Cochrane Library

Cochrane Database of Systematic Reviews

Self-certification versus physician certification of sick leave for reducing sickness absence and associated costs (Review)

Kausto J, Verbeek JH, Ruotsalainen JH, Halonen JI, Virta LJ, Kankaanpää E

Kausto J, Verbeek JH, Ruotsalainen JH, Halonen JI, Virta LJ, Kankaanpää E.
Self-certification versus physician certification of sick leave for reducing sickness absence and associated costs.
Cochrane Database of Systematic Reviews 2019, Issue 5. Art. No.: CD013098.
DOI: 10.1002/14651858.CD013098.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Self-certification versus physician certification of sick leave for reducing sickness absence and associated costs

Johanna Kausto[1], Jos H Verbeek[2], Jani H Ruotsalainen[2], Jaana I Halonen[1], Lauri J Virta[3], Eila Kankaanpää[4]

[1]Finnish Institute of Occupational Health, TYÖTERVEYSLAITOS, Finland. [2]Cochrane Work Review Group, Finnish Institute of Occupational Health, TYÖTERVEYSLAITOS, Finland. [3]Social Insurance Institution of Finland (Kela), Turku, Finland. [4]University of Eastern Finland, Kuopio, Finland

Contact: Johanna Kausto, Finnish Institute of Occupational Health, P.O. Box 40, TYÖTERVEYSLAITOS, FI-00032, Finland. johanna.kausto@ttl.fi.

Editorial group: Cochrane Work Group.
Publication status and date: New, published in Issue 5, 2019.

Citation: Kausto J, Verbeek JH, Ruotsalainen JH, Halonen JI, Virta LJ, Kankaanpää E. Self-certification versus physician certification of sick leave for reducing sickness absence and associated costs. *Cochrane Database of Systematic Reviews* 2019, Issue 5. Art. No.: CD013098. DOI: 10.1002/14651858.CD013098.pub2.

ABSTRACT

Background

From the societal and employers' perspectives, sickness absence has a large economic impact. Internationally, there is variation in sickness certification practices. However, in most countries a physician's certificate of illness or reduced work ability is needed at some point of sickness absence. In many countries, there is a time period of varying length called the 'self-certification period' at the beginning of sickness absence. During that time a worker is not obliged to provide his or her employer a medical certificate and it is usually enough that the employee notifies his or her supervisor when taken ill. Self-certification can be introduced at organisational, regional, or national level.

Objectives

To evaluate the effects of introducing, abolishing, or changing the period of self-certification of sickness absence on: the total or average duration (number of sickness absence days) of short-term sickness absence periods; the frequency of short-term sickness absence periods; the associated costs (of sickness absence and (occupational) health care); and social climate, supervisor involvement, and workload or presenteeism (see Figure 1).

Search methods

We conducted a systematic literature search to identify all potentially eligible published and unpublished studies. We adapted the search strategy developed for MEDLINE for use in the other electronic databases. We also searched for unpublished trials on ClinicalTrials.gov and the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP). We used Google Scholar for exploratory searches.

Selection criteria

We considered randomised controlled trials (RCTs), controlled before-after (CBA) studies, and interrupted time-series (ITS) studies for inclusion. We included studies carried out with individual employees or insured workers. We also included studies in which participants were addressed at the aggregate level of organisations, companies, municipalities, healthcare settings, or general populations. We included studies evaluating the effects of introducing, abolishing, or changing the period of self-certification of sickness absence.

Data collection and analysis

We conducted a systematic literature search up to 14 June 2018. We calculated missing data from other data reported by the authors. We intended to perform a random-effects meta-analysis, but the studies were too different to enable meta-analysis.

Main results

We screened 6091 records for inclusion. Five studies fulfilled our inclusion criteria: one is an RCT and four are CBA studies. One study from Sweden changed the period of self-certification in 1985 in two districts for all insured inhabitants. Three studies from Norway conducted between 2001 and 2014 changed the period of self-certification in municipalities for all or part of the workers. One study from 1969 introduced self-certification for all manual workers of an oil refinery in the UK.

Longer compared to shorter self-certification*for reducing sickness absence in workers*

Outcome: average duration of sickness absence periods

Extending the period of self-certification from one week to two weeks produced a higher mean duration of sickness absence periods: mean difference in change values between the intervention and control group (MD_{change}) was 0.67 days/period up to 29 days (95% confidence interval (95% CI) 0.55 to 0.79; 1 RCT; low-certainty evidence).

The introduction of self-certification for a maximum of three days produced a lower mean duration of sickness absence up to three days (MD_{change} −0.32 days/period, 95% CI −0.39 to −0.25; 1 CBA study; very low-certainty evidence). The authors of a different study reported that prolonging self-certification from ≤ 3 days to ≤ 365 days did not lead to a change, but they did not provide numerical data (very low-certainty evidence).

Outcome: number of sickness absence periods per worker

Extending the period of self-certification from one week to two weeks resulted in no difference in the number of sickness absence periods in one RCT, but the authors did not report numerical data (low-certainty evidence).

The introduction of self-certification for a maximum of three days produced a higher mean number of sickness absence periods lasting up to three days (MD_{change} 0.48 periods, 95% CI 0.33 to 0.63) in one CBA study (very low-certainty evidence).

Extending the period of self-certification from three days to up to a year decreased the number of periods in one CBA study, but the authors did not report data (very low-certainty evidence).

Outcome: average lost work time per 100 person-years

Extending the period of self-certification from one week to two weeks resulted in an inferred increase in lost work time in one RCT (very low-certainty evidence).

Extending the period of self-certification (introduction of self-certification for a maximum of three days (from zero to three days) and from three days to five days, respectively) resulted in more work time lost due to sickness absence periods lasting up to three days in two CBA studies that could not be pooled (MD_{change} 0.54 days/person-year, 95% CI 0.47 to 0.61; and MD_{change} 1.38 days/person-year, 95% CI 1.16 to 1.60; very low-certainty evidence).

Extending the period of self-certification from three days up to 50 days led to 0.65 days less lost work time in one CBA study, based on absence periods lasting between four and 16 days. Extending the period of self-certification from three days up to 365 days resulted in less work time lost due to sickness absence periods longer than 16 days (MD_{change} −2.84 days, 95% CI −3.35 to −2.33; 1 CBA study; very low-certainty evidence).

Outcome: costs of sickness absence and physician certification

One RCT reported that the higher costs of sickness absence benefits incurred by extending the period of self-certification far outweighed the possible reduction in costs of fewer physician appointments by almost six to one (low-certainty evidence).

In summary, we found very low-certainty evidence that introducing self-certification of sickness absence or prolonging the self-certification period has inconsistent effects on the mean number of sickness absence days, the number of sickness absence periods, and on lost work time due to sickness absence periods.

Authors' conclusions

There is low- to very low-certainty evidence of inconsistent effects of changing the period of self-certification on the duration or frequency of short-term sickness absence periods or the amount of work time lost due to sickness absence. Because the evidence is of low or very low certainty, more and better studies are needed.

PLAIN LANGUAGE SUMMARY

Changing the length of time a worker is allowed to take time off work because of illness without a physician's certificate

What was the aim of this review?

To find out if it is possible to affect sickness absence by changing the length of time a worker is allowed to take time off work because of illness without a physician's certificate. We found five studies.

Key messages

We are uncertain whether changing the length of time a worker is allowed to take time off work because of illness without a physician's certificate has any effect on sickness absence, as the included studies reported very different results, and the certainty of the evidence was low to very low.

What was studied in the review?

Sickness absence prevents a person from working and thus may reduce income and causes costs to employers. Usually employers require a physician to certify sickness absence. This may not be meaningful in cases of less severe illnesses that pass quickly with rest. Self-certification of sickness absence is already used at many workplaces for sickness absence periods lasting typically from one day up to two weeks. In this review, we evaluated how change in the length of the self-certification period affects the mean number of sickness absence days, number of sickness absence periods, and the amount of lost work time at workplaces.

Why is this topic important?

Sickness absence is costly to society and to employers. Employers may have to continue paying the sick employee’s salary. After the employer’s obligation to pay has ended, insurance covers sickness benefits. Changing the practice of sickness certification in short sickness absences is expected to change employees' attitudes and behaviour, co-operation and climate at the workplace, and diminish sickness absence. Self-certification makes more healthcare resources available for other purposes.

What are the main results of the review?

We found five studies conducted between 1969 and 2014. One study evaluated the effect of prolonging the self-certification period among all insured workers in a large city and a region in Sweden in 1988. Three studies evaluated the effect of prolonging the self-certification period for employees of a few municipalities in Norway. One study evaluated the effect of introducing self-certification in an organisation in the UK in 1969. The time participants in the intervention groups were allowed to be off work without a doctor's certificate ranged from three days to one year. The included studies measured the effects on the mean number of sickness absence days, the number of sickness absence periods, or on lost work time due to sickness absence periods. All studies compared the effect of the change with practice-as-usual.

Effects on duration of sickness absence periods

Extending self-certification from one week to two weeks increased the mean duration of sickness absence. Introducing self-certification for a maximum of three days reduced the mean duration of sickness absences lasting up to three days. Extending self-certification from one to three days up to a year did not change the duration of sickness absence.

Effects on number of sickness absence periods

Extending self-certification from one week to two weeks did not change the number of sickness absence periods. Introducing self-certification for a maximum of three days increased sickness absence periods lasting up to three days. Extending self-certification from three days to up to a year decreased sickness absence periods.

Effects on lost work time

Extending self-certification from one week to two weeks resulted in an inferred increase in lost work time. Extending self-certification (from zero days to three days and from three days to five days) increased the amount of work time lost due to sickness absence periods lasting up to three days. Extending self-certification from ≤ 3 days to ≤ 50 days and from ≤ 3 days to ≤ 365 days reduced lost work time due to sickness absence periods of 4 to 16 days and > 16 days.

Costs of sickness absence and physician certification

The costs of sickness absence benefits resulting from a longer period of self-certification may be about six times greater than the possible reduction in costs of fewer physician appointments.

How up-to-date is this review?

We searched for studies up to 14 June 2018.

SUMMARY OF FINDINGS

Summary of findings for the main comparison. Longer self-certification compared to shorter self-certification for reducing sickness absence in workers

Longer self-certification compared to shorter self-certification for reducing sickness absence in workers

Patient or population: workers
Setting: work organisations
Intervention: longer self-certification period
Comparison: shorter self-certification period

Outcomes	Impact	№ of participants (studies)	Certainty of the evidence (GRADE)
RCT: Average duration of sickness absence periods in days Follow-up: 6 months	SC ≤ 14 days vs ≤ 7 days led to a 0.67 days longer duration of sickness absence periods ≤ 29 days (95% CI 0.55 to 0.79) in 1 study.	237,391 (1 RCT)	⊕⊕⊖⊖ LOW [1]
CBA: Average duration of sickness absence periods in days Follow-up: 1 year	SC ≤ 3 days vs SC 0 days led to 0.32 days shorter duration of sickness absence periods ≤ 3 days (95% CI −0.39 to −0.25) in 1 study. SC ≤ 365 days vs ≤ 3 days did not lead to a change according to the authors.	1869 (2 observational studies)	⊕⊖⊖⊖ VERY LOW [2 3]
RCT: Number of sickness absence periods per worker Follow-up: 6 months	SC ≤ 14 days led to no difference in the number of sickness absence periods compared to SC ≤ 7 days, but the authors did not report numerical data.	237,391 (1 RCT)	⊕⊕⊖⊖ LOW [1]
CBA: Number of sickness absence periods per worker Follow-up: 1 year	SC ≤ 3 days vs SC 0 days led to an increase in 1 study (0.48 sickness absence periods lasting ≤ 3 days, 95% CI 0.33 to 0.63), and SC ≤ 365 days vs ≤ 3 days did not lead to a considerable difference in the number of sickness absence periods, but the authors did not report numerical data.	1869 (2 observational studies)	⊕⊖⊖⊖ VERY LOW [2 3]
RCT: Average lost work time per 100 person-years Follow-up: 6 months	SC ≤ 14 days vs SC ≤ 7 days led to an inferred increase in lost work time.	237,391 (1 RCT)	⊕⊖⊖⊖ LOW [1 4]
CBA: Average lost work time per 100 person-years Follow-up: 1 year	SC ≤ 3 days vs 0 days led to 0.54 days more lost work time (95% CI 0.47 to 0.61) in 1 study. SC ≤ 5 days vs SC ≤ 3 days led to 1.38 days (95% CI 1.16 to 1.60) more lost work time in absence periods of ≤ 3 and 0.1 days (95% CI could not be calculated) in absence periods of 4 to 16 days in 1 study. SC ≤ 50 days vs SC ≤ 3 days led to 0.65 days less lost work time in 1 study based on absence periods of 4 to 16 days. SC ≤ 365 days vs SC ≤ 3 days led to a decrease of −2.84 days in lost work time (95% CI −3.35 to −2.33) based on absence periods of > 16 days.	17,114 (4 observational studies)	⊕⊖⊖⊖ VERY LOW [2 5]
RCT: Costs of sickness absence and physician certification	The cost of 1 day of sickness absence (SEK 296) was obtained using the average daily compensation paid by employers to employees. Given the cost of a physician appointment (SEK 445), the higher costs of sickness absence benefits (SEK 29,000,000) far outweighed the possible reduction in costs of fewer physician appointments (SEK 4,900,000) during 1 year.	237,391 (1 RCT)	⊕⊕⊖⊖ LOW [1]

CBA: controlled before-after study; **CI:** confidence interval; **RCT:** randomised controlled trial; **SC:** self-certification, which means that a worker is allowed to take short periods (of a predetermined maximum length) of sick leave due to minor illnesses without obtaining an official document from a doctor (i.e. physician certification).

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]We downgraded the certainty of evidence by two levels, that is from high to moderate and then from moderate to low, due to risk of bias. The authors performed randomisation by date of birth, which is an inadequate randomisation method, and they did not blind participants or intervention providers.
[2]We downgraded the certainty of evidence by one level, that is from low to very low, due to risk of bias because the authors did not take critical confounders into account. Given the small number of included studies, we could not assess the possible influence of publication bias. We found no reason to upgrade the certainty of the evidence.
[3]We would have downgraded the certainty of evidence by one more level due to inconsistency because two studies produced opposing results, but we had already reached very low certainty.
[4]We downgraded the certainty of evidence by one level, that is from low to very low, due to indirectness because we inferred the conclusion about lost work time from two other outcomes.
[5]We would have downgraded the certainty of evidence by one more level due to inconsistency because results in studies with less self-certification time showed an increase in lost work time, whereas two studies with much longer self-certification showed a decrease in lost work time, but we had already reached very low certainty.

BACKGROUND

Description of the condition

Sickness absence or sick leave can be defined as "absence from work that is attributed to sickness by the employee and accepted as such by the employer" (Whitaker 2001). 'Sick pay' refers to the continuation of payment of the employee's salary (or part of it) by the employer during the sickness absence. A sickness benefit is usually paid from insurance after the employer's obligation of sick pay has ceased (Spasova 2016). From the perspective of society and the employer, sickness absence has a large economic impact. The incidence of sickness absence and the methods that are used to calculate the costs vary among countries; however, a conservative estimate of the average cost of sickness absence to a nation is 2.5% of gross domestic product (GDP) (Eurofound 2010).

Definitions of short- and long-term sickness absence vary among studies. On one hand, long-term sickness absence has been defined as an absence exceeding 20 to 90 days or longer (Higgins 2012). On the other hand, the National Institute for Health and Care Excellence (NICE) guidance (UK) defines long-term sickness absence as four or more weeks (Gabbay 2011; NICE 2009). A short sickness absence from work is usually a minor disturbance from the perspective of the workplace, with work often taken over by colleagues. For longer-term sickness absence periods, replacements are needed that must be paid in addition to the payment of sickness benefit. It can be inferred that both short- and longer-term sickness absence affect productivity and the economic health of companies as well as that of the country. In this Cochrane Review, we have focused on short-term sickness absence, which is defined as an absence of less than four weeks, based on the NICE guidelines (NICE 2009).

Description of the intervention

An important procedure in the recognition of a need for sickness absence is the certification of illness or reduced work ability by a physician. This means that, in the case of sickness and reduced work ability, an employee needs a medical certificate to be absent from work. The certifying physician is often a general practitioner. It has been reported in Sweden, where self-certification is common, that 9% of visits to general practitioners included sickness certification (Englund 2000). Internationally, there is variation in sickness certification practices. Some countries, such as the Netherlands, do not use any official sickness certification procedures, and employers are free to organise sick leave procedures jointly with their employees. In other countries, there can be a period of varying length at the beginning of the sickness absence during which no medical certificate is needed. Usually the employee notifies his or her supervisor when taken ill, a procedure known as self-certification. The role of the supervisor may vary from merely registering the notification to actually assessing the situation before giving permission to stay at home without a medical certificate. The permitted duration of absence based on self-certification usually ranges from one to seven days (NOSOSCO 2015; Wynne-Jones 2008). As the self-certification practices may vary not only among countries, but also between workplaces within a country, the effects are usually studied at the aggregate level of organisations, companies, municipalities, healthcare settings, or general populations.

