, 32%); (v) presentation of a well-defined research question in answerable form (item 3, 24%); (vi) appropriate measurement of outcomes (item 11, 26%); (vii) appropriate valuation of costs (item 9, 18%); (viii) appropriate measurement of costs (item 8, 15%); (ix) discussion of ethical and distributional issues (item 19, 12%); and (x) declaration of funding information and absence of conflict of interest (item 18, 9%).

With regards to these shortcomings, the main reason for a negative rating for the sensitivity analysis criterion (item 15) was that potential cost drivers were simply not tested (17, 19–22, 24–27, 29, 31, 32, 34, 38, 39, 41–43, 45). The generalizability criterion (item 17) was rated negatively for a similar reason (17, 19–26, 28, 31, 34, 36, 38, 42, 43, 45–48). The lack of an explicitly stated justification in the published study for conducting the economic evaluation from a perspective narrower than the societal perspective and lack of a clear description of "usual care" (or the comparison situation) were the main reasons for a negative score on items 6 (17, 19–23, 25, 26, 29, 31, 32, 34, 38, 39, 41–47) and 2 (18, 19, 21–29, 31, 32, 38–42, 45, 47–49), respectively. Studies were given a negative rating for research question criteria because neither the alternatives nor the studied population (or both) were clearly described in the research questions (20–30, 32, 34, 35, 38, 40–42, 44–46, 49). The main reason for a negative score for appropriate outcome measurement was that data had been extracted either from a workers' compensation or corporate database (16, 19–22, 24, 26, 28, 29, 34, 37, 38, 40–43), or there had been high risk for recall bias (25, 39, 47–49). The criterion relating to the appropriate valuation of

Table 2. Overview of the study population, intervention comparisons, and outcomes of each included study (N=34) per targeted health problem. [ADL=activities of daily living; CTS=carpal tunnel syndrome; ILI=influenza-like illness]

References	Type of prevention	Sector or industry	Study population	Intervention	Comparison	Outcomes
Back injury & pain						
Mitchell et al, 1994 (25)	Primary	Defense	Airforce base warehouse employees	Company policy for back belt use: back belt use mandated for particular job tasks and history of back injury; back injury prevention training for all new hires; annual instruction period for proper lifting techniques; back belt issue paired with two 30-minute instruction session.	Own choice not to use back belt.	Lost work time; limited work days; rate of back injuries per 1000 workers; rate of lost time injuries per 1000 workers; limited work days per 1000 workers.
Shi, 1993 (27)	Primary	Public	County government employees	Back injury prevention program consisting of education, training, physical fitness activities and ergonomic improvement with, before-and-after feedback from health risk assessment.	No program or health risk assessment feedback.	Back pain prevalence; participant satisfaction; risk reduction; workers' compensation claims; medical claims; lost work time.
Hlobil et al, 2007 (16)	Tertiary	Transport	Airline workers	Routine guidance by occupational health physician plus graded activity program supervised by specially trained physiotherapists consisting of two 60-minute treatment sessions/week for a maximum of three months.	Routine guidance by occupational health physician and any other type of care except graded activity.	Sick leave; number of work-disabled workers.
Linton & Bradley, 1992 (23)	Tertiary	Health-care	Female licensed practical nurses	5-week chronic back pain prevention program involving 4-hour physical conditioning per day, ergonomic education, cognitive behavioral intervention, home exercise program.	Information about the study and that promise to be offered the treatment after completion of the study.	Pain intensity; fatigue; anxiety; sleep; ADL; depression; helplessness, medication use; sick leave.
Versloot et al, 1992 (29)	Tertiary	Transport	Bus drivers	Back school program consisting of 3 sessions with 6-month intervals. Session 1: 3 hours covering motivation, personal responsibility, mind-body interactions in relation to illness; stress, coping strategies & relaxation training; body mechanics for sports, working posture and seat adjustment. Sessions 2 & 3: covering experiences of past 6 months, structure & function of the back in relation to symptoms, stress & coping strategies.	Usual care.	Sick leave; quality and perceived effect of program.
Musculoskeletal complaints, injuries & lacerations						
Banco et al, 1997 (17)	Primary	Retail	Grocery store workers	(i) new cutters with a built-in safety guard + 15-minute training on proper use; (ii) usual cutters used + 15-minute training session on proper use.	Usual cutters and no training.	Injury rates per 100 000 man hours; lost work time; workers' compensation claims; medical costs; satisfaction with training; employee preference
Burdorf et al, 2005 (18)	Primary	Construction	Street makers and cement floor layers	Ergonomic equipment (hydraulic clamp/vacuum unit for street making; automated pump or silo/trunk with pump for laying cement flooring).	Street making and laying of cement flooring without new equipment	Physical load; musculoskeletal complaints; sick leave; work performance; overall productivity
Collins et al, 2004 (19)	Primary	Health-care	Nurses & nursing staff of nursing homes	"Best practice" program consisting of mechanical lifts, training to use lifts, zero-lift policy, and a pre-existing medical management program.	Pre-existing medical management program	Lost work time; restricted days; injury rate from handling tasks; injury rate from non-handling tasks; medical costs and disability costs
Engst et al, 2005 (20)	Primary	Health-care	Nurses & nursing staff of an extended care unit of a community hospital	Installation of ceiling lifts with tracking directly into all washrooms plus 1-hour training session by on-site occupational therapist and other general health and safety education.	Usual work environment with mechanical floor lifts, one sit-stander; other general health & safety education and training (excluding specific ceiling lift education & training)	Number of injuries; worker/staff satisfaction; staff perceptions & preferences; physical discomfort; work organization; medical costs; lost work time costs
Melhorn et al, 1999 (24)	Primary	Manufacturing	Sheet metal mechanics of an aircraft manufacturing company	5-step disability management program consisting of individual risk assessment, transitional work period plus usual occupational medicine clinic care.	Usual occupational medicine clinic care	Recordable case incidence rate per 200 000 hours worked; lost time case incidence rate; lost time day severity incidence rate; airplane production; estimated workers' compensation costs

(continued)

