Occupational safety and health enforcement tools for preventing occupational diseases and injuries (Review)

Mischke C, Verbeek JH, Job J, Morata TC, Alvesalo-Kuusi A, Neuvonen K, Clarke S, Pedlow RI

THE COCHRANE COLLABORATION®

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2013, Issue 8

http://www.thecochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Christina Mischke[1], Jos H Verbeek[1], Jenny Job[2], Thais C Morata[3], Anne Alvesalo-Kuusi[4], Kaisa Neuvonen[5], Simon Clarke[6], Robert I Pedlow[7]

[1]Cochrane Occupational Safety and Health Review Group, Finnish Institute of Occupational Health, Kuopio, Finland. [2]Strategic Policy Branch, Safe Work Australia, Canberra, Australia. [3]National Institute for Occupational Safety and Health (NIOSH), Cincinnati, OH, USA. [4]Faculty of Law, University of Turku, Turku, Finland. [5]Cochrane Occupational Safety and Health Review Group, Finnish Institute of Occupational Health, Helsinki, Finland. [6]UK Health and Safety Executive, Merseyside, UK. [7]Research and Evaluation Team, Safe Work Australia, Canberra, Australia

Contact address: Christina Mischke, Cochrane Occupational Safety and Health Review Group, Finnish Institute of Occupational Health, Neulaniementie 4, PO Box 310, Kuopio, 70101, Finland. christina.mischke@ttl.fi. tinamischke@gmx.de.

Editorial group: Cochrane Occupational Safety and Health Group.
Publication status and date: New, published in Issue 8, 2013.
Review content assessed as up-to-date: 20 March 2013.

Citation: Mischke C, Verbeek JH, Job J, Morata TC, Alvesalo-Kuusi A, Neuvonen K, Clarke S, Pedlow RI. Occupational safety and health enforcement tools for preventing occupational diseases and injuries. *Cochrane Database of Systematic Reviews* 2013, Issue 8. Art. No.: CD010183. DOI: 10.1002/14651858.CD010183.pub2.

ABSTRACT

Background

There is uncertainty as to whether and what extent occupational safety and health regulation and legislation enforcement activities, such as inspections, are effective and efficient to improve workers' health and safety. We use the term regulation to refer both to regulation and legislation.

Objectives

To assess the effects of occupational safety and health regulation enforcement tools for preventing occupational diseases and injuries.

Search methods

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE (PubMed), EMBASE (embase.com), CINAHL (EBSCO), PsycINFO (Ovid), OSH update, HeinOnline, Westlaw International, EconLit and Scopus from the inception of each database until January 2013. We also checked reference lists of included articles and contacted study authors to identify additional published, unpublished and ongoing studies.

Selection criteria

We included randomised controlled trials (RCTs), controlled before-after studies (CBAs), interrupted time series (ITS) and econometric panel studies of firms or workplaces evaluating inspections, warnings or orders, citations or fines, prosecution or firm closure by governmental representatives and if the outcomes were injuries, diseases or exposures.

In addition, we included qualitative studies of workers' or employers' attitudes or beliefs towards enforcement tools.

Data collection and analysis

Pairs of authors independently extracted data on the main characteristics, the risk of bias and the effects of the interventions. We expressed intervention effects as risk ratios (RR) or mean differences (MD). We recalculated other effect measures into RRs or MDs. We combined the results of similar studies in a meta-analysis.

Main results

We located 23 studies: two RCTs with 1414 workplaces, two CBAs with 9903 workplaces, one ITS with six outcome measurements, 12 panel studies and six qualitative studies with 310 participants. Studies evaluated the effects of inspections in general and the effects of their consequences, such as penalties. Studies on the effects of prosecution, warnings or closure were not available or were of such quality that we could not include their results. The effect was measured on injury rates, on exposure to physical workload and on compliance with regulation, with a follow-up varying from one to four years. All studies had serious limitations and therefore the quality of the evidence was low to very low. The injury rates in the control groups varied across studies from 1 to 23 injuries per 100 person-years and compliance rates varied from 40% to 75% being compliant.

The effects of inspections were inconsistent in seven studies: injury rates decreased or stayed at a similar level compared to no intervention at short and medium-term follow-up. In studies that found a decrease the effect was small with a 10% decrease of the injury rate. At long-term follow-up, in one study there was a significant decrease of 23% (95% confidence interval 8% to 23%) in injury rates and in another study a substantial decrease in accident rates, both compared to no intervention.

First inspections, follow-up inspections, complaint and accident inspections resulted in higher compliance rates compared to the average effect of any other type of inspections.

In small firms, inspections with citations or with more penalties could result in fewer injuries or more compliance in the short term but not in the medium term.

Longer inspections and more frequent inspections probably do not result in more compliance.

In two studies, there was no adverse effect of inspections on firm survival, employment or sales.

Qualitative studies show that there is support for enforcement among workers. However, workers doubt if the inspections are effective because inspections are rare and violations can be temporarily fixed to mislead inspectors.