How the intervention might work

Sickness absence and return to work may be associated with multiple factors at the levels of the individual, work and work environment, other areas of life, features of the social insurance system, and the larger societal context. Consequently, there are several pathways (practical, attitudinal, and motivational) at various levels (individual employee, the workplace, and (occupational) health care) that may mediate the effect of changing the period of self-certification of sickness absence on short-term sickness absence and associated costs, social climate, supervisor involvement, and workload or presenteeism (working while ill; Figure 1). In a Finnish survey (Hinkka 2018), the majority of participating physicians found that the insufficient availability of healthcare services in the public sector was one reason for prolonged sickness absences. The practice by which the worker has to make an appointment with the physician and the physician has to be available may also prolong sickness absence, while self-certification practice could shorten the absence when access to health care is poor. These kinds of administrative factors may increase the duration of very short-term sick leave. For example, if a physician certificate is required, the appointment with the physician may delay return to work due to physician availability, even if the employee felt healthy enough to return to work earlier. The physician is also likely to make a conservative assessment of the work ability of the worker to prevent presenteeism. A self assessment might therefore decrease the number of days absent from work, because without a medical certificate, the employee may be more likely to return to work as soon as she or he feels fit to work, as no restrictions are set by a medical certificate (Pesonen 2016).

Self-certification can be interpreted by the employee as increased decision latitude (ability to make work-related decisions and personal control over work-related matters), while poor decision latitude at work has been associated with increased sickness absence (Melchior 2003; Michie 2003). It is also possible that employees may interpret self-certification as an indication of the employer's trust, which can in turn result in a stronger commitment to work, decrease in sickness absence, and positive effect on social climate in the workplace (Pesonen 2016). Social climate can be defined as a shared, distinctive, and dynamic perception of social environment that can influence behaviour (Bennett 2010).

The requirement of a sickness certificate from a physician may be thought of as a means to control the use of sick leave because of the existence of a moral hazard. In economic terms, a moral hazard implies that individuals utilise sickness absence (or insurance) more than necessary to maximise their utility (Khan 2009). In fact, the moral hazard may apply especially to short sickness absence periods, because for longer periods it would be more evident that there is a medical cause for the absence. According to this logic, the requirement of a sickness certificate in the case of short-term sickness absence would be a strong intervention to prevent the moral hazard, and changing to a self-certification procedure would possibly lead to a moral hazard problem with an increase in short-term sickness absence periods. On the other hand, it is possible that introducing or extending self-certification will result in presenteeism.

An earlier investigation found that a self-certification practice improved communication between employees and supervisors (Pesonen 2016); good communication between employees and

supervisors can prevent or mitigate the possible moral hazard related to self-certification (Nieuwenhuijsen 2004). The supervisors found that the practice of self-certification was a useful tool in the management of well-being at work. On the other hand, Pesonen 2016 also found that the practice of self-certification increased supervisors' workload. However, self-certification decreased the use of resources in occupational health care and permitted the shift of resources to preventive work. This suggests that if healthcare providers perform less sickness certification, resources that can be used elsewhere, for example for preventive services, will be made available.

Why it is important to do this review

For short periods of sickness absence in particular, physician certification is an expensive process due to the administrative procedures involved. In addition, physician certification is associated with considerable costs and uptake of resources that could be better used in other healthcare areas. It is also known that sickness certification practices vary significantly between physicians (Arrelöv 2005; Englund 2000; Kankaanpää 2014). Literature reviews have concluded that physicians report difficulties with the certification process (Letrilliart 2012; Wynne-Jones 2010). However, as the role of general practitioners in sickness certification differs between countries, there may be variation regarding which aspects of the process are problematic (Winde 2012). Altogether, sickness certification involving a physician has been regarded as a burden to employees, the employer, the healthcare system, and society (Letrilliart 2012).

Sickness certification practices vary between countries, and self-certification is not a common practice in all countries. It is unclear whether and how sickness absence is altered when the requirements of a sickness certificate change. Self-certification will potentially decrease the total number of short-term sick leave days because there will be less administrative delay with the frequently occurring shorter periods, and employees can return to work as soon as they feel well enough to work. However, it is likely that self-certification will not influence the duration of long-term sick leave because the medical condition will be more serious, and assessment and documentation by a physician will be a natural part of healthcare procedures.

A randomised controlled trial conducted in Sweden found that postponing the requirement for a physician's certificate from day eight until day 15 increased the average duration of sickness absences by 6.6% (Hesselius 2005). A cost-benefit analysis showed that costs of increased sickness duration exceeded the benefits from fewer medical appointments. In contrast, a Norwegian study by Mykletun 2014 showed that when the sickness certification entitlement was transferred from physicians to employees (self-certification) for a period of one year, there was a significant decline in sickness absence. Recently, a Finnish case study of six companies indicated that changing from physician certification to a self-certification system did not affect the total number of sick leave days (Pesonen 2016). However, short-term absence days decreased slightly, and the researchers pointed out that there were other benefits, such as a more active role of the supervisor and a 20% to 40% decrease in physician consultations.

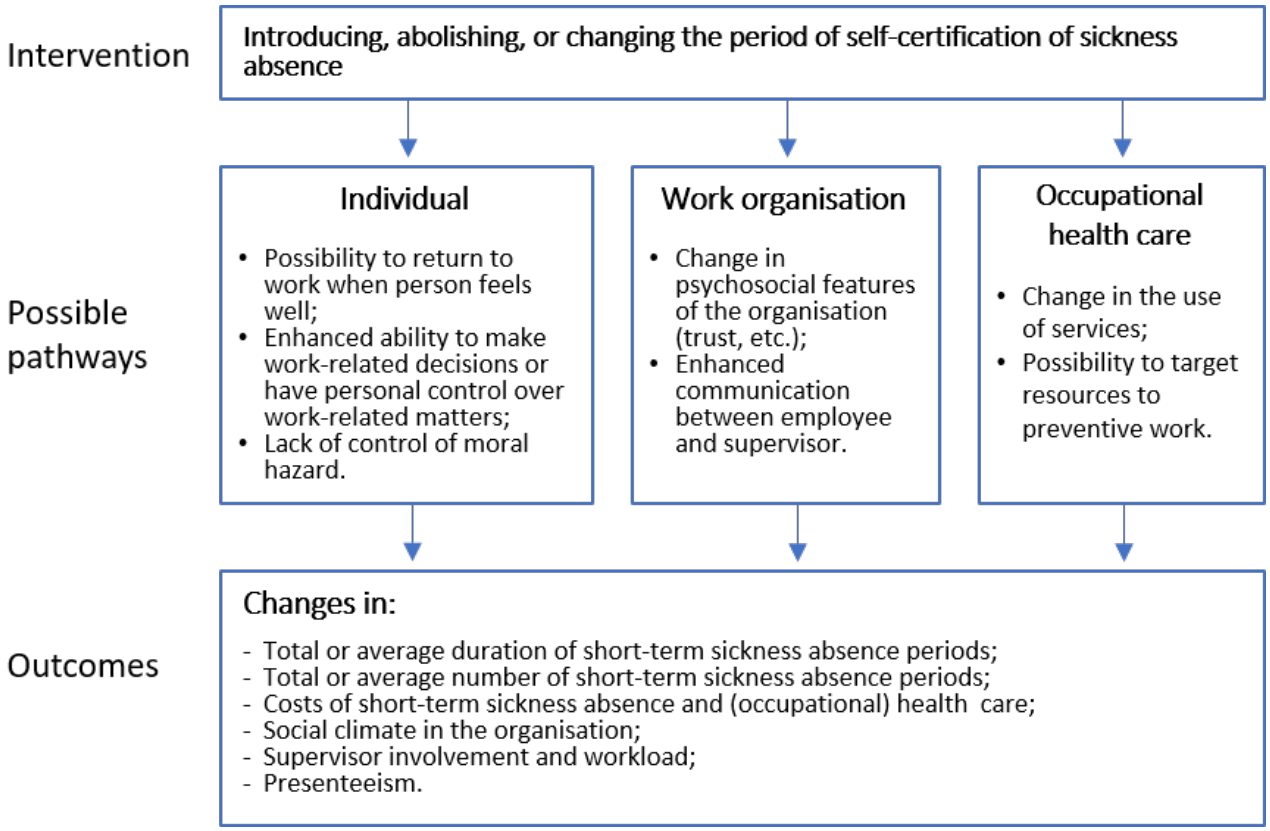
A literature review on rates of sickness certification suggested that the countries with the longest period of self-certification (e.g. the UK and Sweden, with seven days) reported high overall rates of sickness certification when compared to countries with shorter self-certification periods (e.g. Norway, with four days, and Switzerland, with three days) (Wynne-Jones 2008). However, other factors such as the existence of waiting days (i.e. the first days of sickness absence without pay, or the length of the period of sick pay) may also affect sickness absence rates. Due to the uncertainty surrounding the benefits and costs of self-certification, it was important to conduct a systematic review to combine the findings of existing studies. As far as we know, there are no systematic reviews on the effects of different sickness certification practices on sickness absence or related costs. Systematic reviews on sick leave and sickness certification practices are needed, as they can be used in preparation of evidence-based guidelines for health care.

OBJECTIVES

To evaluate the effects of introducing, abolishing, or changing the period of self-certification of sickness absence on:

1. the total or average duration (number of sickness absence days) of short-term sickness absence periods;
2. the frequency of short-term sickness absence periods;
3. the associated costs (of sickness absence and (occupational) health care); and
4. social climate, supervisor involvement, and workload or presenteeism (Figure 1).

Figure 1. Logic model of the intervention



METHODS

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs) where individual employees were randomly assigned either to a group that continued with the current practice of sickness certification (by a physician or by the employee his or herself) or to a group that started a new practice of sickness certification. The change could be from physician certification to self-certification (or vice versa) or to a longer or shorter period of self-certification. Legal and practical constraints will make randomisation difficult at the individual level, but it is conceivable that these constraints could be overcome by using a cluster-randomised design in which groups of workers or whole organisations are randomly assigned to the intervention or the control group, therefore we also included cluster-RCTs.

Because of the difficulty of performing randomised trials evaluating this intervention, we also considered the following observational study designs for inclusion: controlled before-after (CBA) studies (otherwise known as prospective cohort studies or quasi-experimental studies) and interrupted time-series (ITS) studies. Given the fluctuation in sick leave indicators over time, these study designs should be able to control for trends related to factors other than the intervention. Without such caution, it would be difficult to make inferences from the results of the studies.

Controlled before-after studies are easier to perform - considering that the intervention is carried out at the group level - while maintaining reasonable validity. We defined CBA studies as prospective or retrospective studies in which measurements of the outcome were available both before and after the implementation of the intervention for both the intervention and control group, and in which the outcome was measured at the same moment in time for both groups.

Interrupted time-series studies are studies with or without a control group in which the outcome has been measured at least three times before the intervention and at least three times after the intervention. The intervention is applied at a specific well-defined moment in time and is supposed to have an immediate effect, a long-term effect, or both. Because the outcome is measured several times before and after the intervention, it is possible to take time trends into account and thus make up for the lack of a control group (Ramsay 2003).

Types of participants

We included studies involving individual employees or insured workers. We also included studies in which participants were addressed at the aggregate level of organisations, companies, municipalities, healthcare settings, or general populations.

Types of interventions

We included studies evaluating the effects of introducing, abolishing, or changing the period of self-certification of sickness absence. We included any sickness certification practice in which the employee could report sick for a certain number of days without physician certification or certification by any other healthcare professional. Self-certification could be accepted for any disease or restricted to certain types of diseases.

We also included studies that combined self-certification with an intervention related to supervisor role or practices, working conditions (e.g. flexible working conditions), or terms of sickness benefit (e.g. number of waiting days), etc. (i.e. multicomponent interventions).

Types of outcome measures

Primary outcomes

1. The total or average duration (number of sickness absence days) of short-term sickness absence periods.
2. The total or average number of short-term sickness absence periods.

If self-certification was restricted to certain diseases, we also considered the outcomes only for those diseases.

We included any type of measurement of sickness absence, such as administrative data or self reported data.

We reported the results of economic evaluation studies related to introducing, abolishing, or changing the period of self-certification of sickness absence. We evaluated costs from the perspective of society, the employer, and the employee. The cost-related outcome measures were:

1. the costs of short-term sickness absence per employee (average within the organisation);
2. the total cost of short-term sickness absence;
3. the costs of (occupational) health care; and
4. changes in productivity.

We categorised the outcome measurements according to three follow-up times: up to three months, between three and six months, and six months or longer.

Secondary outcomes

We also included studies measuring a change in the social climate, supervisor involvement, and workload or presenteeism, regardless of whether they reported any of the primary outcomes.

Search methods for identification of studies

Electronic searches

We conducted a systematic literature search to identify all potentially eligible published and unpublished studies. We adapted the search strategy developed for MEDLINE for use in the other electronic databases. We imposed no restriction on language of publication. We arranged for the translation of key sections of potentially eligible non-English language papers or that such papers were fully assessed for inclusion by individuals proficient in the language of the publication as needed.

We searched the following electronic databases from inception to present for identifying potential studies.

1. Cochrane Central Register of Controlled Trials (CENTRAL) in the Cochrane Library (Appendix 1, Issue 4 of 12, April 2018).
2. MEDLINE Ovid (Appendix 2, 20 April, 2018).
3. SCOPUS (Appendix 3, 23 May, 2018).
4. PsycINFO Ovid (Appendix 4, 23 May, 2018).
5. EBM Reviews Ovid (Appendix 5, 14 June, 2018).
6. CINAHL EBSCOhost (Cumulative Index to Nursing and Allied Health Literature) (Appendix 6, 30 May, 2018).
7. EconLit ProQuest and EconPapers (Appendix 7, 14 June, 2018).
8. EBSCO Business Source Elite (EBSCOhost) (Appendix 8, 30, May 2018).
9. NIOSHTIC, NIOSHTIC-2, HSELINE, and CISDOC (OSH-UPDATE) (Appendix 9, 30 May, 2018).

We also searched for unpublished trials in the US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (www.clinicaltrials.gov/) on 14 June, 2018 and the World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (www.who.int/ictrp/en/) also on 14 June, 2018. We utilised Google Scholar for exploratory searches.

Searching other resources

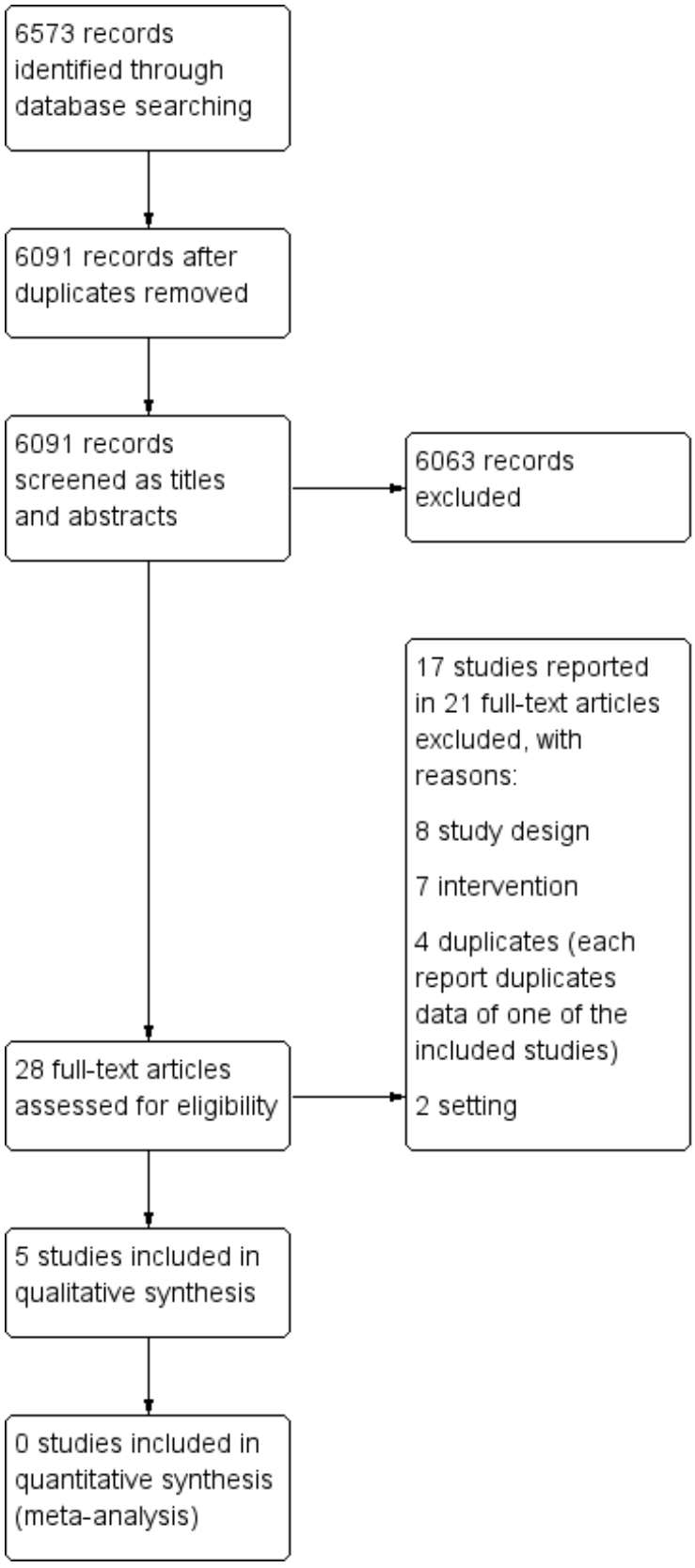
We contacted experts in the field to identify additional unpublished materials. We also searched websites of social security organisations such as the International Social Security Association (ISSA) for potentially eligible studies. We checked reference lists of the included studies for additional references.