Table 2. Continued

References	Type of prevention	Sector or industry	Study population	Intervention	Comparison	Outcomes
Nelson et al, 2006 (26)	Primary	Healthcare	Nurses & nursing staff in high risk nursing home and spinal cord units	Multifaceted back injury prevention program consisting of an ergonomic assessment protocol; patient handling assessment criteria; peer safety leaders; patient handling equipment; after-action review process; and a no lift policy.	No multifaceted program; existing training programs and some equipment to assist with lifting	Injury rate; modified work days; lost work days; job satisfaction; unsafe patient handling acts; staff perceptions; medical costs; workers' compensation costs; lost productivity costs
Spiegel et al, 2002 (28)	Primary	Healthcare	Healthcare workers in an extended care hospital	Installation of a ceiling lift system, adoption of a "no-lift" policy, training and coordination.	No ceiling lift, no "no-lift" policy; presence of other lifting devices not specified	Injury-related time loss; compensation claims
Alexander et al, 1977 (47)	Secondary	Telecom	Prospective hires for telephone company	"No" pre-placement medical examination: examination conducted, however, all results reported as "no risk for work performance or attendance".	Pre-placement medical examination reported as they actually were: A=no risk; B=work restrictions but no risk with appropriate placement; R=risk identified	Lost work time; work performance; number of sick leave periods; number of accidents; number of work force losses; supervisor ratings of employee's overall job performance, recommendation as hire, and job match
Franzblau et al, 2004 (37)	Secondary	Manufacturing	Prospective hires for auto manufacturing company	Pre-placement CTS nerve testing and employment offers still honored despite abnormal test result.	Pre-placement CTS nerve testing and employment offers retracted if abnormal test result	Number of closed workers' compensation claims; CTS claim rate/1000 person-years; medical, lost work time, light-duty time costs
Littleton, 2003 (43)	Secondary	Education	Prospective hires for a university physical plant	Post-offer screening program: persons hired based on results of a post-offer functional capacity screening test based on 5–7 essential tasks; offers withdrawn if failure on any of the 5–7 screening tasks.	No post-offer screening program.	Number of workers' compensation cases
Hochanadel & Conrad, 1993 (21)	Tertiary	Energy	Energy research & manufacturing company employees	On-site physiotherapy clinic offering services related to evaluation, treatment, workstation evaluation, modalities, joint/soft tissue mobilization, progressive resistive exercises, and back school during work hours.	No on-site physiotherapy clinic.	Sick leave rate
Landstad et al, 2002 (22)	Tertiary	Healthcare	Hospital cleaning staff	Action program for prevention & rehabilitation consisting of group, leader, supervisor development; working environment program; suggestions forum; training for better work methods; education on physical and mental development, life skills, stress-management and well-being; rehabilitation.	Customary personnel support according to Swedish legislation.	Sick leave
Mental health problems						
Smoot & Gonzales, 1995 (31)	Primary	Healthcare	Healthcare workers of a psychiatric hospital	Communication training program consisting of 32 hour program in sequential 8-hour sessions; held once per week for 4 weeks at an off-unit training site.	No training.	Sick leave; staff turnover; number of patients' rights complaints; number of episodes of patient restraint & seclusion; number of assaults by patients on staff; staff satisfaction
Lo Sasso et al, 2006 (30)	Tertiary	Diverse sectors	Heterogeneous working population	Enhanced care practice training for physicians and nurses involving 4, 90-minute conference calls over a 2-month period; 8-hour training session for nurse case manager; structured case manager-physician communication.	Usual care; no training for physicians or nurse; no regular case manager contacts.	Effectiveness at work; sick leave
Health, lifestyle & wellness						
Aldana et al, 2006 (42)	Primary	Education	Teachers and support staff of a school district	Wellness program addressing education, weight loss, nutrition, fitness, safety belt & exercise.	No wellness program.	Sick leave; medical costs
Blaze-Temple & Howat, 1997 (32)	Primary	Healthcare	Hospital staff	(i) Counseling via offered employee-assistance program; (ii) self-arranged counseling.	No counseling.	Sick leave; compensated lost work time; turnover costs

(continued)

Table 2. Continued

References	Type of prevention	Sector or industry	Study population	Intervention	Comparison	Outcomes
Proper et al, 2004 (46)	Primary	Public	Civil servants of city government	9 months physical activity counseling involving 7 20-minute sessions to promote physical activity and health dietary habits + general information about lifestyle factors	No counseling; information about lifestyle factors.	Sick leave; physical activity; fitness; musculoskeletal symptoms; body composition; blood pressure; total blood cholesterol
Schultz et al, 2002 (38)	Primary	Manufacturing	Male employees in a manufacturing plant	Health promotion program consisting of annual health risk assessment for 13 health risk factors; on-site health screening; on-site & telephone wellness programs; medical vouchers; telephone nurse counseling line; newsletter.	No use of offered program; knowledge that program exists; newsletter; active health & safety programs within company.	Short-term and long-term lost work time
Shephard, 1992 (45)	Primary	Financial	Insurance company office workers	Worksite fitness center.	No fitness center.	Sick leave, recruitment, employee turnover, productivity, health benefits
Skisak et al, 2006 (41)	Tertiary	Energy	Petrochemical company employees	In-house non-occupational disability management program; training offered to occupational health nurses, corporate case managers, employees, supervisors, timekeepers and human resource managers to improve processes.	No formal internal or external disability management program.	Lost work time, worker satisfaction
Influenza						
Campbell & Rumley, 1997 (35)	Primary	Manufacturing	Textile plant workers	Vaccine Fluzone, manufactured by Connaught laboratories, Inc (North York, Ontario, Canada), 1992–1993 formula, composed of A/Texas/36/91 (H1N1), A/Beijing/353/89 (H3N2) and B/Panama/45/90.	No vaccine	Incidence rate of ILI; sick leave; incidence rate of vaccine side-effects
Cohen et al, 2003 (36)	Primary	Manufacturing	Steelwork employees	Vaxigrip, a purified, inactivated, split virion vaccine. Each 0.5 ml contained 15 micrograms haemagglutinin antigens of A/New Caledonia/20/99 (H1N1)-like; A/Sydney/5/97 (H3N2)-like; and B/Beijing/184/93-like.	No vaccine	Sick leave; attack rate of ILI, ILI episodes; ILI-related sick leave; febrile illness episodes; febrile illness -related sick leave; sick leave rate; medical visits
Colombo et al, 2006 (33)	Primary	Healthcare	District health authority employees	Split vaccine, Aventis Pasteur, source ULSS number 17	No vaccine	Sick leave
Dille, 1999 (39)	Primary	Energy	Nuclear plant employees	"1994–1995 influenza vaccine"	No vaccine	Episodes of ILI; ILI-related sick leave; ILI-complications; medical care
Morales et al, 2004 (44)	Primary	Financial	Bank employees	Inactivated split influenza vaccine	No vaccine	Sick leave; decreased productivity due to ILI episodes; ILI frequency
Samad et al, 2006 (40)	Primary	Energy	Petrochemical company employees	Inactivated, split influenza vaccine (VAXIGRIP, Aventis Pasteur)	No vaccine	Adverse events from vaccine; ILI episodes; sick leave; decreased productivity; percentage of effectiveness in ILI rate reduction; percentage of effectiveness in sick leave rate reduction; rates of ILI; rates of sick leave
Thomson et al, 1999 (34)	Primary	Healthcare	Hospital staff	Vaccine not specified	No vaccine	Sick leave; sick leave rate; incidence of separate sick leave periods
Migraine						
Legg et al, 1997 (48)	Tertiary	Diverse sectors	Heterogeneous working population	Injectable sumatriptan	Usual medication	Lost work time; work performance; lost leisure time
Lofland et al, 2001 (49)	Tertiary	Diverse sectors	Heterogeneous working population	Injection or tablet sumatriptan therapy	Nontriptan therapy	Total disability time; lost work time; lost leisure time

