General discussion
Occupational health aims to preserve, restore and improve the health of workers. Health is a key intermediary to productivity, benefiting the workers, companies and society as a whole. One would expect that occupational health professionals have and can easily attract resources to address workers’ health needs. However, this is not the case, and evidence in the form of economic evaluations is needed to demonstrate the value of occupational health. The aim of this dissertation was to address two aspects of economic evaluations pertaining to occupational health: (i) to explore how the methodology of economic evaluations from a company’s perspective could be improved; and (ii) to generate new evidence of the cost-effectiveness of interventions directed at workers’ health.

This general discussion is divided into four parts. First, the main findings from the three studies exploring how the methodology of economic evaluations from a company’s perspective (chapters 2, 3 & 4) will be summarized along with suggestions for improvement. Second, the main findings from the three economic evaluations (chapters 5, 6 & 7) will be presented and methodological considerations will be discussed. Third, recommendations for practice and research will be presented. The discussion ends with concluding remarks.

I. How can the methodology of economic evaluations from a company’s perspective be improved?

The main findings from our exploration into the methodology of economic evaluations conducted from a company’s perspective were:

- in general, the methodological quality of economic evaluations of occupational health interventions from a company’s perspective is poor. Less than half of the 34 eligible studies in our systematic reviews met more than half of the 19 criteria of a methodological quality checklist. Examples of quality criteria that were least fulfilled included those related to the measurement and valuation of costs, measurement of outcomes and sensitivity analysis, which are all essential aspects of economic evaluation (chapter 2).

- methods used to measure and value changes in health-related productivity vary widely in terms of collecting sick leave data, inclusion of productivity changes due to decreased work performance and price weight used (chapter 3).

- the relevance of various items for estimating the cost of health-related productivity loss from a company’s perspective differ depending on the type of work loss at hand: work presenteeism, short-term absenteeism (< 2 weeks), and long-term absenteeism (> 2 weeks) (chapter 4).

First, our finding that the overall methodological quality of economic evaluations from a company’s perspective was poor is not unique. This may be reflective of the inherent situation for developing scientific fields where there is a lack of specific guidelines. Moreover, this may be due to the fact that this area of study is interdisciplinary, and that research methods and concepts pertaining to economic evaluation (predominantly) come from two scientific streams, health economic and business economics. Fulfilling and not fulfilling a quality criterion in our review may be reflective of similarities and differences between the two streams as the checklist we used was developed in health economics. The differences do not mean that the methodology of one discipline is better than the other. Rather, the methodology from both disciplines converge in
occupational health, and the differences point to criteria to which occupational health researchers should pay (extra) attention in order to move the quality of the economic evaluation evidence in occupational health forward. These include measurement and valuation of costs, measurement of outcomes and sensitivity analysis. We are not aware of a methodological quality checklist from business economics, however, a comparison between such a checklist and the one we applied may provide further insights into relevant quality criteria. This may be the start for developing field-specific quality checklists.

Another consideration is the blurry distinction between two complementary intentions behind economic evaluations: (i) determine whether an intervention is cost-effective for a representative worker or on a specific occupational health problem, and therefore worth considering for implementation; and (ii) determine whether implementation in a specific location for a specific population of workers is worthwhile and affordable. Evaluations conducted with the first intention generate evidence, which forms a vital part of the information needed to make a prioritization and budget allocation decision among a number of occupational health interventions. This was the focus of this dissertation. The second intention can be addressed by so-called budget impact analysis or business case analysis. The results are intended to help with budget planning and reaching targets for health outcomes. The usefulness of this distinction to help clarify and structure further efforts to improve economic evaluation methodology in occupational health warrants discussion. In addition, it is possible that efforts to improve methodology are hindered by confusion and miscommunication caused by researchers, schooled either in the methods from health economics or business economics, using the same terms but with different definitions and usage in mind. Therefore, future efforts should include standardizing terminology.