Data collection and analysis

Selection of studies

We used the study screening tool Covidence to conduct the selection of eligible studies in two stages (Covidence). First, pairs of review authors (JK, JHV, JHR, JIH, LJV) independently screened titles and abstracts of all potentially relevant studies identified by our systematic search. The same authors coded them as 'include' (eligible or potentially eligible/unclear) or 'exclude'. At this stage we excluded all references that clearly did not fulfil our inclusion criteria or that did fulfil our exclusion criteria. Any differences of opinion regarding references coded 'include' were discussed until consensus was reached. At the second stage, we retrieved the full-text study reports or publications, and two review authors (JK and JIH) independently assessed the full-texts and identified studies for inclusion in the review, and recorded reasons for exclusion of the ineligible studies in the 'Characteristics of excluded studies' table. Any disagreements were resolved through discussion or by consulting a third review author (JHR) if necessary. We identified and excluded duplicates and collated multiple reports of the same study so that each study, rather than each report, is the unit of interest in the review. We recorded the selection process in sufficient detail to complete a PRISMA study flow diagram (see Figure 2).

Figure 2. PRISMA study flow diagram.



Had our systematic searches identified studies conducted by authors of this review, review authors who were not involved with the study would have made decisions involving inclusion and exclusion in order to avoid conflicts of interest.

When there was insufficient information about a study to reach a decision on eligibility, we sought to obtain further information from the authors.

We will rerun our systematic searches for trials every two years to enable updating of the review accordingly.

Data extraction and management

We used a data collection form that had been piloted on at least one study in the review to extract study characteristics and outcome data. Two review authors (JK and JIH) extracted the following study characteristics from the included studies.

1. Methods: study design, total duration of study, study location, study setting, withdrawals, and date of study.
2. Participants: N, mean age or age range, sex/gender distribution, occupation, type of work and branch of industry, and inclusion and exclusion criteria.
3. Interventions: description of intervention, comparison, duration, intensity, content of both intervention and control condition, and co-interventions.
4. Outcomes: description of primary and secondary outcomes specified and collected, and time points reported.
5. Notes: funding for trial and notable conflicts of interest of trial authors.

Two review authors (JK and JIH) independently extracted outcome data from the included studies. All data were reported in a useable way. Any disagreements were resolved by consensus or by involving a third review author (JHR). One review author (JK) transferred data into the Review Manager 5 file (RevMan 2014). We double-checked that data were entered correctly by comparing the data presented in the systematic review with that in the study reports. A second review author (JIH) spot-checked study characteristics for accuracy against the trial report. If in future updates of this review we should decide to include studies published in one or more languages in which our author team is not proficient, we will arrange for a native speaker or someone sufficiently qualified in each foreign language to fill in a data extraction form for us.

Assessment of risk of bias in included studies

Risk of bias in randomised controlled trials

Two review authors (JK and JHR) independently assessed risk of bias for each study using the Cochrane 'Risk of bias' assessment tool criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). Any disagreements were resolved by discussion or by involving another review author (JHV or JIH) if necessary. We assessed risk of bias according to the following domains.

1. Random sequence generation (selection bias).
2. Allocation concealment (selection bias).
3. Blinding of participants and personnel (performance bias).
4. Blinding of outcome assessment (detection bias).
5. Incomplete outcome data (attrition bias).
6. Selective outcome reporting.
7. Other bias.

We graded each potential source of bias as high, low, or unclear, and provided a quote from the study report together with a justification for our judgement in the 'Risk of bias' table. We summarised the 'Risk of bias' judgements across different studies for each of the domains listed. We considered blinding separately for self reported or administrative data of sickness absence: for self reported data, we considered the risk of bias due to blinding to be high, and for administrative data we considered it to be low. Where information on risk of bias related to unpublished data or correspondence with a trialist, we noted this in the 'Risk of bias' table. We contacted study authors in the case of insufficient information regarding risk of bias.

Risk of bias in controlled before-after studies

We used the Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) tool for assessing risk of bias in CBA studies (Sterne 2016). The seven domains of bias addressed in the ROBINS-I assessment tool are as follows.

1. Confounding.
2. Selection bias.
3. Bias in classification of interventions.
4. Bias due to deviations from intended interventions.
5. Bias due to missing data.
6. Bias in measurement of outcomes.
7. Bias in selection of the reported results.

We graded each potential source of bias as critical, serious, moderate, low, or no information. We summarised the risk of bias judgements across different studies for each of the domains listed. Our target trial against which we assessed the risk of bias was a trial in which participants that report sick were assigned to physician certification (or a shorter period of self-certification) and self-certification (or a longer period of self-certification) at the start of sick leave. We considered age, gender, and type of job (blue versus white collar) as potential confounders for which studies would have been expected to have adjusted in the design or in the analysis because these variables are related to a longer sickness absence or higher frequency of absence. We assessed waiting days or specific supervisor responsibility for sickness absence as co-interventions that could differ between intervention and control group and have an impact on the primary outcome. We first used the signalling questions as prescribed in the ROBINS-I tool and then assessed the risk of bias if these questions indicated a potential risk of bias.

Risk of bias in interrupted time-series studies

If we include ITS studies in future updates of this review, we will use the risk of bias criteria developed by the Cochrane Effective Practice and Organisation of Care Group (EPOC), which entails assessing the risk of bias in ITS studies against the following criteria.

1. Was the intervention independent of other changes?
2. Was the shape of the intervention effect prespecified?
3. Was the intervention unlikely to affect data collection?
4. Was knowledge of the allocated interventions adequately prevented during the study?
5. Were incomplete outcome data adequately addressed?

6. Was the study free from selective outcome reporting?
7. Was the study free from other risks of bias?

We will judge the risk of bias of ITS studies in all of the above domains to be high, low, or unclear. When one or more domains is judged to be at high risk of bias, the ITS study will be considered to be at high risk of bias.

We considered randomisation (confounding, selection), allocation concealment, and blinded outcome assessment to be key domains. We judged a study to have a high risk of bias overall when one or more key domains had a high risk of bias. Conversely, we judged a study to have a low risk of bias when all key domains had a low risk of bias and none of the other domains had a high risk of bias.

When considering treatment effects, we took into account the risk of bias for the studies that contributed to that outcome.

Assessment of bias in conducting the systematic review

We conducted the review according to the published protocol, Kausto 2018, and have reported any deviations from it in the Differences between protocol and review section of the review.

Measures of treatment effect

For RCTs and CBA studies, we calculated the results of each trial as point estimates, meaning means and standard deviations (SD) because all studies reported relevant outcome data using continuous outcomes. If in future updates of this review we identify studies reporting only effect estimates and their 95% confidence intervals or standard errors, we will enter these data into Review Manager 5 using the generic inverse-variance method (RevMan 2014).

As all results were reported using unambiguous continuous outcomes (number of sickness absence days, number of sickness absence periods, or work time lost due to sickness absence) with the same direction (less is better), there was no need to reverse any scales. When the results could not be plotted, we have described them in the 'Characteristics of included studies' table.

If we include ITS studies in future updates of this review, we will extract data from the original papers and re-analyse them according to recommended methods for analysis of ITS designs for inclusion in systematic reviews and also as recommended for evaluation of law studies (Viscusi 2005). These methods utilise a segmented time-series regression analysis to estimate the effect of an intervention while taking into account secular time trends and any autocorrelation between individual observations. If an included ITS study uses a control group, we will use the difference in rates between the intervention and the control group as the outcome. For each study, we will fit a first-order autoregressive time-series model to the data using a modification of the parameterisation of Ramsay 2003. Details of the model specification are as follows: $Y = \beta_0 + \beta_1 \text{time} + \beta_2 \text{(time-p) I (time > p)} + \beta_3 \text{I (time > p)} + E, E \sim N(0, s_2)$.

For time = 1,...,T, where p is the time of the start of the intervention, I (time ≥ p) is a function which takes the value 1 if time is p or later and zero otherwise, and where the errors E are assumed to follow a first-order autoregressive process (AR1). The β-parameters have the following interpretation: β1 is the pre-intervention slope. β2 is the difference between post- and pre-intervention slopes. β3 is the change in level at the beginning of the intervention period, meaning that it is the difference between the observed level at the first intervention time point and that predicted by the pre-intervention time trend.

We will standardise the data of ITS studies to obtain effect sizes by dividing the outcome and standard error by the pre-intervention SD as recommended by Ramsay 2003.

We will thus have two separate outcomes for an ITS study: the short-term change in the level of the outcome due to the intervention, which can be interpreted as an additive effect, and the long-term change in the trend in time or change of slope indicating an increasing effect of the intervention.

We had two primary outcomes: the mean number of short-term sickness absence periods (< 30 days) and the mean duration of short-term sickness absence periods. The number of short-term periods is a rate because one person can have several short-term sickness periods per year. A rate can also be analysed as continuous data. Because in some of the included studies the baseline rates differed between the intervention and the control group, we have used change values for the mean rate per person-year and calculated the intervention effect as the mean difference in the change values between the intervention and the control group: mean difference $(MD)_{change}$.

For the studies by Taylor 1969, Torsvik 2014, and Saksvik 2001, we calculated the duration of short-term sickness absence periods from the available data. For the duration of the periods we took the reported length that was nearest to the length of the self-certification period. If the self-certification period was up to three days, we extracted data on periods lasting up to three days. We extracted these data on the length of periods close to the length of the self-certification period because it is unlikely that self-certification will influence the length of a sickness period after a physician has certified the sickness absence.

The duration of short-term periods, in which we are interested, is the average duration of a period, calculated as the total of days lost due to short-term absence divided by the number of short-term absence periods. However, studies often reported “a percentage of work time lost of total work time”, which is equal to the average lost work time per worker. An illustrative example would be as follows. If 100 workers should have worked 239 days in a year, and short-term absence periods of up to three days would have accumulated a total lost time of 239 days, this would be reported as 1% lost work time (= 239 lost days/(100 workers*239 workdays)*100). The changes in average lost work time for a specific duration of absence can thus be due to either a change in the number of periods, a change in the average duration of the periods, or both. For the studies that reported percentages, we recalculated the average percentage in lost workdays because this is more intuitive. For this calculation we estimated that there would be 239 workdays in a year in Norway (first including 22 workdays per month during 12 months and then excluding 25 days of holiday). Because one worker may have more than one short sickness absence period during a given study, the average duration of an absence period lasting, for example, up to three days can be longer than three days. For the studies that measured average lost work time, the baseline values were often different in the intervention and control groups, therefore we calculated the change in the outcome from before the intervention to after the intervention for the intervention and

control groups, and used the difference in these change values as the effect of the intervention: MD_{change}.

Unit of analysis issues

If in future updates of this review we include studies that employ a cluster-randomised design and that report sufficient data to be included in the meta-analysis but that do not make an allowance for the design effect, we will calculate the design effect based on a large assumed intracluster correlation of 0.10. We base this assumption of 0.10 as being a realistic estimate by analogy to studies about implementation research. We will follow the methods described in Section 16.3.6 of the *Cochrane Handbook for Systematic Reviews of Interventions* for the calculations (Higgins 2011).

If several active interventions have been compared with no intervention, we will divide the participants in the no-intervention control group by the number of interventions, as recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* for the purposes of meta-analysis (Higgins 2011).

Dealing with missing data

We contacted investigators of all five included studies to verify key study characteristics and to obtain missing numerical outcome data where possible (e.g. when results were reported in figures only). If in future updates of this review we include studies for which we are not able to obtain this information, and we think that the missing data introduce serious bias, we will explore the impact of including such studies in the overall assessment of results by a sensitivity analysis.

If numerical outcome data such as SDs or correlation coefficients were missing and could not be obtained from the study authors, we calculated this information from other available statistics such as P values according to the methods in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Assessment of heterogeneity

We first assessed clinical homogeneity based on the similarity of the intervention, control condition, outcome, population, and follow-up time. We considered all types of participants to be similar. We considered any introduction, abolishment, increase, or decrease in the self-certification period to be similar. We considered all sick leave outcomes as similar (i.e. no restrictions due to diagnosis), although they were measured and reported differently.

Provided sufficient data are included in future updates of this review, we will also test for statistical heterogeneity by means of the Chi^2 test as implemented in the forest plots in Review Manager 5 (RevMan 2014). We will use a significance level of $P < 0.10$ to indicate if there is a problem with heterogeneity. In addition, we will quantify the degree of heterogeneity using the I^2 statistic (where $I^2 > 50\%$ indicates a moderate degree of heterogeneity and $I^2 > 75\%$ indicates a high degree of heterogeneity). When we identify ≥ 75% heterogeneity between studies we will refrain from pooling their results for meta-analysis.

Assessment of reporting biases

We reduced the effects of reporting bias by including studies and not articles. We prevented location bias by searching multiple databases. We prevented language bias by not using any language restrictions. We checked for outcome reporting bias as part of the risk of bias assessment.

Data synthesis

We have presented results separately for RCTs and CBA studies.

Provided sufficient data are included in future updates of this review, we will pool data from studies judged to be clinically homogeneous using Review Manager 5 software (RevMan 2014). If possible, we will combine studies using MDs. If different outcomes do not permit pooling, then we will use standardised mean differences (SMDs). To make the SMDs more readily interpretable for clinicians, we will then recalculate the pooled SMD into an MD by multiplying the SMD by the median SD taken from the included studies using the preferred scale in question. We will pool results from studies with different study designs if the direction and the magnitude of the studies are considered similar.

We will use the generic inverse-variance method as implemented in Review Manager 5 to combine hazard ratios and effect sizes obtained from ITS studies (RevMan 2014).

Given the type of interventions and the conditions under which trials would have been conducted, we expected statistical heterogeneity, therefore we planned to use a random-effects model for meta-analysis. If heterogeneity between studies is low, the results will be similar to those from a fixed-effect model. All estimates will include a 95% confidence interval.

We expected that not all included studies would contain an economic evaluation.

'Summary of findings' table

We have presented a 'Summary of findings' table using all primary and secondary outcomes: mean duration of short-term sickness absence periods, mean number of short-term sickness absence periods, and average lost work time per 100 person-years for our main comparison between longer self-certification and shorter self-certification for reducing sickness absence in workers.

We used the five GRADE considerations (study limitations, consistency of effect, imprecision, indirectness, and publication bias) to assess the quality of a body of evidence as it relates to the studies which contributed data to the prespecified outcomes. We used the methods and recommendations in Chapters 8 and 12 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2017; Schünemann 2017). We intended to use GRADEpro GDT software (GRADEpro GDT 2015) but there were no studies that could be combined and we reported the results in a narrative summary of findings table. We justified all decisions to down- or upgrade the quality of evidence using footnotes.

Subgroup analysis and investigation of heterogeneity

Provided sufficient data are available in future updates of this review, we will compare the effects of interventions on the following three variables in subgroup analyses.

1. The length of increase or decrease in the self-certification period, where we will compare the effects of studies with participants that have up to seven calendar days versus more than seven calendar days of self-certification.

2. The existence of so-called waiting days, i.e. days at the beginning of sickness absence when sick pay is not payable, where we will compare the effects of studies with participants that do not have waiting days versus studies with any number of waiting days.
3. The existence of flexible working conditions such that we will compare studies conducted with participants who have flexible working conditions versus those without flexible working conditions.

We will use the Chi^2 test to test for subgroup interactions in Review Manager 5 (RevMan 2014).

Sensitivity analysis

Provided sufficient data are available in future updates of this review, we will conduct sensitivity analyses to test the robustness of our meta-analysis results by leaving out studies judged to be at high risk of bias. We will also conduct sensitivity analyses for possible assumptions made for missing data or analyses during the review process.

Reaching conclusions

We will base our conclusions only on findings from the quantitative or narrative syntheses of the studies included in this review. We will avoid making recommendations for practice based on anything other than the evidence. We will suggest priorities for future research and outline what the remaining uncertainties are in the area.

RESULTS

Description of studies

Results of the search

We identified a total of 6573 references from CENTRAL (50), MEDLINE (918), SCOPUS (1639), PsycINFO (228), EBM Reviews (43), CINAHL (620), EconLit (710), EBSCO Business Source Elite (1794), EconPapers (71), OSH-UPDATE (417), ClinicalTrials.gov (56), and the WHO ICTRP (27). We identified no references from reference lists of potentially included studies, contact with experts, and searching the ISSA website. After excluding 482 duplicates, we screened the titles and abstracts of 6091 references for potential inclusion in the review. We obtained the full-texts of 28 articles on 17 studies, of which we excluded 21 articles because they did not meet our inclusion criteria for the following reasons: eight articles (study design), seven (intervention), four (duplicates), and two (study setting) (see Characteristics of excluded studies). We included five studies described in five articles in the review. A PRISMA study flow diagram depicting the results of the search and the process of screening and selecting studies is shown in Figure 2.

Included studies

Study designs

We found only one RCT, by Hesselius 2005. The other four included studies used a mixture of study designs that fit our definition of CBA studies (Fleten 2009; Saksvik 2001; Taylor 1969; Torsvik 2014).

Settings

Interventions were carried out among municipality or company employees or insurance claimants in municipalities. Of the five included studies, three were from Norway (Fleten 2009; Saksvik 2001; Torsvik 2014), one was from Sweden (Hesselius 2005), and one was from the UK (Taylor 1969).

Participants

Participants were employees or insurance claimants studied at the aggregate level.