Table 3. Summary of main results of the economic evaluation and quality assessment of the evaluated interventions of each study per targeted health problem. [RCT=randomized controlled trial; NT=not tested statistically; ROI=return-on-investment; EAP=employee assistance program, ICER=incremental cost–effectiveness ratio; 95% CI=95% confidence interval; CAD=Canadian $; AUD=Australian $]

Reference	Study design	Economic evaluation	Authors´ main economic results	Quality criteria fulfilled (%)
Back injury & pain				
Mitchell et al, 1994 (25)	Case control	Financial appraisal	Rate of lost time injury decreased for those injured with belt in place versus without, however limited activity days higher in those injured with belt (NT); costs per injured worker with belt in place consistently higher per category of treatment options, expect physical therapy (NT); total costs per injured worker higher in belt group than no belt group (NT).	26
Shi, 1993 (27)	RCT	Financial appraisal	Total savings from medical costs and sick days was US$251 108 (NT); net benefit of introducing the back injury prevention program was US$161 108 (NT); ROI=179% (NT)	42
Hlobil et al, 2007 (16)	RCT	Financial appraisal	Mean difference in health care costs was €83 (95% CI -467–251); mean difference in net productivity loss costs was €999 (95% CI -1073–3115); cumulative savings over 3-year follow-up was €1661 (95% CI -4154–6913); reference year: 1999	74
Linton & Bradley, 1992 (23)	Uncontrolled before and after	Financial appraisal	No cost savings based on actual days (NT); positive cost savings based on estimated days assuming an increasing trend without program (NT)	42
Versloot et al, 1992 (29)	Controlled before and after	Financial appraisal	Incremental net savings was US$103 400 (NT)	16
Musculoskeletal complaints, injuries & lacerations				
Banco et al, 1997 (17)	RCT	Financial appraisal	Total net savings for Group A stores was US$245 per 100 000 man-hours/store and total net savings per year was US$29 413 (NT); compensation-related cases were virtually eliminated in the Group A stores; total net savings for group B stores was US$106 per 100 000 man-hours/store and total net savings per year was US$12 773 (NT)	32
Burdorf et al, 2005 (18)	Uncontrolled before and after	Financial appraisal	Hydraulic clamp or vacuum unit reduced production costs per cubic meter road by 4–9%, ROI was 0.8–4.8 years, benefits of less sick leave and work performance reduction were €5/day per worker (NT); automated pump or silo/truck with pump increased production costs per cubic meter by 3–10% and benefits of less sick leave and work performance reduction was €6.7/day per worker (NT)	63
Collins et al, 2004 (19)	Uncontrolled before and after	Financial appraisal	Initial program investment recovered by reduction in workers' compensation expenses in slightly less than 3 years (NT); ROI is shorter if the savings in indirect costs (eg, lost wages, costs of hiring & retraining) are considered	26
Engst et al, 2005 (20)	Controlled before and after	Financial appraisal	Payback period was 9.6 years when including all handling claims and 6.5 year when only including lifting/transferring claims (NT)	21
Melhorn et al, 1999 (24)	Controlled before and after	Financial appraisal	Direct savings over 4-year period was over US$5 million (NT); average benefit-to-cost ratio over 4 year period was 16.5:1; range was US$6–26 (NT)	32
Nelson et al, 2006 (26)	Uncontrolled before and after	Financial appraisal	Per year, cost savings of US$204 599 (NT); pay-back period was 3.75 years (NT); rate of ROI was 19% (NT)	26
Spiegel et al, 2002 (28)	Uncontrolled before and after	Financial appraisal	Direct savings alone produced a payback within 4 years and the payback occurs more quickly when the effect of indirect savings or the trend to rising compensation costs is considered (NT); benefits exceed the costs by a factor of >6:1, representing an internal rate of return of 17.9% (NT)	42
Alexander et al 1977 (47)	Controlled before and after	Financial appraisal	To have 1000 trained workers at the end of 1 year, the costs are US$25 000 greater with pre-placement medical evaluation (NT)	47
Franzblau et al, 2004 (37)	Case-control	Financial appraisal	Net lost of US$50 428–357 353 incurred by following a policy of retracting employment offer based on abnormal test results (NT); reference year: 2003	58
Littleton, 2003 (43)	Controlled before and after	Financial appraisal	Cost savings were over US$18 per dollar spent on the program (NT)	37
Hochanadel & Conrad, 1993 (21)	Historical cohort	Financial appraisal	Net savings is US$8.3 million or approximately $830 000 each year (NT); benefit-to-cost ratio for the life of the program >9:1 (NT)	11
Landstad et al, 2002 (22)	Controlled before and after	Financial appraisal	No difference sickness absence costs (not significant); net effect of the intervention for total group was €283.20 and the payback period would be under 4 years (NT); for the young subgroup, the intervention was associated with a net effect of €605.60 (not significant) and a payback period for subgroup would be 1.8 years (not significant)	42

(continued)

Table 3. Continued.