We observed that the final expression of the economic benefit for the company was expressed in monetary terms in five different ways: net benefits or savings, return-on-investment, internal rate of return, payback period, or benefit-to-cost ratio. We recommend expressing the economic benefit in terms of ‘net benefits or savings’ (i.e. net benefits or savings = A-B, where A represents the monetary benefits and B represents the costs) over ratios (i.e. ratio = A/B) because the net benefits or savings are 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. In any case, it is important to define what is understood under the monetary benefits (A) and what under costs (B). Other terms used to denote net benefits or savings are net cost difference and net present value.

Second, in economic evaluations from a company’s perspective, the outcome of interest is change in health-related productivity that can be translated into a monetary value. The findings from systematic review demonstrated that the methods to measure and value changes in health-related productivity varied widely. A certain extent of the variation may reflect the socio-political context in which the study takes place. In particular, studies conducted in countries that have a workers’ compensation system (e.g. Canada and United States) often used data from workers’ compensation databases. In such jurisdictions, there is a differentiation between work-related and non-work-related time loss. In contrast, in countries such as The Netherlands, such a differentiation does not exist and any work time loss due to a health problem is measured as sick leave.
Given the variation arising from context, prescribing one exact method for measurement and valuation is not realistic. Nevertheless, (collaborative) efforts to define standards or guidelines for good practice are warranted.

We offer the following suggestions for improving measurement and valuation. With regards to data collection, we first recommend the use of standardized questionnaires to measure changes in health-related productivity. A number of generic and disease-specific questionnaires have been developed. The questionnaires differ in terms of their scope measuring only work presenteeism to both sick leave, work presenteeism and additional aspects such as major events or compensatory mechanisms. Also, there are tools that measure qualitative changes in work performance and those that quantify changes into time-loss. Data from the latter can be easily translated into monetary values of health-related productivity loss in an economic evaluation. Second, the consensus-based recommendations may be useful as a checklist to identify how current data registration (systems) may be improved for the purpose of routine data collection.

With respect to valuation of time loss from sick leave and work presenteeism, one common price weight should be used for all subjects. This is because the difference in the effect of the intervention is on the difference in change in health-related productivity. The valuation is to help make the effect more interpretable and relevant. The use of worker-specific wages can distort the actual difference. Valuation also differs in terms of the inclusion of factors such as elasticity; compensatory mechanisms; costs for recruiting, hiring and training replacements; and consideration of function characteristics in the form of wage-multipliers. It is important to report their inclusion as well as the rationale in order to provide insight into potential biases. Currently, there is no consensus regarding the inclusion of these factors.

The methodology surrounding measurement and valuation will also benefit from clear reporting. With respect to measurement methods and time units of health-related productivity changes, this would mean extracting the amount of work loss from administrative databases, instead of presenting only costs. An example of how costs may be misleading is a situation where billed charges in an insurance database do not reflect actual cash payments or costs. However, extracting work loss data may be a challenge when dependent on administrative databases originally designed for administrative purposes, such as insurance claims, and not analytic studies. To facilitate transferability, the composition and source of corresponding price weights used to value the health-related productivity changes should be presented. Finally, the socio-political context in which the study takes place should be described so that readers from other jurisdictions can see how the distribution of costs and gains is similar or different to theirs.

The overall quality of future economic evaluations in occupational health may be improved by the following strategies directed at researcher and journal levels. For researchers, acquiring better knowledge of key methodological principles underlying economic evaluations is recommended. Basic training, guidelines, key reference textbooks and quality checklists are useful to this respect. Collaboration with health technology assessment researchers or health economists is an essential prerequisite for conducting and reporting of a high quality economic evaluation. Also, efforts to develop international consensus regarding concepts and methods are needed. For journals, impetus for improvement may come from ensuring journal reviewers are adequately schooled in economic evaluation methodology, and by adopting checklists for submission.
and review (the policy of the British Medical Journal is a good example\(^2\)). For both researchers and journals, attention should be paid to adequate reporting. Researchers should be transparent in their report and checklists may help in choosing what to report. Journals could use the capabilities of internet in response to space constraints, allowing authors to report all relevant aspects of their economic evaluation.