Interventions

In Taylor 1969, the intervention was self-certification introduced to employees for absence periods lasting up to three working days. Upon return to work, employees had to complete a sickness report for the personnel record office. The comparator was physician certification of sickness absence. In the other four studies, the intervention was prolonging the period of self-certification of sickness absence. The comparison was between a longer and a shorter period of self certified sickness absence. In the study by Saksvik 2001, municipality employees were entitled to have one to three self administered sick leave days four times per year at baseline, while the intervention group was entitled to have one to five self administered sick leave days four times per year. In the study by Hesselius 2005, the intervention group was allowed to receive sickness benefits for two weeks without showing a doctor's certificate, while the control group continued the practice-as-usual of being entitled to the benefits without showing a doctor's certificate for one week. In the study by Fleten 2009, the intervention consisted of prolonging the period of self-certification of sickness absence up to 50 days per year, divided into one to 10 periods with a structured workplace follow-up. In the study by Torsvik 2014, the comparison was between self-certification up to one year (intervention) and practice-as-usual, that is a medical certificate was needed either on the third or the eighth day of sickness absence.

Outcomes

The outcomes in each of the included studies were differently measured. Taylor 1969 reported percentage of change in total sickness absence from the year before the experiment to the year following the experiment, proportion of sickness absence periods by duration before and after the experiment, and number of sickness absence periods by duration after the experiment for the intervention and control group. The authors of Saksvik 2001 reported percentages of time lost of total possible work time for self reported sickness absences of one to three days' duration at three time points for both the intervention and control groups. However, the authors did not calculate changes between baseline and the end of follow-up. The authors of Hesselius 2005 reported the length of sickness absence periods for the intervention and control groups by survival analysis (return-to-work curves). They also reported risk ratios between the intervention and control groups at two time points. In Fleten 2009, the authors reported mean durations of sickness absence periods of 1 to 3 days', 4 to 7 days', and 4 to 16 days' duration at several time points before and after the intervention. They reported control group results only partly for sickness absence periods of one to three days. In Torsvik 2014, the authors reported the average absence percentage of contracted workdays and the average length of sickness absence periods for the intervention and control groups before and after the experiment. All included studies had a follow-up time of longer than six months.

Excluded studies

We excluded a total of 17 studies reported in 21 full-text articles. Four articles were duplicates of Hesselius 2005, Torsvik 2014, Preece 2006, and Saksvik 2001. The main reasons for excluding studies were as follows.

1. Study design (Carne 1969; Felder 2008; Gjesdal 2005; Hauge 2017; Herrmann 2015; Ihlebaek 2006; Money 2013; Royneland 2002).
2. Intervention (De Paola 2014; Kaufmann 2010; Pertold 2018; Pettersson-Lidbom 2013; Pollak 2017; Schlotzhauer 1985; Voss 2001).
3. Study setting (Oyeflaten 2011; Preece 2006).

Risk of bias in included studies

Two review authors (JK and JIH) independently assessed the risk of bias using the Cochrane 'Risk of bias' tool criteria outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* for the one included RCT (Higgins 2011). We used the ROBINS-I tool to assess risk of bias for the four included CBA studies (Sterne 2016). The results are summarised in Figure 3, which shows the review authors' judgements about each risk of bias item for each included study, and in Figure 4, which shows the review authors' judgements about each risk of bias item presented as percentages across all included studies.

Figure 3. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

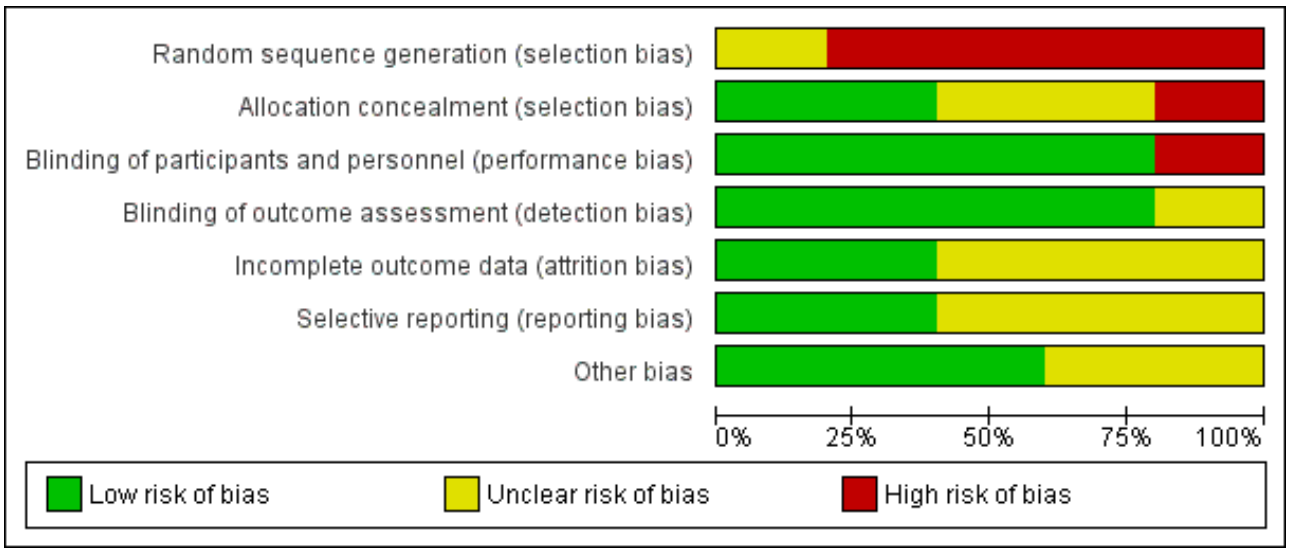
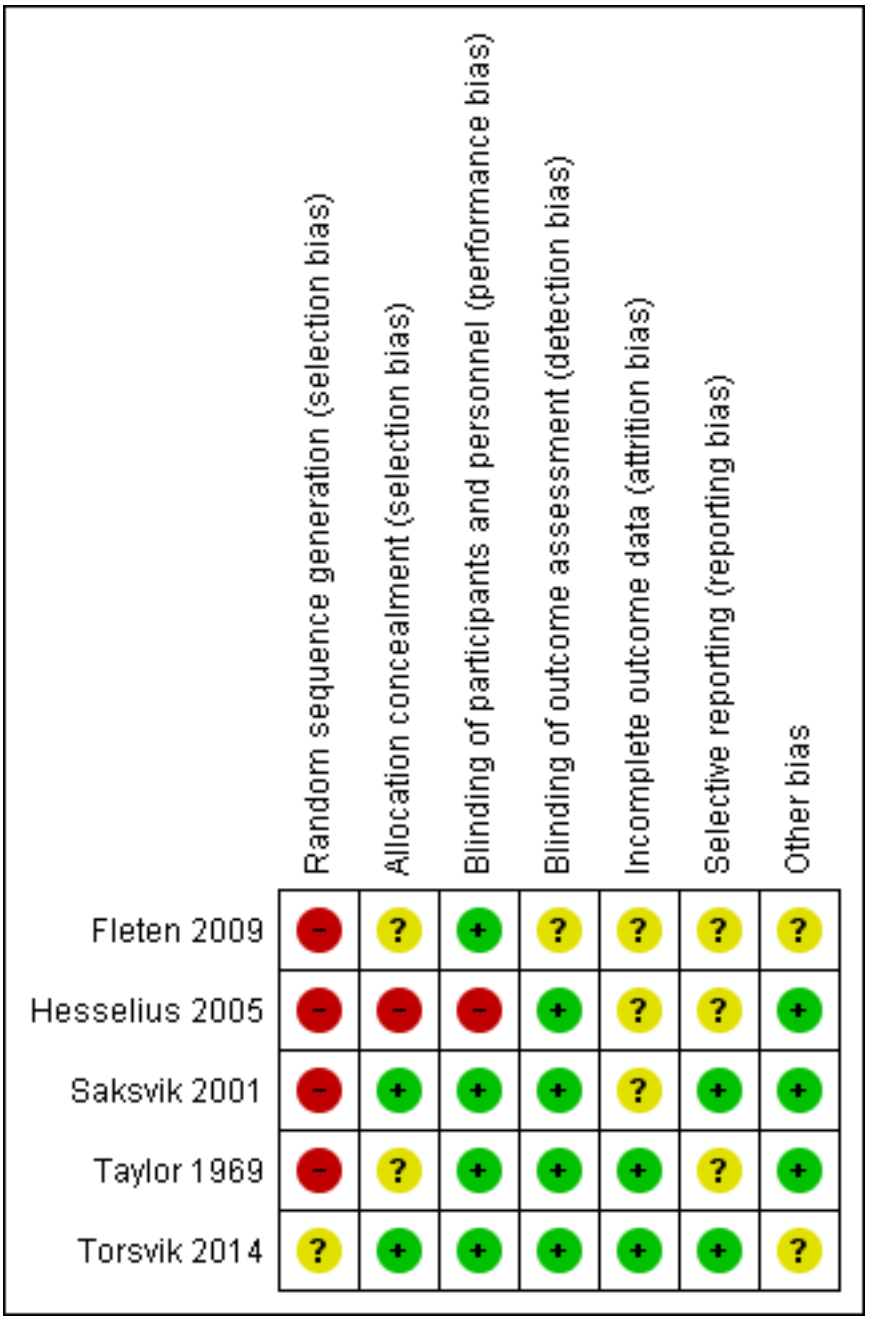


Figure 4. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Allocation

We judged the single included RCT by Hesselius 2005 to have a high risk of selection bias because the authors randomised participants based on their date of birth, which does not guarantee true randomisation, and because they did not report if or how allocation was concealed between randomisation and the start of the study.

There was no information regarding selection bias in Fleten 2009 and Taylor 1969, therefore we judged the risk of allocation bias to be unclear. In Fleten 2009, the selection of participants was based on municipalities and not on participant attributes. The authors also did not report any baseline descriptive information regarding the intervention and control groups. In Taylor 1969, the intervention group comprised all manual labour workers from an oil company, and the control group comprised all staff members of the same company. The authors did not report any baseline descriptive information regarding the intervention and control groups. We judged Saksvik 2001 and Torsvik 2014 to be at low risk of selection bias. In Saksvik 2001, employment in a certain district of a municipality was the selection method for both intervention and control groups. In Torsvik 2014, participants were allocated to intervention and control groups based on the municipalities in which they were employed.

Blinding

We judged the Hesselius 2005 RCT to have a high risk of performance bias because the participants were not blinded, and a low risk of detection bias because the authors obtained outcome data from a register.

Incomplete outcome data

We judged the Hesselius 2005 RCT to have an unclear risk of attrition bias because the authors did not provide any information on the completeness of outcome data.

Selective reporting

We judged the Hesselius 2005 RCT to have an unclear risk of reporting bias because the authors did not report a plan for their analyses, and they reported their findings based on total incidence and prevalence.

In the CBA studies by Taylor 1969, Fleten 2009, and Torsvik 2014, there was no information on selective reporting (no protocol or description of methods), thus we judged reporting bias to be unclear. We judged the Saksvik 2001 CBA study to have a low risk of reporting bias because the authors carried out and reported all planned analyses.

Other potential sources of bias

We judged the Hesselius 2005 RCT to have a low risk of other bias because we found no evidence of other major issues that could have introduced bias.

No studies reported waiting days or specific supervisor responsibility for sickness absence as co-interventions.

Additional sources of bias according to ROBINS-I items

Bias due to confounding

We judged three of the four CBA studies to be at serious risk of bias due to confounding (Fleten 2009; Saksvik 2001; Taylor 1969). Fleten 2009 only conducted a t-test and did not include any confounders in their analyses. Saksvik 2001 used multivariate analysis of variance (MANOVA) to compare means but did not include any confounders other than age in the models. Most importantly, the authors did not adjust for sex and job type. They did not provide information on co-interventions. Taylor 1969 considered only age in his analyses. We judged the fourth CBA study by Torsvik 2014 to be at moderate risk of bias due to confounding because the authors used a balanced panel of municipalities (Mandal versus similar municipalities) based on size and economic variables. However, the authors did not report whether they had included confounding variables in their analyses.

Bias due to missing data

In three of the four CBA studies (Fleten 2009; Saksvik 2001; Taylor 1969), there was no information to base a judgement on bias due to missing data, therefore we judged the risk of bias for this item to be unclear. We judged Torsvik 2014 to have a low risk of bias due to missing data because the authors obtained outcome data for the intervention and control groups from the same national register.

Bias in classification of interventions

We judged all four CBA studies to have a low risk of bias in classification of interventions (Fleten 2009; Saksvik 2001; Taylor 1969; Torsvik 2014). Allocation to either the intervention or the control group was based on employment in a certain municipality or in a certain district of a municipality or occupational status.

Bias due to deviations from intended interventions

We judged three CBA studies to be at low risk of bias due to deviations from intended interventions as there was no reason to believe that the intervention had changed over the follow-up period (Saksvik 2001; Taylor 1969; Torsvik 2014). In Fleten 2009 there was not enough information to base a judgement on deviations from intended interventions, therefore we judged the risk of bias for this item to be unclear.

Bias in measurement of outcomes

We judged Saksvik 2001, Taylor 1969, and Torsvik 2014 to be at low risk of bias in measurement of outcomes as the authors had obtained outcome data from registers. In Fleten 2009 no information was provided to base a judgement because the authors collected outcome data from the control group only when sickness absence lasted longer than three days, therefore we judged the risk of bias in measurement of outcomes to be unclear.

Effects of interventions

See: **Summary of findings for the main comparison** Longer self-certification compared to shorter self-certification for reducing sickness absence in workers

1. Longer self-certification compared to shorter self-certification

1.1 The total or average duration (number of sickness absence days) of short-term sickness absence periods

1.1.1 Evidence from RCTs

The RCT by Hesselius 2005 compared prolonging the period of self-certification from one week to two weeks in two municipalities. The authors presented return-to-work curves from survival analysis. Following the intervention, the average duration of sickness absence periods lasting up to 29 days was 8.91 days (SD 14.6) in the intervention group and 8.24 days (SD 13.9) in the control group (data presented in Hesselius 2009). Based on the reported data, we calculated a mean difference $(MD)_{change}$ between intervention and control group in duration of absence periods lasting less than 29 days. The result shows self-certification producing a higher mean duration of sickness absence, thus favouring physician certification after one week (MD_{change} 0.67 days/period, 95% confidence interval (CI) 0.55 to 0.79; Analysis 1.1).

1.1.2 Evidence from CBA studies

We identified one CBA study that compared duration of sickness absence periods after the introduction of self-certification for a maximum of three days to duration of sickness absence periods before the introduction of self-certification (Taylor 1969). Based on the reported data, we calculated an MD in the change from before to after intervention between intervention and control groups in the duration of sickness absence periods lasting up to three days. The result shows that self-certification produced a lower mean duration of sickness absence, thus favouring self-certification (MD_{change} −0.32 days/period, 95% CI −0.39 to −0.25; Analysis 1.1).

The CBA study by Torsvik 2014 compared prolonging the self-certification period from three days to up to one year (365 days). The authors did not provide numerical values for the length of the absence periods, but wrote that there was no change in the mean duration of sickness absence periods.

1.2 The total or average number of short-term sickness absence periods

1.2.1 Evidence from RCTs

In the RCT by Hesselius 2005, a self-certification period of ≤ 14 days led to no difference in number of sickness absence periods compared to a self-certification period of ≤ 7 days, but the authors did not report numerical data.

1.2.2 Evidence from CBA studies

Self-certification produced a higher mean number of sickness absence periods, thus favouring physician certification (MD_{change} 0.48 periods, 95% CI 0.33 to 0.63; Analysis 1.2) in one older study (Taylor 1969). In order to estimate standard errors (SEs) for the work by Taylor 1969, we applied a distribution for the number of one- to five-day sickness absence periods per person taken from the Finnish Helsinki Health Study conducted in 2000-2002 (Lahelma 2013). We used the following percentages for the number of periods: zero periods 45%; one period 25%; two periods 13%; three periods 10%; four periods 5%; and five periods 2%. This enabled us to calculate the SD for the MD_{change} between intervention and control groups as: √((((0.25*N periods*(1 − mean N periods/worker)^2) + (0.13*N periods*(2 − mean N periods/worker)^2) + (0.10*N periods)*(3 − mean N periods)^2) + (0.05*N periods*(4 − mean N periods)^2) + (0.02*N periods*(5 − mean N periods)))/(N − 1))).

The authors of the Torsvik 2014 CBA study did not provide numerical values for the number of absence periods, but wrote that there was no change in the mean number of sickness absence periods in the intervention group.

1.3 Lost work time due to sickness absence periods per 100 person-years

1.3.1 Evidence from RCTs

Because there was no difference in number of sickness absence periods between intervention and control groups in the Hesselius 2005 RCT, longer sickness absence periods will have led to an increase in lost work time in the intervention group, thus favouring physician certification.

1.3.2 Evidence from CBA studies

For the CBA study by Saksvik 2001, we calculated first SE and then the SD for the mean change in sickness absence periods of one to three days using the 95% CI from the other Norwegian CBA study by Fleten 2009 (Table 1: 95% CI 1.78 to 1.82). We thus calculated SE as: (1.82−1.78)/1.96, and SD as: SE*√N.