Reference	Study design	Economic evaluation	Authors´ main economic results	Quality criteria fulfilled (%)
Mental health problems				
Smoot & Gonzales, 1995 (31)	Controlled before and after	Financial appraisal	Decrease in staff resignations and transfers, sick leave, annual, patient rights' complaints, incidents of restraint & seclusion, assaults on staff (NT); Substantial savings for experimental unit compared to increased expenditures for control unit (NT)	42
Lo Sasso et al, 2006 (30)	RCT	Financial appraisal	Enhanced depression treatment results in an average net benefit to the employer of US$30 per worker in Year 1, US$257 per worker in Year 2 (NT), and the ROI over the 2-year period was 302%; ROI was 406% for 1.26 multiplier job function, 466% for 1.41 multiplier job function, and 675% for 1.93 multiplier job function; ROI under assumptions of sensitivity analyses was 20–132% (NT); reference year: 2000	84
Health, lifestyle & wellness				
Aldana et al, 2005 (42)	Controlled before and after	Financial appraisal	Benefit-to-cost ratio was 15.6:1 (NT); cost savings was $15.60 for every dollar spent on programming; reference year: 2002	37
Blaze-Temple & Howat, 1997 (32)	Controlled before and after	Financial appraisal	EAP program is cost-neutral compared to no counseling; the benefit-cost ratio was 1:1.0 (NT); EAP is not as cost-effective or cost-beneficial as self-arranged counseling; reference years: 1989 & 1990	47
Proper et al, 2004 (46)	RCT	Financial appraisal & cost–effectiveness analysis	No differences in total costs or sick leave costs (not significant); net benefits were €635 (95% C-1885–814); ICER for energy expenditure was €5.20 (95% CI -4.9–27.4) per extra kilocalorie per day per employee; ICER for fitness was €235 (95% CI -10–830) per beat per minute decrease in sub-maximal HR; counseling neither more costly nor more effective for the public health recommendation for physical activity; counseling tending to more costly, more effective for upper-extremity symptoms (not significant)	53
Schultz et al, 2002 (38)	Controlled before and after	Financial appraisal	Savings in disability days associated with participation was US$623 040/year (NT); savings to cost ratio was 2.3, or an annual return of US$2.3 for each dollar spent on program costs (NT)	16
Shephard, 1992 (45)	Controlled before and after	Financial appraisal	Return-on-investment: CAD$6.85 return for each dollar invested (NT); for sensitivity analysis, results expressed as a ratio: CAD$4.80 : $1.00 (NT)	16
Skisak et al, 2006 (41)	Controlled before and after	Financial appraisal	ROI greater than 4:1 based on direct expenditure and savings (NT)	37
Influenza				
Campbell & Rumley, 1997 (35)	Controlled before and after	Financial appraisal	Cost per saved lost workday was US$22.36 for a company savings of US$2.58 per dollar invested in the vaccination program	53
Cohen et al, 2003 (36)	RCT	Financial appraisal	Cost savings per vaccinated employee from the reduction in influenza-like illness -related absenteeism was AUD$20.93 and the global cost savings were AUD$5,652.31 (NT); cost savings per vaccinated employee from the reduction in febrile illness-related absenteeism was AUD$58.36 and global cost savings were AUD$15 406.74; reference year: 2000	84
Colombo et al, 2006 (33)	Controlled before and after	Financial appraisal	Benefit-to-cost ratio of the vaccination program was €4.20, meaning €1 invested returns €4 saved through less absenteeism from work (NT); savings were about €55 per employee for a total of €5900 (NT)	95
Dille, 1999 (39)	Case-control	Financial appraisal	Direct savings in potential health care costs avoided estimated as US$45.72/person vaccinated (NT); indirect savings estimated as US$38.12/person vaccinated (NT); combined cost savings = US$83.84/person vaccinated (NT); reference year: 1994	58
Morales et al, 2004 (44)	Controlled before and after	Financial appraisal	Net benefits/savings were US$6.40–25.98 per vaccinated employee based on labor costs alone (NT); net benefits/savings were US$89.30–237.80 when operating income also considered (NT); vaccination program cost saving for vaccination coverage above 20% and influenza-like illness rates above 10%	68
Samad et al, 2004 (40)	Controlled before and after	Financial appraisal	Employer savings were US$53 per vaccinated employee when labor costs only were considered (NT); savings increased to US$899.70 when operating income of each employee was also considered (NT)	58
Thomson et al, 1999 (34)	Case-control	Financial appraisal	Net benefit between AUD$1.55–5.80 depending on the average staff level (NT)	32
Migraine				
Legg et al, 1997 (48)	Uncontrolled before and after	Financial appraisal	Benefit-to-cost ratio was US$435 to $43.78 or 10:1 and the net benefit was US$435 – US$43.78 = US$391 (NT)	42
Lofland et al, 2001 (49)	Uncontrolled before and after	Financial appraisal & cost–effectiveness analysis	Benefit-to-cost ratio was 2.4:1.0, and ICER for company's perspective was US$59/disability day averted (NT)	53

Table 4. Summary of the percentage of studies fulfilling each Consensus on Health Economic Criteria (CHEC-list) quality criterion along with percentage of studies fulfilling each criterion over time.