**II. Are the evaluated interventions cost-effective?**

The main findings of the three economic evaluations conducted in this dissertation were, respectively:

- A graded activity intervention for airline workers on sick leave due to non-specific subacute low back pain may potentially result in cost savings in terms of productivity loss compared to usual care (chapter 5, the AGAS study).
- A general practitioner-based minimal intervention for workers with stress-related sick leave was not cost-effective in improving quality of life or reducing sick leave compared to usual care (chapter 6, the MISS study).
- A supervisor case management intervention was not cost-effective in improving quality of life or reducing productivity loss in working mothers following childbirth compared to common practice (chapter 7, the Mom-at-work study).

**Methodological considerations**

**Perspective**

We conducted economic evaluations from a societal perspective and a company’s perspective. The main advantage and disadvantage of societal perspective are the converse for the company’s perspective. The main advantage of the societal perspective is that all costs and consequences, regardless of who pays or receives, are taken into consideration. It is most complete, providing insight into (re)distributional or substitution effects and improving generalizability of the findings\(^8\). However, the results are not directly interpretable for company decisions because certain costs and consequences that are not relevant from the company’s point of view may be included. For example, resource use from the health care sector, occupational health and alternative care sector, patients and their families, and changes from both paid and unpaid productivity would be included in an analysis from a societal perspective. From a (Dutch) company’s perspective, the analysis would be limited to costs of occupational health care and paid productivity. Also, the choice of price weight may differ between perspectives. For example, in the case of paid productivity loss from sick leave, a price weight based on the national average may be used in an analysis from a societal perspective, whereas from a company’s perspective, the price weight may be based on an average of the participating companies.

In the first two applied studies (chapters 5 and 6), we limited the analysis to one perspective. However, there is no restriction on the number of perspectives that can be taken. In the discussion of our third applied study (chapter 7), we supplemented the findings from a societal perspective with an estimation from a company’s perspective. Given the complexity of decision making in occupational health (compared to the health care arena) with multiple stakeholders\(^2\) and the fact that decisions are based on an interaction between stakeholders rather than one decision authority, future studies should use a combination of perspectives\(^2\). Being able to make head-to-head comparisons between perspectives will facilitate decision making as sources of agreements and
discrepancies will become visible, and negotiation points between parties will be identified.

On a practical note, during the design phase, a thorough review of current legislation is indicated to understand the distribution of burden. For example, as in the case of the Mom-at-work study (chapter 7), legislation may spread the burden of sick leave benefits across two stakeholders, and this will affect how data need to be collected and analyzed. To determine which specific stakeholder perspective warrants particular attention, it may be helpful to ask oneself the following questions: (i) where will the intervention be implemented; (ii) who decides if the intervention will be implemented; (iii) who will pay for the intervention; (iv) who will benefit from the intervention; (v) which outcome(s) is(are) the intervention designed to affect; and (vi) who is funding the study. In any case, a societal perspective will always be warranted as the same categories of data will be collected regardless of changes in legislation within a country, which can affect who pays or gains. This also facilitates comparison of studies across countries with different legislation.

Study population
In the AGAS study, workers on sick leave due to subacute low back pain were targeted. The study was performed within a single airline company and the majority of subjects were male, blue-collar workers. Also, the majority of employees had direct service-related jobs (e.g. airplane maintenance technicians, pilots, baggage handlers) where the possibility of postponing production until later or “doing more with less” were not viable options. Although the study was conducted within a single company, we assume that the results are generalizable to other companies with similar populations of male, blue-collar workers.

The MISS study population was reflective of the patient population in general practice: heterogeneous, typically having a mixture of psychological and somatic symptoms. Current primary care diagnostic or dimensional systems to distinguish between patients with subacute psychopathology that is clearly related to stress (SMD) from those with a mixture of symptoms fall short. We expected that patients with depressive, anxiety or somatic symptoms would still benefit from the MISS as a result of more accurate diagnosis and focused treatment for these problems in addition to attention for the stress-related component. To prevent selection bias, the researchers, not the GPs, recruited patients. Also, to prevent contamination, we conducted the randomization at the level of GPs instead of patients. This randomization procedure had an additional advantage in that it safeguarded GPs from possible ethical dilemmas related to withholding a potentially beneficial treatment. Furthermore, GPs were only informed of their patients’ participation in the study 2-month after enrolment. Thus, during the first two months of treatment when the most treatment was provided, knowledge of enrolment in the study could not have influenced a GP’s behaviour.