Based on the data reported in Saksvik 2001 and Taylor 1969, we were able to calculate MDs between change in intervention and control groups in lost work time due to sickness absence periods shorter than three days. In both cases, self-certification resulted in more work time lost due to sickness absence periods lasting less than three days, thus favouring physician certification (Saksvik 2001: MD_{change} 1.38, 95% CI 1.16 to 1.60; and Taylor 1969: MD_{change} 0.54, 95% CI 0.47 to 0.61; Analysis 1.3). In Saksvik 2001, the MD_{change} for absence periods of 4 to 16 days was 0.1 days (95% CI could not be calculated).

The CBA study by Fleten 2009 compared the effects of prolonging the period of self-certification from one to three days up to 50 days per year divided into one to 10 periods per year. For this study, we calculated SDs based on available data in the report as: SE*√N. Following the intervention, average lost work time due to absence periods lasting between four to 16 days decreased from 8.47 days to 7.82 days in the intervention group; the duration remained stable in the control group, but more detailed information on this was not provided. The change in lost work time was −0.65 days (when change in the control group was assumed to be 0 days), thus favouring self-certification. This effect may be attributable to a reduction in the number of sickness absence periods or to a reduction in their average length.

In the CBA study by Torsvik 2014, the authors compared self-certification up to one year and practice-as-usual (a medical certificate was needed either on the third or the eighth day of sickness absence). The authors presented the absenteeism data as mean per cent of sickness absence of all contracted working days per year. They included only sickness absence periods lasting 16 days or more in their analyses. From the percentages of lost total possible work time provided, we calculated an MD between change in the intervention and control groups in lost work time due to sickness periods longer than 16 days. We calculated SDs based on the SE and number of participants as reported by the authors as: SE*√N. The result shows self-certification resulting in less work time lost due to sickness absence periods longer than 16 days, thus favouring self-certification (MD_{change} −2.84 days, 95% CI −3.35 to −2.33; Analysis 1.3).

1.4 Cost of sickness absence

1.4.1 Evidence from RCTs

The authors of Hesselius 2005 valued one day of sickness absence by using the average daily compensation (SEK 296) paid by employers to employees. At the time of conducting the intervention in 1988 this covered 90% to 100% of wages. In the one-week self-certification group, 20.6% of participants had sickness absences longer than seven days, whereas in the two-week self-certification group 11.6% had sickness absences longer than 14 days. The increase in the length of the self-certification period therefore led to a reduction in physician certifications of nine percentage points. The authors presented the cost of a physician's appointment as SEK 445. They calculated that the higher costs of sickness absence benefits (SEK 29,000,000) far outweighed the possible reduction in costs of fewer physician appointments in the Gothenburg area (SEK 4,900,000).

1.4.2 Evidence from CBA studies

We found no CBA studies using this outcome measure.

DISCUSSION

Summary of main results

We included five studies (one RCT and four CBA studies) evaluating the effects of introduction of or a change in the length of self-certification of sickness absence. One study from the 1960s evaluated the effect of introducing self-certification, and the other four, later, studies evaluated the effect of prolonging the self-certification period. The included studies measured the effects of the intervention on the number of sickness absence periods, the number of sickness absence days per sickness absence period, or on lost work time due to sickness absence periods. The results were inconsistent. See Summary of findings for the main comparison for an overview of the results.

Effects on average duration of sickness absence periods

There is low-certainty evidence from one RCT by Hesselius 2005 showing that extending the period of self-certification from one week to two weeks increased the mean duration of sickness absence. There is very low-certainty evidence from one old CBA study by Taylor 1969 showing that the introduction of self-certification for a maximum of three days reduced the mean duration of sickness absence. There is very low-certainty evidence from one CBA study by Torsvik 2014 showing that extending the self-certification period from one to three days up to one year did not lead to a change in the mean duration of sickness absence. The authors of Torsvik 2014 did not report data to support this finding.

Effects on number of sickness absence periods per worker

There is low-certainty evidence from one RCT by Hesselius 2005 that extending the period of self-certification from one week to two weeks resulted in no difference in number of sickness absence periods. Unfortunately, the authors did not report data to support this finding. There is very low-certainty evidence from one old CBA study by Taylor 1969 showing that the introduction of self-certification for a maximum of three days increased the mean number of sickness absence periods. There is very low-certainty evidence from another CBA study by Torsvik 2014 showing that extending the period of self-certification from three days to up to a year did not change the number of sickness absence periods. The authors of Torsvik 2014 did not report data to support this finding.

Effects on average lost work time per 100 person-years

There is low-certainty evidence from one RCT by Hesselius 2005 showing that extending the period of self-certification from one week to two weeks resulted in an inferred increase in lost work time. There is very low-certainty evidence from two CBA studies by Saksvik 2001 and Taylor 1969 from different decades with very different data that could not be pooled, which showed that extending the period of self-certification (from zero days to three days in Taylor 1969 and from three days to five days in Saksvik 2001) increased the amount of work time lost due to sickness absence periods lasting less than three days. There is very low-certainty evidence from two CBA studies by Fleten 2009 and Torsvik 2014 showing that extending the period of self-certification (from between one to three days up to 50 days per year divided into one to 10 periods per year in Fleten 2009, and from between three to eight days up to one year in Torsvik 2014) reduced the amount of work time lost due to sickness absence periods lasting four to 16 days and 16 days or longer, respectively.

Costs of sickness absence and physician certification

There is low-certainty evidence from one RCT by Hesselius 2005 showing that the costs of sickness absence benefits resulting from a longer period of self-certification were about six times the possible reduction in costs of fewer physician appointments. Hesselius 2005 used a valuation of SEK 445 for a physician appointment, which is extremely high compared to the average compensation per day received by the participating workers (SEK 296, which compensated 90% to 100% of lost wages).

However, the authors provided no information on the costs of potential laboratory tests, X-rays, prescriptions, control appointments, or other costs related to the required physician appointment, nor could we make a reliable assumption of these costs. Consequently, our cost-benefit calculation is far from ideal.

In summary, we found very low-certainty evidence that introducing self-certification of sickness absence or prolonging the self-certification period has inconsistent effects on the mean number of sickness absence days, the number of sickness absence periods, and the amount of lost work time due to sickness absence periods.

We found no studies on the effects of self-certification on social climate, supervisor involvement, workload or presenteeism.

Overall completeness and applicability of evidence

Only five studies met our inclusion criteria. Three of the five studies were carried out in Norway, one in Sweden, and one in the UK. Differences in disability policies, social security systems and working life, and cultures between countries might reduce the generalisability of the findings to other countries with different social contexts. Two of the studies were rather old. Taylor 1969 introduced self-certification in the 1960s, and Hesselius 2005 reported a trial conducted in the 1980s, when working life and culture were quite different from how they are currently. This is important because self-certification is assumed to have an effect on behaviour, which is also dependent on cultural factors.

Quality of the evidence

We judged the single included RCT to have a high risk of bias overall because it had a high risk of bias for sequence generation, allocation concealment, and blinding of participants and personnel. The study used alternation to allocate participants to the intervention or control group, which yields a predictable sequence. We downgraded the certainty of evidence from the single RCT by two levels, that is from high to low, due to risk of bias. We downgraded the evidence from this study on lost work time to very low due to indirectness because we inferred the result from two other outcomes.

We downgraded the certainty of the evidence from the CBA studies from low to very low due to risk of bias because the authors did not take into account critical confounders. . Given the small number of included studies, we could not assess the possible influence of publication bias.

Potential biases in the review process

We carried out systematic searches in the most important electronic databases. We also conducted a search for unpublished trials. We attempted to minimise selection bias by using no time or language restrictions. While the possibility remains that we have

missed relevant studies, we think this is unlikely. Due to the lack of reporting details in the studies, and because we could not obtain additional data from study authors, we had to make assumptions in order to calculate effect sizes. In order to calculate SDs of the number of sickness absence periods per worker for Taylor 1969, we had to assume a distribution of the number of short-term sickness absence periods per person. We took this information from a more recent Finnish study (Lahelma 2013). As this might not have been appropriate for a much older British study, we performed a sensitivity analysis. We found that with a distribution of almost all workers reporting sick and more frequently, the SDs doubled, but the MD was still significantly different from zero, therefore we believe that these assumptions have not biased our results.

Self-certification is not the only factor that affects the duration and frequency of sickness absence. These are also influenced by such factors as business cycle (Pichler 2015), industry and occupation (Leinonen 2018), as well as healthcare regulations (e.g. waiting times, pay schemes). In this Cochrane Review we focused on changes in self-certification and subsequent changes in sickness absence. We assume that the context (e.g. business cycle, industry and occupation, compensation level, and waiting times) was very similar for the intervention and control groups.

Agreements and disagreements with other studies or reviews

We are not aware of other systematic reviews on the topic we have addressed in this review. Our conclusions differ from those made by the authors of the studies included in our review. Some studies reported that self-certification of sickness absence reduced the length of sickness absence or the number of sickness absence periods, whereas others reported increases.

AUTHORS' CONCLUSIONS

Implications for practice

Overall, we found low- to very low-certainty evidence of inconsistent effects of self-certification of sickness absence on the examined outcomes.

Implications for research

Given the inconsistency and very low-certainty of the evidence identified, and the widespread practice of self-certification of sickness absence, further studies are needed.

As the example of Hesselius 2005 shows, conducting randomised trials in this area is possible. To prevent confounding and selection bias, studies should employ randomisation. The preferred study design is cluster-randomisation to prevent 'contamination' of intervention and control workers, as reported by Hesselius 2013. If a non-randomised design is used, participants in the intervention and control groups should have similar age and sex distributions and work in the same industrial sector, and have similar sickness absence compensation schemes (e.g. similar waiting-day practices). Study authors should reportthe number of participants in both the intervention and control groups and their age, sex,occupation, level of education, economic circumstance, disease status, and health behaviours, such as smoking, so as to judge the representativeness of the sample.

When the intervention consists of prolonged self-certification periods, authors should report if this also implies restrictions on the use of occupational health services for these shorter absence periods. Control groups should include participants for whom there is no change in self-certification practice until the end of intervention follow-up.

Authors should report the effects on the number of short-term sickness absence periods per worker, the duration of these periods, and the average lost work time. Outcomes must be measured similarly in both the intervention and control groups, which could be realised by including participants working for the same employer. Adverse outcomes such as overdiagnosis, overtreatment, and medicalisation of mild health symptoms and psychosocial problems should be considered if a physician's certificate is needed from the first day of absence.

For policymakers, the main implication from this Cochrane Review is that changes to current systems of sickness certification must be evaluated with a controlled study design to increase the evidence base on the effects of sickness certification.

ACKNOWLEDGEMENTS

We thank Katri Larmo at Helsinki University Library as well as Heikki Laitinen and Kaisa Hartikainen at University of Eastern Finland Library for developing, testing, and running the search strategies.

We also thank Cochrane Work Editors Riitta Sauni and Alex Burdorf and external peer referees Kari-Pekka Martimo, Gwenllian Wynne-Jones, and Evelyn Aaviksoo for their comments on the review. We also thank Lisa Winer for copy editing the text of the review.

REFERENCES

References to studies included in this review

Fleten 2009 *{published data only}*

Fleten N, Krane L, Johnsen R. Extended self-certification - a step towards more appropriate sickness absence?. *Norsk Epidemiologi* 2009;**19**(2):223-8.

Hesselius 2005 *{published data only}*

Hesselius P, Johansson P, Larsson L. Monitoring sickness insurance claimants: evidence from a social experiment. Upsala:IFAU - Institute for Evaluation of Labour Market and Education Policy; 2005. Working Paper Series 2005:15,.

Hesselius P, Johansson P, Vikström J. Is the individual's sickness absence affected by the sickness absence of the surroundings? [Påverkas individen av omgivningens sjukfrånvaro?]. *Ekonomisk Debatt* 2008;**7**:44-53. [www.nationalekonomi.se/filer/pdf/36-7-phpjjv.pdf]

Hesselius P, Johansson P, Vikström J. Social behaviour in work absence. *Scandinavian Journal of Economics* 2013;**115**(4):995-1019.

Hesselius P, Nilsson JP, Johansson P. Sick of your colleagues' absence?. *Journal of the European Economic Association* 2009;**7**(2-3):583–94. [DOI: 10.1162/JEEA.2009.7.2-3.583]

Saksvik 2001 *{published data only}*

Saksvik PO, Nytrø K. Improving subjective health and reducing absenteeism in a natural work life-intervention. *Scandinavian Journal of Psychology* 2001;**42**(1):17-24. [DOI: 10.1111/1467-9450.00210]

Taylor 1969 *{published data only}*

Taylor PJ. Self-certification for brief spells of sickness absence. *British Medical Journal* 1969;**1**(5637):144-7.

Torsvik 2014 *{published data only}*

Torsvik G, Vaage K. Gatekeeping versus monitoring: evidence from a case with extended self-reporting of sickness absence. Munich: CESifo Group; 2014. CESifo Working Paper Series 5113,.

References to studies excluded from this review

Carne 1969 *{published data only}*

Carne S. Sick absence certification. Analysis of one group practice in 1967. *British Medical Journal* 1969;**1**(5637):147-9.

De Paola 2014 *{published data only}*

De Paola M, Scoppa V, Pupo V. Absenteeism in the Italian public sector: the effects of changes in sick leave policy. *Journal of Labor Economics* 2014;**32**(2):337-60.

Felder 2008 *{published data only}*

Felder S. To wait or to pay for medical treatment? Restraining ex-post moral hazard in health insurance. *Journal of Health Economics* 2008;**27**(6):1418-22. [DOI: 10.1016/j.jhealeco.2008.06.002]

Gjesdal 2005 *{published data only}*

Gjesdal S. Sickness absence incidence in Norway 1975-2002. *Tidsskrift for den Norske Lægeforening* 2005;**125**(6):742-5.

Hauge 2017 *{published data only}*

Hauge KE, Ulvestad ME. Having a bad attitude? The relationship between attitudes and sickness absence. *IZA Journal of Labor Policy* 2017;**6**:11. [DOI: 10.1186/s40173-017-0088-y]

Herrmann 2015 *{published data only}*

Herrmann WJ, Haarmann A, Baerheim A. Regulations of sickness certification as a factor for increased health care utilization in Germany. *Zeitschrift fur Evidenz, Fortbildung und Qualitat im Gesundheitswesen* 2015;**109**(8):552-9. [DOI: 10.1016/j.zefq.2015.10.004]

Ihlebaek 2006 *{published data only}*

Ihlebaek C, Hansson TH, Laerum E, Brage S, Eriksen HR, Holm SH, et al. Prevalence of low back pain and sickness absence: a "borderline" study in Norway and Sweden. *Scandinavian Journal of Public Health* 2006;**34**(5):555-8.

Kaufmann 2010 *{published data only}*

Kaufmann T, Wäschle R, Bauer M, Schüpfer G. Management of short-term absence in a hospital: empirical investigations for implementation of an intervention protocol. *Der Anaesthesist* 2010;**59**(5):433-42. [DOI: 10.1007/s00101-010-1717-7]

Money 2013 *{published data only}*

Money A, Hussey L. A comparative analysis of self-reported and medically certified incidence data on work-related illness. HSE: Bootle, UK: HSE Books, Research Report 954.

Oyeflaten 2011 *{published data only}*

Oyeflaten I. Subjectivity in certification of sick leave. *Tidsskrift for den Norske Laegeforening* 2011;**131**(13-14):1308-10. [DOI: 10.4045/tidsskr.10.1446]

Pertold 2018 *{published data only}*

Pertold F, Westergaard-Nielsen N. Firm insurance and sickness absence of employees. *International Journal of Manpower* 2018;**39**(1):133-51. [DOI: 10.1108/IJM-05-2016-0097]

Pettersson-Lidbom 2013 *{published data only}*

Pettersson-Lidbom P, Thoursie PS. Temporary disability insurance and labor supply: evidence from a natural experiment. *Scandinavian Journal of Economics* 2013;**115**(2):485-507. [DOI: 10.1111/j.1467-9442.2012.01746.x]

Pollak 2017 *{published data only}*

Pollak C. The impact of a sick pay waiting period on sick leave patterns. *European Journal of Health Economics* 2017;**18**(1):13-31. [DOI: 10.1007/s10198-015-0755-0]

Preece 2006 *{published data only}*

Preece R. Do first-day absence schemes really work? Real evidence to prove the effectiveness of such arrangements is hard to come by. *Occupational Health* 2006;**58**(7):10.

Royneland 2002 *{published data only}*

Royneland O, Hartvig P. Systematic evaluation of sickness certification II. *Tidsskrift for den Norske Laegeforening* 2002;**122**(12):1235.

Schlotzhauer 1985 *{published data only}*

Schlotzhauer DL, Rosse JG. A five-year study of a positive incentive absence control program. *Personnel Psychology* 1985;**38**(3):575-85. [DOI: 10.1111/j.1744-6570.1985.tb00561.x]

Voss 2001 *{published data only}*

Voss M, Floderus B, Diderichsen F. Changes in sickness absenteeism following the introduction of a qualifying day for sickness benefit - findings from Sweden Post. *Scandinavian Journal of Public Health* 2001;**29**(3):166-74.

Additional references

Arrelöv 2005

Arrelöv BE, Borgquist L, Svärdsudd KF. Influence of local structural factors on physician's sick-listing practice: a population-based study. *European Journal of Public Health* 2005;**15**(5):470-4.

Bennett 2010

Bennett JB. Social climate research. In: Weiner IB, Craighead WE editor(s). The Corsini Encyclopedia of Psychology. Wiley, 2010.

Covidence [Computer program]

Veritas Health Innovation. Covidence systematic review software. Version accessed 10 June 2018. Melbourne, Australia: Veritas Health Innovation, 2018.