CHEC-list quality criteria	Percentage of total articles (N=34) (%)	Percentage per time period			
		<1995 (N=7) (%)	1995–1999 (N=8) (%)	2000–2004 (N=10) (%)	2005–2007 (N=9) (%)
1. Is the study population clearly described?	35	14	13	50	56
2. Are competing alternatives clearly described?	32	14	38	20	56
3. Is a well-defined research question posed in answerable form?	24	0	38	30	22
4. Is the economic study design appropriate to the stated objective?	97	86	100	100	100
5. Is the chosen time horizon appropriate in order to include relevant costs and consequences?	62	57	50	70	67
6. Is the actual perspective chosen appropriate?	38	14	38	40	56
7. Are all important and relevant costs for each alternative identified?	74	57	75	80	78
8. Are all costs measured appropriately in physical units?	15	0	13	20	22
9. Are costs valued appropriately?	18	0	25	10	33
10. Are all important and relevant outcomes for each alternative identified?	85	57	88	90	100
11. Are all outcomes measured appropriately?	26	14	25	30	33
12. Are outcomes valued appropriately?	53	29	63	50	67
13. Is an incremental analysis of costs and outcomes of alternatives performed?	79	57	75	90	89
14. Are all future costs and outcomes discounted appropriately?	50	14	75	60	44
15. Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	41	14	25	60	56
16. Do the conclusions follow from the data reported?	53	71	13	60	67
17. Does the study discuss the generalizability of the results to other settings and patient/client groups?	41	29	38	30	67
18. Does the article indicate that there is no potential conflict of interest of study researcher(s) and funder(s)?	9	0	0	10	22
19. Are ethical and distributional issues discussed appropriately?	12	14	13	10	11

costs was scored negatively if the cost prices or the reference year was not clearly stated or if the main costs were calculated using tariffs rather than data on actual resources consumed (17–29, 31, 34, 35, 38, 40–49). A negative rating was given for appropriate measurement of costs if the physical units or the data collection method was not clearly stated, or if the validity of the instrument used to collect cost data was questionable (16, 17, 19–29, 32, 34, 35, 38–48). In most cases, the criterion related to the discussion of ethical and distributional issues was rated negatively because neither issue or just one of the two was addressed (16–25, 27–29, 34–36, 38, 40, 42, 45, 46, 49) and (26, 31, 33, 37, 39, 41, 44, 48), respectively. Finally, a negative rating for the criterion related to conflict of interest was due to either lack of a declaration of (no) competing interests (17, 20, 22–24, 26, 30, 40, 48, 49) or a lack of both a declaration and information about funding (18, 19, 21, 25, 27–29, 31, 32, 34–36, 38, 39, 41–43, 45–47).

A comparison of the methodological quality revealed a trend of a higher proportion of the studies meeting the quality criteria over time (table 4). While there are insufficient numbers of studies for each subgroup of health problems, a comparison of subgroups consisting of five or more studies suggested that the overall methodological quality of economic evaluations of influenza vaccines was higher than that of other health problems (table 5).

Discussion

Thirty-four studies were included in this systematic review of the methodological quality of economic evaluations from a corporate perspective. While a positive trend over time was observed, less than half of the studies met more than 50% of methodological quality criteria, and only three studies met more than 75% of the criteria. In the following, we discuss the implication of poor methodological quality, strategies to improve the quality of future economic evaluations, the strengths and limitations of our review, and additional considerations.

Implication of poor methodological quality

Measures of quality are strongly associated with aspects of study design and conduct that may be potential sources of selection, performance, attrition, detection, reporting, or other bias (50). A "no" coding for the following CHEC-list criteria may have an impact on the "other" category of risk of bias (i) item 5: appropriate time horizon; (ii) items 8 and 9: all costs measured and valued appropriately; (iii) items 11 and 12: all outcomes measured and valued appropriately; (iv) and item 14: appropriate discounting. Empirical evidence of the association between the estimated magnitudes of effect and a "no" coding on the aforementioned criteria are currently lacking. However, there is consensus that these items are

Table 5. Summary of the percentage of studies fulfilling each Consensus on Health Economic Criteria (CHEC) list quality criterion per targeted health problem. [MSK = musculoskeletal complaints, injury & lacerations]

CHEC-list quality criteria	Reference (N=34) (%)	Percentage per health problem			
		Back pain & injury (N=5)	MSK (N=12)	Life-style & wellness (N=6)	Influ-enza (N=7)
1. Is the study population clearly described?	35	40	17	33	57
2. Are competing alternatives clearly described?	32	40	25	0	71
3. Is a well-defined research question posed in answerable form?	24	20	25	0	43
4. Is the economic study design appropriate to the stated objective?	97	100	92	100	100
5. Is the chosen time horizon appropriate in order to include relevant costs and consequences?	62	60	58	67	86
6. Is the actual perspective chosen appropriate?	38	40	33	0	57
7. Are all important and relevant costs for each alternative identified?	74	80	67	50	86
8. Are all costs measured appropriately in physical units?	15	0	0	0	29
9. Are costs valued appropriately?	18	20	0	17	43
10. Are all important and relevant outcomes for each alternative identified?	85	80	92	83	86
11. Are all outcomes measured appropriately?	26	20	8	17	57
12. Are outcomes valued appropriately?	53	20	25	67	86
13. Is an incremental analysis of costs and outcomes of alternatives performed?	79	60	83	67	100
14. Are all future costs and outcomes discounted appropriately?	50	20	33	33	100
15. Are all important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?	41	20	33	17	71
16. Do the conclusions follow from the data reported?	53	60	67	50	43
17. Does the study discuss the generalizability of the results to other settings and patient/client groups?	41	60	17	33	71
18. Does the article indicate there is no potential conflict of interest of study researcher(s) and funder(s)?	9	20	0	0	29
19. Are ethical and distributional issues discussed appropriately?	12	0	17	17	0

important. Most of these items are from textbooks of clinical epidemiology and health economics.