In the Mom-at-work study, selection bias may be a concern. First, we had difficulty recruiting companies to participate in this study. Participation may be reflective of the degree of problem recognition within the company, and in turn be reflective of the current workplace culture, policies for maternal health issues, workers’ health and sick leave, and motivation to improve the situation. As such, the participating companies may represent an optimal setting and any observed effect could be an overestimation. However, we did not observe an effect, therefore, selection bias in this sense is not
likely. Second, the women in the study had higher levels of education compared to the
general Dutch female population. Higher levels of education are associated with higher
socioeconomic status, and better post-partum health outcomes have been reported for
such groups of women\textsuperscript{29}. It is unclear if the intervention would have been more effective
in a population with a lower level of education.

\textit{Study power}

All sample sizes were calculated based on detecting differences in effects, not costs.
Although the direction of the findings from the cost analysis in the AGAS study (chapter
5) was robust and constantly in favour of the graded activity intervention for the entire
3-year follow-up period and the intervention group returned to work significantly earlier,
the mean cost differences were not statistically significant. This is a common problem
in economic evaluations alongside trials, which are underpowered for skewed data
distributions, such as costs of health utilization or productivity loss. Insufficient power
to detect relevant differences due to skewed cost data is possible\textsuperscript{30}. If studies would be
powered to detect cost-differences, they generally become too large, too expensive and
impossible to conduct. Modeling studies could solve this problem, but they are only
as good as the data that are included. If relevant data are not available, prospective
collection of data alongside RCTs seems to be the most valid solution.

\textit{Missing data}

Each economic evaluation had missing data. To avoid magnifying the problem of
low power and inefficiency, missing data should be imputed. Methods for imputing
missing data can be divided into naïve and principled methods. Naïve methods include
no imputation (complete cases analysis, available case analysis), last-outcome carried
forward method, and hot-decking\textsuperscript{31}. In the AGAS study, we imputed missing data using
the hot-deck method. Recently, multiple imputation has been recommended as the
method of choice\textsuperscript{32 33}. In the MISS and Mom-at-work studies, we handled missing data
by using multiple imputation. While multiple imputation is recommended for handling
missing data, we acknowledge that it is not a solution to prevent missing data in the first
place. Even with principled methods of imputation, the estimate of costs and effects are
less reliable than with a complete data set. If the number of dropouts and/or missings
is large and differs between intervention groups, even the best imputation techniques
will not provide a valid estimate of cost-effectiveness. Researchers should always put
as much time and effort as possible into collecting an “as complete as possible” data set.

\textit{Adequacy of the intervention}

In the MISS study, lack of contrast and poor implementation may have affected the
adequacy of the intervention. With respect to lack of contrast between the two intervention
groups, we could not rule out 100% that usual care in the study may not have been the
same as usual care in “real life”. In the study, usual care of GPs may have been influenced
by altered patient behavior due to patients making a connection between stress and
sick leave from the recruitment information and screening questionnaire. However, we
believe that any difference would be minor as patients received the information and
questionnaire after contact with the GP was already initiated. Also, in addition to the
measures taken to prevent contamination, GPs randomized to the usual care group did
not receive any extra information or instruction on treatment, and actual treatment was
left to their discretion. Because patients in both groups were provided with the same limited information about the study, such an effect would have occurred similarly in both groups.

With regards to poor implementation, this was possible. The training provided to the GPs in the MISS group may not have been sufficient for mastery of the desired skills, resulting in inadequate implementation of the intervention. Implementation may be improved by more training as well as greater dialogue between GPs and occupational physicians in order to establish a stronger culture of collaboration in the care of workers with stress-related sick leave. Future intervention trials should pay particular attention to providing adequate training of professionals in order to ensure sufficient mastery of the new treatment method or approach.