Englund 2000

Englund L, Svärdsudd K. Sick-listing habits among general practitioners in a Swedish county. *Scandinavian Journal of Primary Health Care* 2000;**18**(2):81-6.

EPOC

Cochrane Effective Practice and Organisation of Care Group (EPOC). Suggested risk of bias criteria for EPOC reviews. epoc.cochrane.org/epoc-resources-review-authors (accessed 7 February 2018).

Eurofound 2010

Eurofound. Absence from work. www.eurofound.europa.eu/sites/default/files/ef_files/docs/ewco/tn0911039s/tn0911039s.pdf (accessed 6 February 2018).

Gabbay 2011

Gabbay M, Taylor L, Sheppard L, Hillage J, Bambra C, Ford F, et al. NICE guidance on long-term sickness and incapacity. *British Journal of General Practice* 2011;**61**(584):e118-24.

GRADEpro GDT 2015 [Computer program]

McMaster University (developed by Evidence Prime Inc). GRADEpro GDT. Hamilton (ON): McMaster University (developed by Evidence Prime Inc), 2015.

Hesselius 2013

Hesselius P, Johansson P, Vikström J. Social behaviour in work absence. *Scandinavian Journal of Economics* 2013;**115**(4):995-1019.

Higgins 2011

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

Higgins 2012

Higgins A, O'Halloran P, Porter S. Management of long term sickness absence: a systematic realistic review. *Journal of Occupational Rehabilitation* 2012;**22**(3):322-32.

Higgins 2017

Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors). Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). The Cochrane Collaboration, 2017. Available at www.training.cochrane.org/handbook.

Hinkka 2018

Hinkka K, Niemelä M, Autti-Rämö I, Palomäki H. Physicians' experiences with sickness absence certification in Finland. Scandinavian Journal of Public Health 2018 [Epub ahead of print]. [DOI: 10.1177/1403494818758817]

Kankaanpää 2014

Kankaanpää A. Sick leave prescribing practices in Finland. Turku, Finland:Turun yliopiston julkaisuja;2014. Annales Universitatis Turkuensis D 1132.

Khan 2009

Khan J, Rehnberg C. Perceived job security and sickness absence: a study on moral hazard. *European Journal of Health Economics* 2009;**10**:421-8.

Lahelma 2013

Lahelma E, Aittomäki A, Laaksonen M, Lallukka T, Martikainen P, Piha K, et al. Cohort profile: the Helsinki Health Study. *International Journal of Epidemiology* 2013;**42**(3):722-30. [DOI: 10.1093/ije/dys039]

Leinonen 2018

Leinonen T, Viikari-Juntura E, Husgafvel-Pursiainen K, Solovieva S. Cause-specific sickness absence trends by occupational class and industrial sector in the context of recent labour market changes: a Finnish panel data study. *BMJ Open* 2018;**8**(4):1-11. [DOI: 10.1136/bmjopen-2017-019822]

Letrilliart 2012

Letrilliart L, Barrau A. Difficulties with the sickness certification process in general practice and possible solutions: a systematic review. *European Journal of General Practice* 2012;**18**(4):219-28.

Melchior 2003

Melchior M, Niedhammer I, Berkman LF, Goldberg M. Do psychosocial work factors and social relations exert

independent effects on sickness absence? A six year prospective study of the GAZEL cohort. *Journal of Epidemiology and Community Health* 2003;**57**:285-93.

Michie 2003

Michie S, Williams S. Reducing work related psychological ill health and sickness absence: a systematic literature review. *Occupational and Environmental Medicine* 2003;**60**(1):3-9.

Mykletun 2014

Mykletun A, Torsvik G, Vaage K. Gatekeeping versus monitoring: evidence from a case with extended self-reporting of sickness absence. Bergen, Norway: Department of Economics, University of Bergen. 2014 Working Papers in Economics,.

NICE 2009

National Institute for Health and Care Excellence (NICE). Public health guidance 19. Management of long-term sickness absence and incapacity for work. guidance.nice.org.uk/PH19 (accessed 6 February 2018).

Nieuwenhuijsen 2004

Nieuwenhuijsen K, Verbeek JH, de Boer AG, Blonk RW, van Dijk FJ. Supervisory behaviour as a predictor of return to work in employees absent from work due to mental health problems. *Occupational and Environmental Medicine* 2004;**61**(10):817-23.

NOSOSCO 2015

NOSOSCO, Project group. Sickness absence in the Nordic countries. Copenahgen, Denmark: Nordic Council of Ministers, NOMESCO-NOSOSCO: 2015. Nordic Social Statistical Committee report 59.

Pesonen 2016

Pesonen S, Halonen JI, Liira J. Self-reporting - a study on implementing and the effects of self-reporting of sick leaves [Omailmoitus - tutkimus sairauspoissaolojen omailmoituksen käyttöönotosta ja vaikutuksista]. Helsinki, Finland: Finnish Institute of Occupational Health; 2016.

Pichler 2015

Pichler S. Sickness absence, moral hazard, and the business cycle. *Health Economics* 2015;**24**(6):692-710. [PUBMED: 24737552]

Ramsay 2003

Ramsay CR, Matowe L, Grilli R, Grimshaw JM, Thomas RE. Interrupted time series designs in health technology assessment: lessons from two systematic reviews of behavior change strategies. *International Journal of Technology Assessment in Health Care* 2003;**19**(4):613-23.

RevMan 2014 [Computer program]

The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Schünemann 2017

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, et al on behalf of the Cochrane Applicability and Recommendations Methods Group. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors). Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). The Cochrane Collaboration, 2017. Available at www.training.cochrane.org/handbook.

Spasova 2016

Spasova S, Bouget D, Vanhercke B. Sick pay and sickness benefit schemes in the European Union: background report for the Social Protection Committee's In-Depth Review on sickness benefits. Brussels, Belgium: European Commission; 2016.

Sterne 2016

Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016;**355**:i4919.

Viscusi 2005

Viscusi WK, Harrington JE, Vernon JM (editors). Economics of regulation and anti-trust. Regulation of Workplace Health and Safety.. Massachusetts Institute of Technology, 2005.

Whitaker 2001

Whitaker SC. The management of sickness absence. *Occupational and Environmental Medicine* 2001;**58**(6):420-4.

Winde 2012

Winde LD, Alexanderson K, Carlsen B, Kjeldgård L, Wilteus AL, Gjesdal S. General practitioners' experiences with sickness certification: a comparison of survey data from Sweden and Norway. *BMC Family Practice* 2012;**13**:10.

Wynne-Jones 2008

Wynne-Jones G, Mallen CD, Welsh V, Dunn KM. Rates of sickness certification in European primary care: a systematic review. *European Journal of General Practice* 2008;**14**(3-4):99-108.

Wynne-Jones 2010

Wynne-Jones G, Mallen CD, Main CJ, Dunn KM. What do GPs feel about sickness certification? A systematic search and narrative review. *Scandinavian Journal of Primary Health Care* 2010;**28**(2):67-75.

References to other published versions of this review

Kausto 2018

Kausto J, Verbeek JH, Ruotsalainen JH, Halonen JI, Virta LJ, Kankaanpää E. Self-certification versus physician certification of sick leave for reducing sickness absence and associated costs. *Cochrane Database of Systematic Reviews* 2018, Issue 8. [DOI: 10.1002/14651858.CD013098]

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Fleten 2009

Methods	**Study design:** Controlled before-after study **Study grouping:** Parallel group
Participants	**Baseline characteristics** **Inclusion criteria:** Intervention group: employees of Kristiansand municipality. Control group: employees of Arendal municipality **Exclusion criteria:** Not provided **Pretreatment:** Not provided
Interventions	**Intervention characteristics** Prolonging the period of self-certification of sickness absence (from 24 days per year divided into 3 to 24 periods) until 50 days per year, divided into 1 to 10 periods with a structured workplace follow-up
Outcomes	*Total number of short-term sickness absence periods (1 to 3 days)* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Unit of measure**: A sickness absence period • **Direction**: Lower is better • **Data value**: Change from baseline • **Notes**: Intervention: change from 2001 to 2003 (26%). Control: change from 2001 to 2003 (6%) *Mean length of short-term sickness absence periods* • **Outcome type**: Continuous outcome *Proportion of self reported periods of 1 to 3 days of all 1- to 3-day periods* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Scale**: Percentage • **Range**: 0 to 100 • **Direction**: Higher is better • **Data value**: Change from baseline • **Notes**: Intervention: from 87% to 93%. Control: stable 90% *Number of short-term sickness absence periods per employee* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Unit of measure**: Number of periods per employee • **Direction**: Lower is better • **Data value**: Endpoint *Total number of sickness absence periods (4 to 16 days)* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Unit of measure**: A period • **Direction**: Lower is better

Fleten 2009 *(Continued)*

- **Data value**: Endpoint

Total number of sickness absence periods (4 to 7 days)

- **Outcome type**: Continuous outcome
- **Reporting**: Partially reported
- **Unit of measure**: A period
- **Direction**: Lower is better
- **Data value**: Endpoint

Identification	**Sponsorship source:** Not provided **Country:** Norway **Setting:** Intervention carried out in a municipality of 5700 employees (with a control municipality, number of employees not given). **Authors' names:** Fleten N, Krane L, Johnsen R **Institution:** Institutt for Samfunnsmedisin, Universitetet i Tromsø **Email:** nils.fleten@uit.no **Address:** Institutt for Samfunnsmedisin, Universitetet i Tromsø, N-9037 Tromsø
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Judgement Comment: ROBINS-I (confounding): risk of bias Serious. Only t-test used, no confounders included in the analyses. No comparison group was available for sickness absences that lasted more than 3 days. We judged this to correspond to high risk of bias according to the RCT assessment.
Allocation concealment (selection bias)	Unclear risk	Judgement Comment: ROBINS-I (selection): risk of bias No information. Selection was based on municipalities, not participants, no descriptive information provided on the intervention and control groups to permit comparisons. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (classification of interventions): risk of bias Low. All Kristiansand municipality employees were included in the intervention group, and all Arendal employees were included in the control group. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	Judgement Comment: ROBINS-I (deviation from intended interventions): risk of bias No information. No information about the changes in the intervention and control groups. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement Comment: ROBINS-I (missing data): risk of bias No information. No information about missing data was provided. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Selective reporting (reporting bias)	Unclear risk	Judgement Comment: ROBINS-I (selective reporting): risk of bias No information. No protocol for the statistical analyses provided. We judged this to correspond to unclear risk of bias according to the RCT assessment.

Fleten 2009 *(Continued)*

Other bias	Unclear risk	Judgement Comment: ROBINS-I (measurement of outcomes): risk of bias No information as there was no comparison group for sickness absence lasting more than 3 days. We judged this to correspond to unclear risk of bias according to the RCT assessment.

Hesselius 2005

Methods	**Study design:** Randomised controlled trial **Study grouping:** Parallel group **Description:** A randomly assigned treatment group was allowed to receive sickness benefits for 2 weeks without showing a doctor's certificate, instead of 1 week as usual. Randomisation was performed using date of birth. All insured born on an even date were asked to show a doctor's certificate after 2 weeks, whereas insured born on an uneven date had to show one already after 1 week.
Participants	**Baseline characteristics** Treatment group was allowed to receive sickness benefits for 2 weeks without showing a doctor's certificate. **Inclusion criteria:** Living in Gothenburg or in Jämtland in Sweden in the second half of 1988. Being an insurance claimant. **Exclusion criteria:** Not reported **Pretreatment:** There were no significant differences between the intervention and control groups with respect to any of the characteristics: gender, age, and income distributions, benefit cap, as well as average age and average sickness absence prior to the experiment were all equal between intervention and control groups. Jämtland: intervention group mean age 38.9, females 47.5%; control group mean age 38.7 years, females 47.6%. Gothenburg: intervention group mean age 38.1, females 48.2%; control group mean age 38.2 years, females 48.1%.
Interventions	**Intervention characteristics** • Treatment group was allowed to receive sickness benefits for 2 weeks without showing a doctor's certificate. • Control group had to show a doctor's certificate after 1 week of sickness absence.
Outcomes	*Risk of exit from sickness absence on day 7* • **Outcome type**: Continuous outcome • **Reporting**: Fully reported • **Unit of measure**: Risk Ratio • **Direction**: Higher is better • **Data value**: Endpoint • **Notes**: Jämtland: intervention group RR = 0.37. Gothenburg: intervention group RR = 0.34 *Risk of exit from sickness absence on day 14* • **Outcome type**: Continuous outcome • **Reporting**: Fully reported • **Unit of measure**: Risk Ratio • **Direction**: Higher is better • **Data value**: Endpoint • **Notes**: Jämtland: intervention group RR = 3.12. Gothenburg: intervention group RR = 3.39 *Hazard ratio for ending sickness absence on day 7*

Hesselius 2005 *(Continued)*

- **Outcome type**: Continuous outcome
- **Reporting**: Fully reported
- **Unit of measure**: Hazard ratio (from Figures)
- **Direction**: Higher is better
- **Data value**: Endpoint
- **Notes**: Jämtland: intervention group HR ˜ 0.15, control group HR ˜ 0.40. Gothenburg: intervention group HR ˜ 0.16, control group HR ˜ 0.47

Hazard ratio for ending sickness absence on day 14

- **Outcome type**: Continuous outcome
- **Reporting**: Fully reported
- **Unit of measure**: Hazard ratio (from Figures)
- **Direction**: Higher is better
- **Data value**: Endpoint
- **Notes**: Jämtland: intervention group HR ˜ 0.35, control group HR ˜ 0.13. Gothenburg: intervention group HR ˜ 0.41, control group HR ˜ 0.12

Identification	**Sponsorship source:** Wallander and Hedelius Foundation **Country:** Sweden **Setting:** Randomised trial (social experiment) **Author's name:** Hesselius Patrik **Institution:** IFAU and Uppsala University **Email:** patrik.hesselius@ifau.uu.se **Address:** P.O. Box 513, 75120 Uppsala
Notes	Double reporting: Hesselius P, Johansson P, Vikström J. Social behaviour in work absence. *Journal of Economics* 2013;115(4):995-1019

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Judgement Comment: Sequence generation was based on date of birth (even or uneven). This is not a true randomisation. We judged this domain as at high risk of bias according to the RCT assessment.
Allocation concealment (selection bias)	High risk	Judgement Comment: The authors do not report how or if they concealed allocation between randomisation and the start of the study. We judged this domain as at high risk of bias according to the RCT assessment.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Judgement Comment: Non-blinded experiment. We judged this domain as at high risk of bias according to the RCT assessment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Judgement Comment: Register-based outcome. We judged this domain as at low risk of bias according to the RCT assessment.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Judgement Comment: No information provided on completeness of outcome data.

Hesselius 2005 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	Judgement: No analysis plan provided. Total incidence and prevalence used as basis for reported findings.
Other bias	Low risk	Judgement Comment: No evidence of other biases.

Saksvik 2001

Methods	**Study design:** Controlled before-after study **Study grouping:** Parallel group
Participants	**Baseline characteristics** **Inclusion criteria:** Healthcare sector employees whose responsibility was to provide services and assistance to elderly and handicapped patients of a district in a municipality. Intervention group: from 15 units, N eligible 267, of which 167 participants answered before and after questionnaires. Control group: employees of an administrative unit from the same municipality, but from another district. N eligible 355, of which 100 answered before and after questionnaires. **Exclusion criteria:** Not reported **Pretreatment:** Not reported
Interventions	**Intervention characteristics** Changing from practice allowing 1 to 3 self administered sick leave days 4 times/year to practice allowing 1 to 5 self administered sick leave days 4 times/year
Outcomes	Per cent lost time of total possible work time due to *short-term self reported absenteeism (1 to 3 days or 1 to 5 days)* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Scale**: Per cent • **Range**: 0 to 100 • **Unit of measure**: Per cent lost time of total possible work time • **Direction**: Lower is better • **Data value**: Values for 3 time points • **Notes**: No clear difference between intervention and control group in the development of absenteeism between the pre- and postintervention measurements. Intervention had no effect on absenteeism. Per cent lost time of total possible work time due to *middle-/long-term sickness absence (4 to 16 days or 6 to 16 days)* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Scale**: Per cent • **Range**: 0 to 100 • **Unit of measure**: Per cent lost time of total possible work time • **Direction**: Lower is better • **Data value**: Values for 3 time points Per cent lost time of total possible work time due to *long-term sickness absence (over 16 days)* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported

Saksvik 2001 *(Continued)*

- **Scale**: Per cent
- **Range**: 0 to 100
- **Unit of measure**: Per cent lost time of total possible work time
- **Direction**: Lower is better
- **Data value**: Values for 3 time points

Per cent lost time of total possible work time due to *total sickness absenteeism*

- **Outcome type**: Continuous outcome
- **Reporting**: Partially reported
- **Scale**: Per cent
- **Range**: 0 to 100
- **Unit of measure**: Per cent lost time of total possible work time
- **Direction**: Lower is better
- **Data value**: Values for 3 time points

Identification	**Sponsorship source:** Not provided **Country:** Norway **Setting:** A municipality **Author's name:** Per Oystein Saksvik **Institution:** Norwegian University of Science and Technology **Email:** per.saksvik@svt.ntnu.no **Address:** Håkon Magnussons gt. 1B, N-7491 Trondheim, Norway
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Judgement Comment: ROBINS-I (confounding): risk of bias Serious. MANOVA used to compare the sample means, but confounders other than age not included in the models. Sex and job type not adjusted. No information on co-interventions. We judged this to correspond to high risk of bias according to the RCT assessment.
Allocation concealment (selection bias)	Low risk	Judgement Comment: ROBINS-I (selection): risk of bias Low. Employment in a certain district of a municipality was the selection method for both intervention and control groups. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (classification of intervention): risk of bias Low. Intervention status was based on employment in a certain district of the same municipality. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (deviation from intended interventions): risk of bias Low. No reason to believe that the intervention changed over the follow-up period. We judged this to correspond to low risk of bias according to the RCT assessment.
Incomplete outcome data (attrition bias)	Unclear risk	Judgement Comment: ROBINS-I (missing data): risk of bias No information. Outcome data from employee registers. No information provided on vari-

Saksvik 2001 *(Continued)*

All outcomes		ables/data that were missing. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Selective reporting (reporting bias)	Low risk	Judgement Comment: ROBINS-I (selective reporting): risk of bias Low. The planned analyses were reported. We judged this to correspond to low risk of bias according to the RCT assessment.
Other bias	Low risk	Judgement Comment: ROBINS-I (measurement of outcome): risk of bias Low. Absenteeism measure based on municipality records. Within the same municipality the procedure likely to be similar within districts, i.e. intervention and control groups, but no information on this was provided. We judged this to correspond to low risk of bias according to the RCT assessment.