Using the results from economic evaluations with poor methodological quality to advise companies on how to allocate resources for OSH interventions may result in inappropriate decisions (51). The fact that statistical analysis of the cost differences and joint cost–effect estimates was seldom conducted, adds to the risk of misleading conclusions. Of the studies which met less than 75% of the quality criteria and did not conduct a statistical analysis, 15 (17, 19, 24, 26, 27, 29, 31, 35, 39–41, 44, 45, 48, 49) concluded that the intervention of interest was cost-effective while 6 (20, 21, 28, 38, 42, 43) did not make a conclusive statement of cost-effectiveness but reported positive cost savings or benefits. Although our review did not seek or find evidence of publication bias (or other forms of reporting bias), we note that the negative implications for decision-making is also, in principle, strengthened by the risk of publication bias, which may lead to overestimates of treatment effects, net cost savings, and/or cost-effectiveness (52).

Improving the quality of future economic evaluations

Our findings of poor methodological quality are not unique to the studies assessed in this review. This problem has been signaled in, for example, reviews of economic evaluations including financial appraisals targeting other specified sets of occupational health, mental health, stroke, and other healthcare interventions (8–10, 53, 54). Across these reviews, common strategies have been proposed to improve the methodological quality of future economic evaluations, and these strategies are equally applicable here. In sum, the strategies can be targeted at the researcher, the journal, or both. For researchers, acquiring better knowledge of key methodological principles underlying economic evaluations from basic training, key reference textbooks, use of practice guidelines or quality checklists, and collaboration with health technology assessment researchers or health economists are recommended. For journals, the impetus for improvement may come from ensuring journal reviewers are adequately schooled in economic evaluation methodology, and by adopting checklists for submissions as is the policy of the British Medical Journal. For both, attention should be paid to reporting in terms of transparency by researchers and use of website capabilities by journals in response to space constraints.

Based on the findings of our review, specific attention should be paid to improving the following five aspects related to internal validity. It should be noted that each of these aspects have implications for

generalizability (51, 55). First, economic evaluations from a corporate perspective should include an explicit description of the study population and the competing alternatives. With regard to study population, a clear presentation of clinical characteristics, inclusion and exclusion criteria, and drop-outs during follow-up is required to know if potential biases may taint the findings (11). Furthermore, corporate characteristics such as size and sector as well as descriptions of job functions should be provided. Explicit descriptions of the alternatives are needed in order to judge whether or not a meaningful comparison of interventions has been chosen (4).

Second, an explicit statement of the perspective is required as the chosen approach influences the selection of costs and outcomes (4). When a narrower perspective (such as a corporate perspective) is chosen, the rationale for not using the broader societal perspective should be provided. While economic evaluations for specific decision-makers are necessary and warranted (5–7), it may be short-sighted not to take note of the costs and consequences affecting other stakeholders as well as the broader socio-political context in which the study takes place. A presentation of all socially relevant costs and outcomes in a disaggregated form may facilitate the extrapolation of findings to other settings (55).

Third, attention needs to be paid to how costs and outcomes are measured and how the former are valued. For both costs and health-related work productivity outcomes, the measurement tools used for data collection should be clearly reported and the tools should be valid. In addition, the physical units of costs and changes in health-related work productivity should be reported. Caution needs to be exercised in limiting lost work productivity data to that extracted from insurance databases. This is because these data only reflect the lost work time of approved cases. In recent years, a number of measurement tools have been developed to measure changes in work productivity from health-related absenteeism and presenteeism (56, 57). The cost prices used for valuation along with their sources and the index year of the evaluation need to be clearly stated. Whenever possible, cost items should be valued based on the *actual resources consumed* as charges and tariffs do not always represent the actual unit costs (11, 58). For a similar reason, cost data from workers' compensation claims or other insurance forms should also be used with caution as the full cost of a claim is often not billed to the company in the form of higher insurance premiums (9). Furthermore, costs and consequences beyond one year should be discounted, that is, reduced to reflect that what is spent or saved in the future should not weigh as heavily in the decision-making process as what is spent or saved today (4).

Fourth, economic evaluations require that assumptions are made. Thus, it is necessary to test these assumptions in a sensitivity analysis. All variables are, in theory, candidates for the sensitivity analysis. However, if the variables are certain or preliminary analyses have shown that their impact on the results is minimal, then these variables may be excluded. Again, authors should provide justification for their choices so that readers can judge the plausibility of tested parameters for themselves (4). In economic evaluations from a corporate perspective, testing the assumptions behind how the changes in health-related productivity are valued is particularly relevant as this is the main outcome of interest.

Fifth, greater attention needs to be paid to characterizing the uncertainty around the cost estimates and joint cost–effect estimates. To quantify the precision of the cost and joint cost–effect estimates, non-parametric bootstrapping is the recommended statistical technique for dealing with the highly skewed nature of cost data (59). In this review, only two studies conducted a cost–effectiveness analysis. Should future economic evaluations from a company's perspective involve a cost–effectiveness analysis, then a method of visually representing uncertainty – known as cost–effectiveness acceptability curves – may assist with the interpretation of results (60). However, it probably is very difficult for chief executive officers, chief financial officers, and managers to understand the statistics of economic evaluations. Implementation of findings of economic evaluations is important. Increasing the likelihood that companies will read, understand, and use the results of economic evaluations is one of the main challenges in field of OSH.

Strengths and limitations

The main strength of this study is that it is the first review focusing on the methodological quality of economic evaluations of OSH interventions conducted from a corporate perspective. Furthermore, we conducted the methodological appraisal using a standardized quality checklist based on consensus among experts in economic evaluation. A limitation of the CHEC-list may be that not all criteria are independent. For example, it is more difficult to code "yes" against criterion 6 ("Is the actual perspective chosen appropriate?") if criterion 3 ("Is a well-defined research question posed in answerable form?") is coded "no". However, the main aim of the CHEC-list is to assess the risk of bias of economic evaluations. Regardless of this potential dependency of topics, the risk of bias is higher in studies that do not meet these criteria. Also, because the CHEC-list does not add up to a total score, this potential dependency will not have an impact on the results of a systematic review in which the CHEC-list is used. Also, although we systematically searched four databases as well as the reference lists of included studies and our own database, we cannot guarantee that we captured all eligible

studies. However, given that our finding is not uncommon, we do not believe that our main conclusion and the relevance of the recommendations for improving the methodological quality in future economic evaluations would be significantly altered with additional studies.