In the Mom-at-work study, it was possible that intensity of the intervention was too low. The one-time phone call at 6-weeks may have been insufficient as new problems may arise between 6-weeks and the official end of the maternity leave. Follow-up contact, for example, at 8- or 9-weeks and just prior to expected return-to-work, may be improvements.

Choice of outcome
The quality-adjusted life year (QALY) is recommended as the standard outcome for economic evaluations from a societal perspective. It is a measure that captures health gains from both reduced morbidity and reduced mortality, and combines these two aspects into a single measure based on the relative desirability of the different outcomes. The composite, generic nature of the QALY and the incorporation of the notion that some health outcomes are more valued than others are its advantages over program-specific measures such as pain (symptom) reduction or disability-days-averted. Decision makers with limited budgets need to be able to make comparisons so that they maximize health gain and determine which interventions to reduce or eliminate to free up funding for the new one.

However, generic QALYs may be insensitive for assessing the effects of interventions involving persons with mental health problems as in the MISS study. To mitigate this problem, inclusion of a disease-specific instrument is recommended. Also, while the QALY is recommended for economic evaluations from a societal perspective, this recommendation comes from the health care literature. Whether or not this recommendation can be applied directly to occupational health is questionable. Perhaps a generic measure capturing the quality and quantity of working life is warranted. An example is a recently conceptualized, so-called productivity-adjusted life year, which aims to express the amount of health and productive time lost due to poor working conditions and associated illness.

Estimating costs
The first issue with estimating costs is measurement. Measurement strategies may be prospective using cost diaries or retrospective using questionnaires, and data collection may be continuous or at discrete moments. Prospective, continuous data collection would be ideal, however, not always feasible in practice. In our studies, we made choices based on cost drivers and timing of greatest resource condition given the health problem.

In the AGAS study, we used cost diaries to collect health care utilization data. Health care use data were collected prospectively for the first 3 months, and then at 6 and 12
months. Sick leave data were extracted from the database of the company. In the Mom-at-work study, we used questionnaires at discrete moments spanning the entire follow-up year. All subjects received questionnaires that covered the first 24-weeks post-partum and the last 4 weeks from 48-52 weeks post-partum. Between 24- and 48-weeks, we collected extra data from women who continued to have problems in the preceding period. For those without problems, we assumed resource use to be zero. In the MISS study, we collected data on GP consultations, medications, lab tests and sick leave continuously, and the remaining resource use data were collected at four discrete moments. In both the AGAS and MISS studies, we linearly interpolated from the observed months (5 and 4 months, respectively) to 12 months by assuming a linear time trend in resource use. This may have resulted in an overestimation or underestimation in these studies if resource use between two consecutive measurement moments was not linear over time. However, in both these studies, these costs were a small proportion of the total costs and the over- or underestimation may have occurred in both groups; therefore, we do not expect that this would alter the overall findings.

The second issue relates valuation, that is, to choosing price weights to translate the units of resource consumption into costs. In making our choices, we tried to strike a balance between the desire for standardization on the one hand, and respecting the specific context of the studies on the other. To value health care sector, other sector and patient/family resource use, we used standard price weights whenever possible. Standard price weights are preferable to, for example, tariffs because tariffs contain an element of negotiation such that the level of the tariff does not correspond with the actual cost of providing the particular service. Also, the use of standard price weights facilitate the comparability of the results with other studies. To determine the price weight to value paid productivity loss, we took the specific context of the studies into heavier consideration. For the AGAS study, which was conducted from a company’s perspective and took place in a single company, we calculated a price weight specific for the study population. The MISS and Mom-at-work studies were conducted from a societal perspective, therefore, we determined price weights based on national data. Because the study population of the MISS study included both genders and a wide range of ages, we used an overall national average. For the Mom-at-work study, the price weight was based on national averages for women and weighted for age groups from 15-24 years through to 45-54 years.