Taylor 1969

Methods	**Study design:** Controlled before-after study **Study grouping:** Parallel group
Participants	**Baseline characteristics** **Inclusion criteria:** Intervention group included manual hourly paid employees of a workplace. **Exclusion criteria:** Not reported **Pretreatment:** Low turnover rates in both (employee and staff) groups. Age structure remains substantially unchanged.
Interventions	**Intervention characteristics** Introduction of self certificate. Manual hourly paid employees could have absence up to 3 working days without a doctor's certificate, but they had to complete a sickness report upon return to work.
Outcomes	*Per cent change in sickness absence periods* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Range**: 0 to 100 • **Unit of measure**: Percentage change in sickness absence periods • **Direction**: Smaller increase is better/larger decrease is better • **Data value**: Change from baseline • **Notes**: The labour rates in 1966 showed a rise in periods of all durations of 25% (i.e. +25%) over the rate in 1964 in the intervention group, stable in control group. *Per cent change in total time lost* • **Outcome type**: Continuous outcome • **Reporting**: Partially reported • **Range**: 0 to 100 • **Unit of measure**: Percentage change in total time lost • **Direction**: Smaller increase is better/larger decrease is better • **Data value**: Change from baseline • **Notes**: Fall of 15% (i.e. −15%) in total time lost in the intervention group, stable in control group.
Identification	**Sponsorship source:** Not reported **Country:** United Kingdom

Taylor 1969 *(Continued)*

	Setting: Intervention among manual workers and staff employees of an oil refinery
	Author's name: Taylor PJ
	Institution: Institute of Occupational Health, London School of Hygiene
	Email: Not available
	Address: 43 Russell Square, London W.C.1.
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Judgement Comment: ROBINS-I (confounding): risk of bias Serious. Only age considered in the analyses. We judged this to correspond to high risk of bias according to the RCT assessment.
Allocation concealment (selection bias)	Unclear risk	Judgement Comment: ROBINS-I (selection): risk of bias No information. All manual labour workers from an oil company were included in the intervention group, and all staff members of the same company were included in the control group. No proper descriptive data provided for these groups. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (classification of intervention): risk of bias Low. All labour were included in the intervention group, and all staff were included in the control group. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (deviation from intended interventions): risk of bias Low. No reason to believe that the intervention changed over the follow-up period. We judged this to correspond to low risk of bias according to the RCT assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (missing data): risk of bias Low. Authors used "comprehensive medical and sickness absence records". No reference to missing data. We judged this to correspond to low risk of bias according to the RCT assessment.
Selective reporting (reporting bias)	Unclear risk	Judgement Comment: ROBINS-I (selective reporting): risk of bias No information. Information is insufficient to make judgement: no protocol, no description of methods. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Other bias	Low risk	Judgement Comment: ROBINS-I (measurement of outcome): risk of bias Low. Sickness absence records held by the employer were used, and the intervention and control groups were within the same company. We judged this to correspond to low risk of bias according to the RCT assessment.

Torsvik 2014

Methods	**Study design:** Controlled before-after study
	Study grouping: Parallel group
Participants	**Baseline characteristics**

Torsvik 2014 *(Continued)*

	Inclusion criteria: Intervention group: employment in Mandal municipality. Control group: employment in 63 municipalities of the same size range and similar municipality characteristics as Mandal **Exclusion criteria:** Not provided **Pretreatment:** Mean number of employees before the experiment: 1032 in Mandal and 861 in control group; after experiment: 1205 in Mandal and 1006 in control group. Fraction of women before experiment: 0.83 in Mandal and 0.81 in control group; after experiment: 0.83 in Mandal and 0.82 in control group
Interventions	**Intervention characteristics** Prolonging the period of self-certification of sickness absence up to 1 year (the whole benefit period) with a structured workplace follow-up.
Outcomes	*Change in percentage of sickness absence (periods lasting more than 16 days) of contracted working days* • **Outcome type**: Continuous outcome • **Reporting**: Fully reported • **Scale**: Per cent • **Range**: 0 to 100 • **Unit of measure**: % • **Direction**: Smaller increase is better/larger decrease is better • **Data value**: Change from baseline • **Notes**: In the intervention group percentage points of absence of contracted days was 1.25 lower in the post- vs pre-reform period; the corresponding change in the control group was 0.17. *Difference-in-difference in means of absence % of contracted workdays between intervention and control group* • **Outcome type**: Continuous outcome • **Reporting**: Fully reported • **Scale**: Per cent • **Range**: 0 to 100 • **Unit of measure**: % • **Direction**: Smaller increase is better/larger decrease is better • **Data value**: Change from baseline • **Notes**: Difference-in-difference in means (of absence days/contracted days) between intervention and control groups was 1.08 percentage points.
Identification	**Sponsorship source:** Not reported **Country:** Norway **Setting:** Self-certification reform carried out among employees of a municipality in Norway **Authors' names:** Gaute Torsvik, Kjell Vaage **Institution:** University of Oslo, University of Bergen **Email:** gaute.torsvik@econ.uio.no **Address:** -
Notes	Double reporting: Mykletun 2014

Risk of bias

Bias	Authors' judgement	Support for judgement

Torsvik 2014 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	Judgement Comment: ROBINS-I (confounding): risk of bias Moderate. A balanced panel of municipalities was used (Mandal vs similar municipalities) based on size and economic variables. However, no information on whether confounding variables were included in the analyses. We judged this to correspond to unclear risk of bias according to the RCT assessment.
Allocation concealment (selection bias)	Low risk	Judgement Comment: ROBINS-I (selection): risk of bias Low. Participants chosen based on employment in the intervention municipality; selection method in the control group was the same. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of participants and personnel (performance bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (classification of intervention): risk of bias Low. Intervention status was the same for all employees in the intervention and control municipalities. We judged this to correspond to low risk of bias according to the RCT assessment.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (deviation from intended interventions): risk of bias Low. Intervention was applied at the municipality level. However, no information on co-interventions in each municipality or adherence to intervention was provided. No reason to believe the intervention changed over the follow-up period. We judged this to correspond to low risk of bias according to the RCT assessment.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Judgement Comment: ROBINS-I (missing data): risk of bias Low. Outcome data for intervention and control groups obtained from the same national register (NAV). We judged this to correspond to low risk of bias according to the RCT assessment.
Selective reporting (reporting bias)	Low risk	Judgement Comment: ROBINS-I (selective reporting): risk of bias No information. No protocol or statistical analysis plan available. We judged this to correspond to low risk of bias according to the RCT assessment.
Other bias	Unclear risk	Judgement Comment: ROBINS-I (measurement of outcome): risk of bias Low. Outcome data were based on register (NAV) for both intervention and control groups. We judged this to correspond to unclear risk of bias according to the RCT assessment.

HR: hazard ratio; MANOVA: multivariate analysis of variance; RCT: randomised controlled trial; RR: risk ratio

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Carne 1969	Study design
De Paola 2014	Intervention
Felder 2008	Study design
Gjesdal 2005	Study design
Hauge 2017	Study design
Herrmann 2015	Study design
Ihlebaek 2006	Study design

Study	Reason for exclusion
Kaufmann 2010	Intervention
Money 2013	Study design
Oyeflaten 2011	Study setting
Pertold 2018	Intervention
Pettersson-Lidbom 2013	Intervention
Pollak 2017	Intervention
Preece 2006	Study setting
Royneland 2002	Study design
Schlotzhauer 1985	Intervention
Voss 2001	Intervention

DATA AND ANALYSES

Comparison 1. Longer self-certification versus shorter self-certification

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Duration of periods	3		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
1.1 Periods < 3 days	2		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Periods < 29 days	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Number of periods/worker	1		Mean Difference (IV, Fixed, 95% CI)	Totals not selected
2.1 Periods < 3 days	1		Mean Difference (IV, Fixed, 95% CI)	0.0 [0.0, 0.0]
3 Lost work time due to period of sickness absence	3		Mean Difference (Fixed, 95% CI)	Totals not selected
3.1 Periods < 3 days	2		Mean Difference (Fixed, 95% CI)	0.0 [0.0, 0.0]
3.2 Periods > 16 days	1		Mean Difference (Fixed, 95% CI)	0.0 [0.0, 0.0]

Analysis 1.1. Comparison 1 Longer self-certification versus shorter self-certification, Outcome 1 Duration of periods.

Study or subgroup	Self-certification N	Self-certification Mean(SD)	Physician certification N	Physician certification Mean(SD)	Mean Difference Fixed, 95% CI	Mean Difference Fixed, 95% CI
1.1.1 Periods < 3 days						
Saksvik 2001	215	2.7 (1.5)	1519	1.3 (1.5)		1.38[1.16,1.6]
Taylor 1969	1364	-0.4 (0.5)	505	-0 (0.7)		-0.32[-0.39,-0.25]
1.1.2 Periods < 29 days						
Hesselius 2005	116115	8.9 (14.7)	121276	8.2 (13.9)		0.67[0.55,0.79]

Favours self-certif. Favours physician certif.

Analysis 1.2. Comparison 1 Longer self-certification versus shorter self-certification, Outcome 2 Number of periods/worker.

Study or subgroup	Self-certification N	Self-certification Mean(SD)	Physician certification N	Physician certification Mean(SD)	Mean Difference Fixed, 95% CI	Mean Difference Fixed, 95% CI
1.2.1 Periods < 3 days						
Taylor 1969	1364	0.5 (1.2)	505	0 (1.1)		0.48[0.36,0.6]

Favours self-certif. Favours phys certif.

Analysis 1.3. Comparison 1 Longer self-certification versus shorter self-certification, Outcome 3 Lost work time due to period of sickness absence.

Study or subgroup	Self-certification N	Physician certification N	Mean Difference (SE)	Mean Difference IV, Fixed, 95% CI	Mean Difference IV, Fixed, 95% CI
1.3.1 Periods < 3 days					
Saksvik 2001	0	0	1.4 (0.112)		1.38[1.16,1.6]
Taylor 1969	1364	505	0.5 (0.034)		0.54[0.47,0.61]
1.3.2 Periods > 16 days					
Torsvik 2014	0	0	-2.8 (0.26)		-2.84[-3.35,-2.33]

Favours self-certif. Favours physician certif.

APPENDICES

Appendix 1. Detailed Cochrane Central Register of Controlled Trials (CENTRAL) search strategy

Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley Online Library)
Date Run: May 30, 2018 (from inception to search date)

ID Search Hits

#1 ((self* or oneself) near/2 (certif*)) or ((self* or oneself) near/2 declar*) or ((self* or oneself) near/2 administ* near/2 (sick* or ill*)) or ((sick* or ill*) near/2 certif*) or (uncertif* near/2 (sick* or ill*)) or (absen* near/2 certificat*) or (medical* near/2 certificat*) 103

#2 ((moral* or ethic*) near/2 (hazard* or absen*)) 10

#3 MeSH descriptor: [Morals] explode all trees 802

#4 MeSH descriptor: [Absenteeism] explode all trees 549

#5 #3 and #4 3

#6 #1 or #2 or #5 116

#7 MeSH descriptor: [Sick Leave] explode all trees 589

#8 ("medical leave*") 294

#9 (sick* near/2 leave*) 1551

#10 (sick* near/2 absen*) 359

#11 MeSH descriptor: [Return to Work] explode all trees 184

#12 (return* near/2 work*) 2300

#13 ((fit or sick*) near/1 (note* or notif*)) 15

#14 MeSH descriptor: [Insurance Claim Review] explode all trees 98

#15 (insur* near/1 claim*) 253

#16 MeSH descriptor: [Insurance, Health] explode all trees 1601

#17 ((sick* or health) near/1 insur*) 2651

#18 MeSH descriptor: [Health Benefit Plans, Employee] explode all trees 42

#19 (emplo* near/3 medical* near/3 (care or caring)) 18

#20 MeSH descriptor: [Occupational Health Services] explode all trees 423

#21 (occupational near/2 health near/2 service*) 603

#22 #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 7860

#23 #22 and #6 34

#24 MeSH descriptor: [Costs and Cost Analysis] explode all trees 26088

#25 (cost* near/2 analy*) 33694

#26 (cost* near/2 benef*) 24682

#27 (cost* near/2 effect*) 36970

#28 (cost* near/2 utilit*) 5076

#29 (economic* near/2 eval*) 22054

#30 (economic* near/2 analy*) 12023

#31 (economic* near/2 model*) 2664

#32 (cost* near/2 stud*) 16053

#33 (cost* near/2 effic*) 3078

#34 #24 or #25 or #26 or #27 or #28 or #29 or #30 or 30 #31 or #32 or #33 53644

#35 #6 and #34 36

#36 #23 or #35 50

Cochrane Central Register of Controlled Trials : Issue 5 of 12, May 2018

There are 15 results from 1275166 records for your search on #36 - #23 or #35 in Trials

Appendix 2. Detailed MEDLINE search strategy

Database: Medline (Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) 1946 to Present)

Search date: June 14, 2018

MEDLINE(R) Daily and Ovid MEDLINE(R) <1946 to Present>

Search Strategy:

1 (((self* or oneself) adj2 certif*) or ((self* or oneself) adj2 declar*) or ((self* or oneself) adj2 administ* adj2 (sick* or ill*)) or ((sick* or ill*) adj2 certif*) or (uncertif* adj2 (sick* or ill*)) or (absen* adj2 certificat*) or (medical* adj2 certificat*)).mp. (2306)

2 ((moral* or ethic*) adj2 (hazard* or absen*)).mp. (406)

3 exp Morals/ and Absenteeism/ (55)

4 1 or 2 or 3 (2760)

5 Sick Leave/ or "medical leave*".mp. (5411)

6 (sick* adj2 leave*).mp. (7894)

7 (sick* adj2 absen*).mp. (2654)

8 Return to Work/ (1710)

9 (return* adj2 work*).mp. (10871)

10 ((fit or sick*) adj1 (note* or notif*)).mp. (109)

11 Insurance Claim Review/ or (insur* adj1 claim*).mp. (11947)

12 exp Insurance, Health/ (137976)

13 ((sick* or health) adj1 insur*).mp. (73002)

14 exp Health Benefit Plans, Employee/ (10033)

15 (emplo* adj3 medical* adj3 (care or caring)).mp. (105)

16 exp Occupational Health Services/ (10306)

17 (occupational adj2 health adj2 service*).mp. (10953)

18 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 (193653)

19 4 and 18 (832)

20 exp "Costs and Cost Analysis"/ (216468)

21 (cost* adj2 analy*).mp. (129293)

22 (cost* adj2 benef*).mp. (86010)

23 (cost* adj2 effect*).mp. (120738)

24 (cost* adj2 utilit*).mp. (5001)

25 (economic* adj2 eval*).mp. (11705)

26 (economic* adj2 analy*).mp. (7013)

27 (economic* adj2 model*).mp. (11499)

28 (cost* adj2 stud*).mp. (8256)

29 (cost* adj2 effic*).mp. (14583)

30 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 (326287)

31 4 and 30 (216)

32 19 or 31 (918)

Appendix 3. Detailed SCOPUS search strategy

Scopus

Search date: May 23, 2018 (from inception to search date)

1 TITLE-ABS-KEY(((self* or oneself) W/2 certif*) or ((self* or oneself) W/2 declar*) or ((self* or oneself) W/2 administ* W/2 (sick* or ill*)) or ((sick* or ill*) W/2 certif*) or (uncertif* W/2 (sick* or ill*)) or (absen* W/2 certificat*) or (medical* W/2 certificat*)) (4724)

2 TITLE-ABS-KEY((moral* or ethic*) W/2 (hazard* or absen*)) (5081)

3 TITLE-ABS-KEY(moral* and absenteeism) (274)

4 1 or 2 or 3 (10002)

5 TITLE-ABS-KEY("medical leave*") (5206)

6 TITLE-ABS-KEY(sick* W/2 leave*) (8820)

7 TITLE-ABS-KEY(sick* W/2 absen*) (3496)

8 TITLE-ABS-KEY(return* W/2 work*) (16258)

9 TITLE-ABS-KEY((fit or sick*) W/1 (note* or notif*)) (245)

10 TITLE-ABS-KEY(insur* W/1 claim*) (12796)

11 TITLE-ABS-KEY((sick* or health) W/1 insur*) (152802)

12 TITLE-ABS-KEY(health W/3 benefit* W/3 plan*) (9763)

13 TITLE-ABS-KEY(emplo* W/3 medical* W/3 (care or caring)) (263)

14 TITLE-ABS-KEY(occupational W/2 health) (84560)

15 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 (266608)