Additional considerations

Two observations about the economic analyses warrant further attention by stakeholders in occupational health research. The first observation is that the predominant form of economic evaluation was a financial appraisal. From a health economics perspective, a financial appraisal represents a partial form of economic evaluation, and is, therefore, less appropriate for informing decisions. In particular, the incompleteness of financial appraisals can be traced back to the fact that the health benefits are not included in the monetary expression of the consequences by using principles of willingness-to-pay. The monetary consequences are limited to those related to healthcare use and increased productive output. From a practical standpoint, however, it may be that financial appraisal will suffice for corporate decision-makers. Such an assumption should be tested as it means that the savings from improved health are ignored. The omission of health improvements from the equation may be related to the fact that methodology to value health improvement in monetary terms is still evolving. The second observation is that within the financial appraisals, the final expression of the economic benefit for the company was expressed in five different ways: (i) net benefits or savings, (ii) ROI, (iii) internal rate of return, (iv) payback period, or (v) benefit-to-cost ratio. An expression of the difference in monetary benefits and program costs (ie, referred to as "net benefits or saving" in the included studies and net present value in health economic terms) is preferred over ratios because the net present value is straightforward to calculate and interpret. Ratios are sensitive to what is placed in the numerator and denominator, and unless the contents of the numerator and denominator are clearly described, ratios from different studies cannot be meaningfully compared (4).

Concluding remarks

While exceptions were identified, the overall methodological quality of the identified economic evaluations from a corporate perspective was poor. In particular, attention should be paid to the measurement and valuation of costs, sensitivity analysis, and characterization of uncertainty around the cost and cost–effect estimates. Also, the sufficiency of well-conducted financial appraisals for informing company decisions with regard to occupational health interventions should be tested.

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement 200549 EcOSH and financial support from the Dutch Ministry of Social Affairs and Employment.

The authors extend their gratitude to Ingrid Riphagen for her time and expertise in developing and running the primary search strategy in Embase.com. The authors also thank Stefan IJmker for his assistance with conducting the literature search in RILOSH.

References

1. van der Beek AJ, Frings-Dresen MH, van Dijk FJ, Houtman IL. Priorities in occupational health research: a Delphi study in The Netherlands. Occup Environ Med. 1997;54:504–10.
2. Leigh JP. Expanding research on the economics of occupational health [editorial]. Scand J Work Environ Health. 2006;32(1):1–4.
3. Burdorf A. Economic evaluation in occupational health – its goals, challenges, and opportunities [editorial]. Scand J Work Environ Health. 2007;3(3)3:161–4.
4. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford (United Kingdom): Oxford University Press; 2005.
5. Brouwer WB, van Exel NJ, Baltussen RM, Rutten FF. A dollar is a dollar is a dollar – or is it? Value Health. 2006;9:341–7.
6. Berger ML, Murray JF, Xu J, Pauly M. Alternative valuations of work loss and productivity. J Occup Environ Med. 2001;43:18–24.
7. Young AE, Wasiak R, Roessler RT, McPherson KM, Anema JR, van Poppel MN. Return-to-work outcomes following work disability: stakeholder motivations, interests and concerns. J Occup Rehabil. 2005;15:543–56.
8. Niven KJ. A review of the application of health economics to health and safety in healthcare. Health Policy. 2002;61:291–304.
9. 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;16:367–92.
10. Jefferson T, Demicheli V, Vale L. Quality of systematic reviews of economic evaluations in health care. JAMA. 2002;287:2809–12.
11. 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;21:240–5.
12. Brug J, van der Maas PJ, Roscam Abbing EW. Primaire preventie [Primary prevention]. In: Mackenbach JP, van der

Maas PJ, editors. Volksgezondheid en gezondheidszorg [Public health and health care]. 4th ed. Maarsen (the Netherlands): Elsevier Gezondheidszorg; 2008. p 189–237.

13. Green S, Higgins J. Glossary: Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [Internet]. The Cochrane Collaboration; 2005 [cited 7 February 2008]. Available from: http://www.cochrane.org/resources/handbook/

14. Domholdt E. Physical therapy research: principles and applications. Philadelphia (PA): WB Saunders Co; 1993.

15. Greenberg PE, Birnbaum HG, Kessler RC, Morgan M, Stang P. Impact of illness and its treatment on workplace costs: regulatory and measurement issues. J Occup Environ Med. 2001;43:56–63.

16. Hlobil H, Uegaki K, Staal JB, de Bruyne MC, Smid T, van Mechelen W. Substantial sick-leave costs savings due to a graded activity intervention for workers with non-specific sub-acute low back pain. Eur Spine J. 2007;16:919–24.

17. Banco L, Lapidus G, Monopoli J, Zavoski R. The safe teen work project: a study to reduce cutting injuries among young and inexperienced workers. Am J Ind Med. 1997;31:619–22.

18. Burdorf A, Windhorst J, Van der Beek AJ, Van der Molen HF, Meerding WJ. De kosten en baten van nieuwe arbeidsmiddelen bij straatmaken en vloerleggen [The costs and benefits of ergonomic equipment for street making and laying cement flooring]. Tijdschr Toegepast Arbowetenschap. 2005;1:3–11.

19. Collins JW, Wolf L, Bell J, Evanoff B. An evaluation of a "best practices" musculoskeletal injury prevention program in nursing homes. Injury Prev. 2004;10:206–11.

20. Engst C, Chhokar R, Miller A, Tate RB, Yassi A. Effectiveness of overhead lifting devices in reducing the risk of injury to care staff in extended care facilities. Ergonomics. 2005;48:187–99.

21. Hochanadel CD, Conrad DE. Evolution of an on-site industrial physical therapy program. J Occup Med. 1993;35:1011–6.

22. Landstad BJ, Gelin G, Malmquist C, Vinberg S. A statistical human resources costing and accounting model for analysing the economic effects of an intervention at a workplace. Ergonomics. 2002;45:764–87.

23. Linton SJ, Bradley LA. An 18-month follow-up of a secondary prevention program for back pain: Help and hindrance factors related to outcome maintenance. Clin J Pain. 1992;8:227–36.