The third issue relates to estimating the cost of health-related productivity loss. In the AGAS study, we based our estimation of the cost of health-related productivity loss on the number of sick leave days. However, it is not clear how accurate such an estimation is. First of all, the actual level of production loss may be influenced by the type of compensation mechanisms that exist within a particular job. For example, work normally performed by the absent employee may be completed by colleagues or made up upon return to work during usual working hours. In this situation, the absenteeism of the employee would not lead to productivity loss and therefore, should not be considered a cost. In the AGAS study, however, the majority of employees had service-related jobs (e.g. airplane maintenance technicians, pilots, baggage handlers) where the possibility of postponing production until later or “doing more with less” were not viable options. Furthermore, the way in which the cost of health-related productivity loss is estimated can lead to different results. In the AGAS study, we estimated productivity loss in four different ways: gross, net and with an assumption of either 25% or 50% decreased work
performance. On the one hand, the gross estimation is likely to be an overestimation as workers who return to their original duties on a part-time basis, or who perform alternative job tasks, conduct work in some format. On the other hand, the net cost estimation may be an underestimation as workers who are still not fully recovered may not be 100% productive. The sensitivity analysis in which a 25% or 50% decrease in work performance was assumed, albeit arbitrarily, takes this into consideration. The inaccuracy of lost productivity estimation could be partially solved by the use of questionnaires for measuring health-related work performance.

In the MISS study, we did not include work presenteeism or unpaid labor as part of our assessment of worker productivity. Before and after periods of sick leave, work performance and unpaid labor of individuals with stress-related mental health problems may be reduced. It is possible that the MISS intervention and usual care can have different effects on these components, which were not captured. Thus, the picture of functional recovery and different effect of the interventions may be incomplete. Future research studies should include these aspects.

In the Mom-at-work study, we measured productivity loss arising from sick leave and work presenteeism using the Health and work Performance Questionnaire (HPQ). We observed that 58% of the women reported work presenteeism, suggesting that work presenteeism is a common problem. This is in line with findings from other studies investigating other working populations and health problems. For example, Aronsson et al. found that 37% of the Swedish workforce experienced work presenteeism. Among workers with high physical load jobs and health problem, 50% reported work presenteeism. We also found that the contribution of sick leave and work presenteeism to total productivity loss hours was roughly 50:50. That work presenteeism can represent a considerable proportion of total productivity loss, is in line with findings from Burton et al. in which the ratio of sick leave:work presenteeism was 40:60. Future studies on return-to-work on diseases or disorders where it is expected that workers may stay at work or early return to work before they are fully recovered should include work presenteeism.

**Time horizon**

An adequate length of follow-up is needed to assess the full impact of a given intervention. In randomized controlled trials, a follow-up of 1-year is common as the cost and time requirements for longer periods can be prohibitive. However, for clinical problems with a multi-year time horizon, a 1-year follow-up may be insufficient to capture all downstream consequences of the interventions in question. This is a consideration for the MISS study. Extrapolation by means of decision models may offer a way to bridge the gap between the observations from trials and what may be expected over the long term. However, we did not find any differences in effect between the MISS intervention and control groups after one year follow-up. It is very unlikely that a difference would occur at a later stage. With respect to the Mom-at-work study, the focus was on health problems and sick leave following childbirth and involved a minimal intervention. As such, a 1-year follow-up is sufficient. Long-term follow-up is a relevant issue with respect to low back pain. In the AGAS study, we evaluated the effect of the intervention on sick leave in the 2nd and 3rd year post-intervention. The mean difference in the cost of health-related productivity loss was in favour of the graded activity group, though not statistically significant.
Generalizability
All three economic evaluations were conducted within pragmatic study designs in which the (new) intervention was compared to usual care and evaluated in “real life” versus “ideal” circumstances. This facilitates the generalizability of the findings to similar Dutch companies and Dutch general practice. The use of a societal perspective in the MISS and Mom-at-work studies also increases the generalizability of their respective study findings.

The generalizability of findings may be limited by system differences between The Netherlands and other countries. With respect to the Mom-at-work study, differences in maternity leave policy should be noted. For example, the length of maternity leave in The Netherlands is 16 weeks with full pay whereas mothers in the UK are entitled to 52 weeks of paid and unpaid leave. In the USA, paid maternity leave is not standard: among the best U.S. employers, 7% did not offer any paid leave, 17% between 1-4 weeks, 48% between 5-8 weeks, 20% between 9-12 weeks, and 8% more than 12-weeks. Also, differences in the organization of primary and occupational health care, and the older average age of Dutch women giving birth may result in different patterns of resource use.