16 4 and 15 (1283)

17 TITLE-ABS-KEY(cost* W/2 analy*) (323156)

18 TITLE-ABS-KEY(cost* W/2 benef*) (210308)

19 TITLE-ABS-KEY(cost* W/2 effect*) (396452)

20 TITLE-ABS-KEY(cost* W/2 utilit*) (13176)

21 TITLE-ABS-KEY(economic* W/2 eval*) (39362)

22 TITLE-ABS-KEY(economic* W/2 analy*) (78072)

23 TITLE-ABS-KEY(economic* W/2 model*) (41202)

24 TITLE-ABS-KEY(cost* W/2 stud*) (33086)

25 TITLE-ABS-KEY(cost* W/2 effic*) (75365)

26 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 (790811)

27 4 and 26 (465)

28 16 or 27 (1639, final result)

Appendix 4. Detailed PsycINFO Ovid search strategy

Database: **PsycINFO**

Search date: May 23, 2018 (from inception to search date)

---

1 (((self* or oneself) adj2 certif*) or ((self* or oneself) adj2 declar*) or ((self* or oneself) adj2 administ* adj2 (sick* or ill*)) or ((sick* or ill*) adj2 certif*) or (uncertif* adj2 (sick* or ill*)) or (absen* adj2 certificat*) or (medical* adj2 certificat*)).mp. (614)

2 ((moral* or ethic*) adj2 (hazard* or absen*)).mp. (413)

3 exp Morals/ and Absenteeism/ (3)

4 1 or 2 or 3 (1027)

5 Employee Leave Benefits/ or "medical leave*".mp. (1058)

6 (sick* adj2 leave*).mp. (1374)

7 (sick* adj2 absen*).mp. (1138)

8 Return to Work/ (1284)

9 (return* adj2 work*).mp. (2988)

10 ((fit or sick*) adj1 (note* or notif*)).mp. (26)

11 (insur* adj1 claim*).mp. (448)

12 exp Health Insurance/ (10062)

13 ((sick* or health) adj1 insur*).mp. (9136)

14 exp Employee Health Insurance/ (756)

15 (emplo* adj3 medical* adj3 (care or caring)).mp. (39)

16 exp Occupational Health/ and service*.mp. (366)

17 (occupational adj2 health adj2 service*).mp. (175)

18 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 (20368)

19 4 and 18 (187)

20 exp "COSTS AND COST ANALYSIS"/ (24217)

21 (cost* adj2 analy*).mp. (17523)

22 (cost* adj2 benef*).mp. (7647)

23 (cost* adj2 effect*).mp. (15174)

24 (cost* adj2 utilit*).mp. (724)

25 (economic* adj2 eval*).mp. (1819)

26 (economic* adj2 analy*).mp. (1611)

27 (economic* adj2 model*).mp. (1515)

28 (cost* adj2 stud*).mp. (1594)

29 (cost* adj2 effic*).mp. (2156)

30 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 (44595)

31 4 and 30 (53)

32 19 or 31 (228)

Appendix 5. Detailed EBM Reviews Ovid search strategy

Database: **EBM Reviews** - Cochrane Database of Systematic Reviews <2005 to June 7, 2018>, EBM Reviews - ACP Journal Club <1991 to May 2018>, EBM Reviews - Database of Abstracts of Reviews of Effects <1st Quarter 2016>, EBM Reviews - Cochrane Clinical Answers <May 2018>, EBM Reviews - Cochrane Central Register of Controlled Trials <May 2018>, EBM Reviews - Cochrane Methodology Register <3rd Quarter 2012>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>

Search date: 14.6.2018 (from inception to search date)

---

1 (((self* or oneself) adj2 certif*) or ((self* or oneself) adj2 declar*) or ((self* or oneself) adj2 administ* adj2 (sick* or ill*)) or ((sick* or ill*) adj2 certif*) or (uncertif* adj2 (sick* or ill*)) or (absen* adj2 certificat*) or (medical* adj2 certificat*)).mp. (108)

2 ((moral* or ethic*) adj2 (hazard* or absen*)).mp. (18)

3 exp Morals/ and Absenteeism/ (2)

4 1 or 2 or 3 (126)

5 Sick Leave/ or "medical leave*".mp. (794)

6 (sick* adj2 leave*).mp. (1488)

7 (sick* adj2 absen*).mp. (410)

8 Return to Work/ (161)

9 (return* adj2 work*).mp. (2354)

10 ((fit or sick*) adj1 (note* or notif*)).mp. (14)

11 Insurance Claim Review/ or (insur* adj1 claim*).mp. (237)

12 exp Insurance, Health/ (1537)

13 ((sick* or health) adj1 insur*).mp. (2568)

14 exp Health Benefit Plans, Employee/ (41)

15 (emplo* adj3 medical* adj3 (care or caring)).mp. (16)

16 exp Occupational Health Services/ (402)

17 (occupational adj2 health adj2 service*).mp. (576)

18 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 (7749)

19 4 and 18 (28)

20 exp "Costs and Cost Analysis"/ (25208)

21 (cost* adj2 analy*).mp. (33658)

22 (cost* adj2 benef*).mp. (25096)

23 (cost* adj2 effect*).mp. (37714)

24 (cost* adj2 utilit*).mp. (5056)

25 (economic* adj2 eval*).mp. (21789)

26 (economic* adj2 analy*).mp. (10476)

27 (economic* adj2 model*).mp. (2300)

28 (cost* adj2 stud*).mp. (11433)

29 (cost* adj2 effic*).mp. (3773)

30 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 (54547)

31 4 and 30 (31)

32 19 or 31 (43)

Appendix 6. Detailed CINAHL EBSCOhost search strategy

CINAHL	Search date: May 30, 2018

#	Query	Limiters/Expanders	Results
S32	S19 OR S31	Search modes - Boolean/Phrase	243
S31	S4 AND S30	Search modes - Boolean/Phrase	56
S30	S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29	Search modes - Boolean/Phrase	86,758
S29	(cost* N2 effic*)	Search modes - Boolean/Phrase	2,842
S28	(cost* N2 stud*)	Search modes - Boolean/Phrase	3,061
S27	(economic* N2 model*)	Search modes - Boolean/Phrase	873
S26	(economic* N2 analy*)	Search modes - Boolean/Phrase	1,711
S25	(economic* N2 eval*)	Search modes - Boolean/Phrase	2,663
S24	(cost* N2 utilit*)	Search modes - Boolean/Phrase	1,008
S23	(cost* N2 effect*)	Search modes - Boolean/Phrase	22,536
S22	(cost* N2 benef*)	Search modes - Boolean/Phrase	19,606
S21	(cost* N2 analy*)	Search modes - Boolean/Phrase	29,278
S20	(MH "Costs and Cost Analysis+")	Search modes - Boolean/Phrase	67,920
S19	S4 AND S18	Search modes - Boolean/Phrase	219
S18	S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17	Search modes - Boolean/Phrase	97,473
S17	(occupational N2 health N2 service*)	Search modes - Boolean/Phrase	4,830

(Continued)

S16	(MH "Occupational Health Services+")	Search modes - Boolean/Phrase	5,692
S15	(emplo* N3 medical* N3 (care or caring))	Search modes - Boolean/Phrase	61
S14	(MH "Health Benefit Plans, Employee")	Search modes - Boolean/Phrase	189
S13	((sick* or health) N1 insur*)	Search modes - Boolean/Phrase	43,256
S12	(MH "Insurance, Health+")	Search modes - Boolean/Phrase	65,557
S11	(insur* N1 claim*)	Search modes - Boolean/Phrase	618
S10	((fit or sick*) N1 (note* or notif*))	Search modes - Boolean/Phrase	100
S9	(return* N2 work*)	Search modes - Boolean/Phrase	3,987
S8	(MH "Job Re-Entry")	Search modes - Boolean/Phrase	4,735
S7	(sick* N2 absen*)	Search modes - Boolean/Phrase	1,084
S6	(sick* N2 leave*)	Search modes - Boolean/Phrase	3,913
S5	(MH "Sick Leave") OR ("medical leave*")	Search modes - Boolean/Phrase	3,842
S4	S1 OR S2 OR S3	Search modes - Boolean/Phrase	784
S3	(MH "Morals+") AND (MH "Absenteeism")	Search modes - Boolean/Phrase	19
S2	((moral* or ethic*) N2 (hazard* or absen*))	Search modes - Boolean/Phrase	145
S1	((self* or oneself) N2 (certif*)) or ((self* or oneself) N2 declar*) or ((self* or oneself) N2 administ* N2 (sick* or ill*)) or ((sick* or ill*) N2 certif*) or (uncertif* N2 (sick* or ill*)) or (absen* N2 certificat*) or (medical* N2 certificat*)	Search modes - Boolean/Phrase	620

Appendix 7. Detailed EconLit ProQuest, EconPapers search strategies

EconLit 14.6.2018 (from inception to search date)

Set#	Searched for	Databases	Results
S1	(((self* or oneself) NEAR/2 (certif*)) or ((self* or oneself) NEAR/2 declar*) or ((self* or oneself) NEAR/2 administ* NEAR/2 (sick* or ill*)) or ((sick* or ill*) NEAR/2 certif*) or (uncertif* NEAR/2 (sick* or ill*)) or (absen* NEAR/2 certificat*) or (medical* NEAR/2 certificat*)) OR ((moral* or ethic*) NEAR/2 (hazard* or absen*))	EconLit	5109
S2	("medical leave*") OR (sick* NEAR/2 leave*) OR (sick* NEAR/2 absen*) OR (return* NEAR/2 work*) OR ((fit or sick*) NEAR/1 (note* or notif*)) OR (insur* NEAR/1 claim*) OR su(Health Insurance, Public and Private) OR ((sick* or health) NEAR/1 in-	EconLit	7882

(Continued)

	sur*) OR (emplo* NEAR/3 medical* NEAR/3 (care or caring)) OR (occupational NEAR/2 health NEAR/2 service*)		
S3	S1 AND S2	EconLit These databases are searched for part of your query.	326
S4	(cost* NEAR/2 analy*) OR (cost* NEAR/2 benef*) OR (cost* NEAR/2 effect*) OR (cost* NEAR/2 utilit*) OR (economic* NEAR/2 eval*) OR (economic* NEAR/2 analy*) OR (economic* NEAR/2 model*) OR (cost* NEAR/2 stud*) OR (cost* NEAR/2 effic*)	EconLit	123744
S5	S1 AND S4	EconLit These databases are searched for part of your query.	431
S6	S3 OR S5	EconLit These databases are searched for part of your query.	710

EconPapers 16.8.2018

94 documents matched the search for ("sick* absence" OR "sick* leave" OR "medical leave") AND (certif* OR note* OR notif*OR administ*).

16 duplicates

7 registered authors notifications

71 results

Appendix 8. Detailed EBSCO Business Source Elite (EBSCOhost) search strategy

EBSCOhost Business Source Elite

Search date: May 30, 2018 (from inception to search date)

1 ((self* or oneself) N2 certif*) or ((self* or oneself) N2 declar*) or ((self* or oneself) N2 administ* N2 (sick* or ill*)) or ((sick* or ill*) N2 certif*) or (uncertif* N2 (sick* or ill*)) or (absen* N2 certificat*) or (medical* N2 certificat*) (1624)

2 (moral* or ethic*) N2 (hazard* or absen*) (4044)

3 MH ABSENTEEISM (Labor) (3119)

4 1 or 2 or 3 (8714)

5 MH SICK leave or "medical leave*" (6845)

6 (sick* N2 leave*) (4094)

7 (sick* N2 absen*) (1390)

8 MH EMPLOYMENT reentry (885)

9 (return* N2 work*) or reemploy* or re-employ* (6530)

10 ((fit or sick*) N1 (note* or notif*)) (242)

11 MH DEFINED contribution health benefit plans or (insur* N1 claim*) (16480)

12 ((sick* or health) N1 insur*) (95315)

13 (health N3 benefit N3 plan*) (1987)

14 (emplo* N3 medical* N3 (care or caring)) (1974)

15 MH INDUSTRIAL hygiene (16041)

16 (occupational N2 health N2 service*) (2658)

17 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 (142187)

18 4 and 17 (1445)

19 MH COST analysis (9128)

20 (cost* N2 analy*) (46709)

21 (cost* N2 benef*) (48005)

22 (cost* N2 effect*) (57909)

23 (cost* N2 utilit*) (3125)

24 (economic* N2 eval*) (4590)

25 (economic* N2 analy*) (53950)

26 (economic* N2 model*) (39756)

27 (cost* N2 stud*) (6054)

28 (cost* N2 effic*) (14838)

29 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 (176596)

30 4 and 29 (399)

31 18 or 30 (1794, final result)

Appendix 9. Detailed OSH-UPDATE search strategy

OSH Update (databases: NIOSHTIC, NIOSHTIC-2, HSELINE, CISDOC)

Search date: May 30, 2018 (from inception to search date)

1 GW{((self* or oneself) and certif*) or ((self* or oneself) and declar*) or ((self* or oneself) and administ* and (sick* or ill*)) or ((sick* or ill*) and certif*) or (uncertif* and (sick* or ill*)) or (absen* and certificat*) or (medical* and certificat*)} (6605)

2 GW{(moral* or ethic*).-2.(hazard* or absen*)} (40)

3 GW{moral* and absenteeism} (382)

4 1 or 2 or 3 (6773)

5 GW{medical leave*} (140)

6 GW{sick*.-2.leave*} (1362)

7 GW{sick*.-2.absen*} (2329)

8 GW{return*.-2.work*} (1862)

9 GW{(fit or sick*).-1.(note* or notif*)} (61)

10 GW{insur*.-1.claim*} (249)

11 GW{(sick* or health).-1.insur*} (962)

12 GW{health.-3.benefit*.-3.plan*} (14)

13 GW{emplo*.-3.medical*.-3.(care or caring)} (26)

14 GW{occupational.-2.health} (80583)

15 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 (84865)

16 4 and 15 (947)

17 GW{cost*.-2.analy*} (1549)

18 GW{cost*.-2.benef*} (1397)

19 GW{cost*.-2.effect*} (1873)

20 GW{cost*.-2.utilit*} (19)

21 GW{economic*.-2.eval*} (301)

22 GW{economic*.-2.analy*} (442)

23 GW{economic*.-2.model*} (123)

24 GW{cost*.-2.stud*} (173)

25 GW{cost*.-2.effic*} (227)

26 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 (4827)

27 4 and 26 (80)

28 16 or 27 (989)

29 DC{OUNIOC or OUNIOS or OUHSEL or OUCISD} (580458)

30 28 and 29 (417, final result)

CONTRIBUTIONS OF AUTHORS

Conceiving the protocol and review: JK, JHV, JHR, JIH, EK, LJV.

Designing the protocol and review: JK, JHV, JHR, JIH, EK, LJV.

Co-ordinating the protocol and review: JK.

Designing search strategies: JK, JHV, JHR, JIH, EK, LJV.

Writing the protocol and review: JK.

Securing funding for the protocol and review: JK, JHV, LJV.

Performing previous work that was the foundation of the current study: JK, JHV, JIH.

DECLARATIONS OF INTEREST

Johanna Kausto: None known.

Jos H Verbeek: None known.

Jani H Ruotsalainen: None known.

Jaana I Halonen: None known.

Lauri J Virta: None known.

Eila Kankaanpää: None known.

SOURCES OF SUPPORT

Internal sources

- Finnish Institute of Occupational Health, Finland.

  Salaries for Johanna Kausto, Jos Verbeek, Jani Ruotsalainen, and Jaana Halonen.
- Social Insurance Institution of Finland, Finland.

  Salary for Lauri Virta.
- University of Eastern Finland, Finland.

  Salary for Eila Kankaanpää.

External sources

- Social Insurance Institution of Finland, Finland.

  A research grant of EUR 54,000 to the Finnish Institute of Occupational Health to enable Johanna Kausto to co-ordinate and conduct this review.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We had planned to compile an additional GRADE table showing our decisions about the certainty of evidence and their justifications. Instead, we have presented our grading of the certainty of evidence in Summary of findings for the main comparison. For the CBA studies, we had planned to plot the outcome measurements both at baseline and follow-up to ensure that baseline imbalances were considered, but instead of using plotted values, we used change from before to after values for the calculations.

We had planned to consider three follow-up time periods: up to three months, between three and six months, and more than six months, and to analyse the included studies accordingly. However, as all included studies employed follow-up times longer than six months, this differentiation was not possible.

We had also planned to compare the effects of interventions in subgroup analyses, but no such data were available from the included studies.

We had planned to report the change in costs of short-term sickness absence per employee and total costs of short-term sickness absence, however the study by Hesselius 2005 reported only total costs. We had also planned to take into account variation in items included in the cost estimates (administrative costs of handling short-term sickness absence, value of lost production, and costs of health care) and in the methods used to value these items, but there was no need to consider variation in any of these items as only one study reported costs.

INDEX TERMS

Medical Subject Headings (MeSH)

*Absenteeism; *Employment; *Physical Examination; *Sick Leave [statistics & numerical data]; Certification; Physicians; Randomized Controlled Trials as Topic; Time Factors

MeSH check words

Humans