24. Melhorn JM, Wilkinson L, Gardner P, Horst WD, Silkey B. An outcomes study of an occupational medicine intervention program for the reduction of musculoskeletal disorders and cumulative trauma disorders in the workplace. J Occup Environ Med. 1999;41:833–46.

25. Mitchell LV, Lawler FH, Bowen D, Mote W, Asundi P, Purswell J. Effectiveness and cost-effectiveness of employer-issued back belts in areas of high risk for back injury. J Occup Med. 1994;36:90–4.

26. Nelson A, Matz M, Chen F, Siddarthan K, Lloyd J, Fragala G. Development and evaluation of a multifaceted ergonomics program to prevent injuries associated with patient handling tasks. Int J Nurs Stud. 2006;43:717–33.

27. Shi L. A cost-benefit analysis of a California county's back injury prevention program. Public Health Rep. 1993;108:204–11.

28. Spiegel J, Yassi A, Ronald LA, Tate RB, Hacking P, Colby T. Implementing a resident lifting system in an extended care hospital: demonstrating cost-benefit. AAOHN J. 2002;50:128–34.

29. Versloot JM, Rozeman A, Van Son AM, Van Akkerveeken PF. The cost-effectiveness of a back school program in industry: a longitudinal controlled field study. Spine. 1992;17:22–7.

30. Lo Sasso AT, Rost K, Beck A. Modeling the impact of enhanced depression treatment on workplace functioning and costs: a cost-benefit approach. Med Care. 2006;44:352–8.

31. Smoot SL, Gonzales JL. Cost-effective communication skills training for state hospital employees. Psychiatr Serv. 1995;46:819–22.

32. Blaze-Temple D, Howat P. Cost benefit of an Australian EAP. Employee Assist Q. 1997;12:1–24.

33. Colombo GL, Ferro A, Vinci M, Zordan M, Serra G. Cost-benefit analysis of influenza vaccination in a public healthcare unit. Ther Clin Risk Manag. 2006;2:219–26.

34. Thomson P, Cuddeford G, Mitchell P. Hospital staff absenteeism following an influenza immunisation program. J Occup Health Saf Aust N Z. 1999;15:231–42.

35. Campbell DS, Rumley MH. Cost-effectiveness of the influenza vaccine in a healthy, working-age population. J Occup Environ Med. 1997;39:408–14.

36. Cohen P, Darling C, Hampson A, Downs K, Tasset-Tisseau A. Influenza vaccination in an occupational setting: effectiveness and cost-benefit study. J Occup Health Saf Aust N Z. 2003;19:167–82.

37. Franzblau A, Werner RA, Yihan J. Preplacement nerve testing for carpal tunnel syndrome: is it cost effective? J Occup Environ Med. 2004;46:714–9.

38. Schultz AB, Lu C, Barnett TE, Yen LT, McDonald T, Hirshland D, et al. Influence of participation in a worksite health-promotion program on disability days. J Occup Environ Med. 2002;44:776–80.

39. Dille JH. A worksite influenza immunization program: impact on lost work days, health care utilization, and health care spending. AAOHN J. 1999;47:301–9.

40. Samad AH, Usul MHBH, Zakaria D, Ismail R, Tasset-Tisseau A, Baron-Papillon F, et al. Workplace vaccination against influenza in Malaysia: does the employer benefit? J Occup Health. 2006;48:1–10.

41. Skisak CM, Bhojani F, Tsai SP. Impact of a disability management program on employee productivity in a petrochemical company. J Occup Environ Med. 2006;48:497–504.

42. Aldana SG, Merrill RM, Price K, Hardy A, Hager R. Financial impact of a comprehensive multisite workplace health promotion program. Prev Med. 2005;40:131–7.

43. Littleton M. Cost-effectiveness of a prework screening program for the University of Illinois at Chicago physical plant. Work. 2003;21:243–50.

44. Morales A, Martinez MM, Tasset-Tisseau A, Rey E, Baron-Papillon F, Follet A. Costs and benefits of influenza vaccination and work productivity in a Colombian company

from the employer's perspective. Value Health. 2004;7:433–41.

45. Shephard RJ. Long term impact of a fitness programme – the Canada Life Study. Ann Acad Med Singapore. 1992;21:63–8.

46. 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(1):36–46.

47. Alexander RW, Brennan JC, Maida AS, Walker RJ. The value of preplacement medical examinations for nonhazardous light duty work. J Occup Med. 1977;19:107–12.

48. Legg RF, Sclar DA, Nemec NL, Tarnai J, Mackowiak JI. Cost benefit of sumatriptan to an employer. J Occup Environ Med. 1997;39:652–7.

49. Lofland JH, Kim SS, Batenhorst AS, Johnson NE, Chatterton ML, Cady RK, et al. Cost-effectiveness and cost-benefit of sumatriptan in patients with migraine. Mayo Clin Proc. 2001;76:1093–101.

50. Higgins JPT, Altman DG, editors. Assessing risk of bias in included studies. In: Higgins JPT, Altman DG, editors. Cochrane Handbook for Systematic Reviews of Interventions. Chichester (United Kingdom): John Wiley & Sons, 2008.

51. Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ. 2001;323:42–6.

52. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet. 1991;337:867–72.

53. Evers SM, Van Wijk AS, Ament AJ. Economic evaluation of mental health care interventions: a review. Health Econ. 1997;6:161–77.

54. Evers SM, Ament AJ, Blaauw G. Economic evaluation in stroke research: a systematic review. Stroke. 2000;31:1046–53.

55. 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;21:165–71.

56. Kessler RC, Barber C, Beck A, Berglund P, Cleary PD, McKenas D, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med. 2003;45:156–74.

57. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. Expert Rev Pharmacoecon Outcomes Res. 2005;5:23–8.

58. Oostenbrink JB, Bouwmans CAM, Koopmanschap MA, Rutten FFH. Dutch manual for costing: methods and standard costs for economic evaluations in health care. Diemen (the Netherlands): Health Insurance Council (CVZ); 2004.

59. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York (NY): Chapman & Hall; 1993.

60. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3:309–19.

Received for publication: 21 October 2009