III. Recommendations

For occupational health practice
The observations from the main and ancillary analyses from the MISS study underscore that (preclinical/mild) mental health problems seen in primary care are heterogeneous and cannot be treated in the same way. Continued efforts to improve general practitioners’ differential diagnostic skills and provide focused treatment are warranted.

Based on the findings of the economic evaluations, widespread implementation of the GP-based minimal intervention for workers with stress-related sick leave and the supervisor case management intervention for working mothers following childbirth are not indicated. Although the findings were not statistically significant, implementation of the graded activity intervention for workers with subacute low back pain may be warranted.

Occupational health professionals need to be aware that a number of occupational health interventions exist, which have not yet been evaluated for effectiveness and cost-effectiveness or persist in use even after health claims are unfounded. Evidence should precede implementation for two practical reasons: already scarce resources will become scarcer, and it is always more difficult to stop the use of something once it has been introduced.

For occupational health research
Given that decision-making in occupational health is often based on interactions between stakeholders, efforts to improve current methodology and conduct economic evaluations should be in collaboration with stakeholders. Also, researchers should be active in knowledge translation activities and participate in efforts to teach users of results to be critical consumers. Research methods, terminology and concepts used in economic evaluations of occupational health interventions (mainly) come from health economics and business economics. Future initiatives to improve the methodological quality of economic
evaluations of occupational health interventions should also aim to clarify the similarities and differences between these two scientific streams. This is so that confusion and miscommunication do not hinder progress.

Methodological issues that warrant further inquiry include the conceptualization of compensation mechanisms, and the interface between what stakeholders identify as relevant, as in our recommendations, and how to measure (via measurement tools) and to value productivity loss.

A practical lesson for future intervention trials is that sufficient attention should be paid to providing adequate training of professionals in order to ensure sufficient mastery of the new treatment method or approach.

Strategies to improve the methodology of future economic evaluations in occupational health may 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, impetus for improvement may come from ensuring journal reviewers are adequately schooled in economic evaluation methodology, and by adopting checklists for submissions.
- For both, attention should be paid to reporting in terms of transparency by researchers and the use of internet capabilities by journals in response to space constraints.

Finally, further impulses to improve overall methodology may come from taking lessons from methodological developments in health economics. For example, application of the net benefit regression model\textsuperscript{52} or cost-consequence analysis\textsuperscript{53}.

IV. Conclusion

Conducting economic evaluations in occupational health is a complex endeavor. A future task for the field of occupational health should be to clarify methods, terminology and concepts, and establish best practice guidelines for economic evaluations of occupational health interventions. These guidelines should provide explicit recommendations for reporting key aspects of methodology underlying the measurement and valuation of changes in health-related productivity. The use of standardized methodological quality assessment tools such as the CHEC-list\textsuperscript{24,54} can also be helpful. Improvements in reporting will positively affect transparency and simplify the process of determining the comparability of studies. In addition, it will also facilitate the transferability of results.

Two recent events demonstrate how economic evaluations of occupational health and safety are a research priority not only in The Netherlands, but internationally. The first is the 2008 publication of the textbook, “Economic evaluation of interventions for occupational health and safety: Developing good practice”, edited by E. Tompa, A.J. Culyer and R. Dolinschi\textsuperscript{23}. The second is the organization of a series of workshops in 2009 on economic evaluation in occupational safety and health (ECOSH) funded by the European Union\textsuperscript{55}. The aims of the workshops include reaching consensus on optimal methods of economic evaluation in occupational safety and health. Therefore... stay tuned for the developing story.
References


Kurtz LC, Sword W, Ciliska D. Urban women’s socioeconomic status, health service needs and utilization in the four weeks after postpartum hospital discharge: findings of a Canadian cross-sectional survey. BMC Health Serv Res 2008;8:203.


Oostenbrink JB, Al MJ. The analysis of incomplete cost data due to dropout. Health Econ 2005;14:763-76.


Economic evaluation of interventions for occupational health