Authors' conclusions

There is evidence that inspections decrease injuries in the long term but not in the short term. The magnitude of the effect is uncertain. There are no studies that used chemical or physical exposures as outcome. Specific, focused inspections could have larger effects than inspections in general. The effect of fines and penalties is uncertain. The quality of the evidence is low to very low and therefore these conclusions are tentative and can be easily changed by better future studies. There is an urgent need for better designed evaluations, such as pragmatic randomised trials, to establish the effects of existing and novel enforcement methods, especially on exposure and disorders.

PLAIN LANGUAGE SUMMARY

Inspections to prevent occupational diseases and injuries

In most countries, government-related inspectors check if workplaces comply with regulation, such as WorkSafeBC in British Columbia in Canada, the Occupational Health and Safety Administration (OSHA) in the USA or the Labour Inspectorate in other countries. Inspections are costly and do not reach all workplaces. It is unclear how effectively these inspections reduce occupational diseases and injuries.

To review the evidence on the effect of inspections we searched for studies until January 2013.

We found 23 studies. Two studies were randomised controlled trials with 1414 workplaces. Fifteen non-randomised studies analysed injury rates of firms obtained from large administrative databases. Six studies with more than 340 participants in total reported on the opinions of workers or employers.

Two studies randomly allocated inspections or no inspections to workplaces. After one year follow-up the non-fatal injury rate in one study and the frequency of physical overload in the other study were still similar in both study groups. Another five similar but lower

quality studies had inconsistent results at short and medium-term follow-up. Two other non-randomised studies found that after more than three years inspections decreased injuries and accidents by 23% compared to no inspections and there was no effect on the firms' productivity.

Specific inspections resulted in higher compliance rates. Inspections with penalties could result in fewer injuries and more compliance in the short term in small firms. Longer inspections and more frequent inspections probably do not result in more compliance.

Two studies did not find a harmful effect of inspections on firm lifetime or employment.

Qualitative studies showed that there is support for enforcement among workers. However, workers doubt if inspections are effective because they are rare and violations can be temporarily fixed to mislead the inspectors.

We concluded that inspections decrease injuries in the long term but probably not in the short term. The evidence is of low to very low quality because the results across studies are inconsistent and studies are observational and do not take into account other factors that could affect the results. In addition, the magnitude of the effect is uncertain because it varies from a 3 to 23 per cent decrease in injury rates. Because the quality of the evidence is low, future studies can easily change our conclusions. There is an urgent need for large-scale randomised trials to evaluate different types of inspection methods on exposure, disorders and injuries.

SUMMARY OF FINDINGS FOR THE MAIN COMPARISON *[Explanation]*

Inspection compared to no intervention for preventing occupational diseases and injuries

Patient or population: firms potentially subject to inspection
Setting: verification of compliance with occupational health and safety legislation
Intervention: inspection
Comparison: no intervention

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participating firms (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk	Corresponding risk				
	No intervention	Inspection				
Fatal and non-fatal injuries in RCT, short-term follow-up WC claims Follow-up: mean 21 months	**Moderate**		**RR 1.04** (0.9 to 1.21)	1402 (1 study)	⊕⊕◯◯ **low**[1,2]	
	41 per 1000	**43 per 1000** (37 to 50)				
Fatal and non-fatal injuries in CBA, medium-term follow-up WC claims Follow-up: mean 24 months	**Moderate**		**RR 0.87** (0.75 to 1.02)	818 (1 study)	⊕⊕◯◯ **low**[1]	
	31 per 1000	**27 per 1000** (23 to 32)				
Fatal and non-fatal injuries in CBA, long-term follow-up WC claims Follow-up: mean 48 months	**Moderate**		**RR 0.77** (0.64 to 0.92)	818 (1 study)	⊕⊕◯◯ **low**	

	31 per 1000	**24 per 1000** (20 to 29)			
Fatal and non-fatal crashes, ITS-level Crash data Follow-up: mean 36 months	The median level of fatal and non-fatal crashes was **2.99 crashes per 100 trucks**	The mean level of fatal and non-fatal crashes in the year after the intervention was **2.42 standard deviations lower** (2.88 to 1.96 lower)		6200 (1 study)	⊕⊕◯◯ **low**
Fatal and non-fatal crashes, ITS-slope Crash data Follow-up: mean 36 months	The median fatal and non-fatal crashes was **2.99 crashes per 100 trucks**	The trend of fatal and non-fatal crashes after the intervention was **0.89 standard deviations lower** (0.98 to 0.8 lower)		6200 (1 study)	⊕⊕◯◯ **low**

The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CBA: controlled before-after study; **CI:** confidence interval;**ITS:** interrupted time series; **RCT:** randomised controlled trial; **RR:** risk ratio; **WC**: Workers' Compensation

GRADE Working Group grades of evidence
High quality: Further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: We are very uncertain about the estimate.

[1]Compliance with inspections unclear; no blinding.
[2]Wide confidence interval.

BACKGROUND

Occupational health and safety legislation and regulation is often regarded as the backbone of the management of health and safety risks at work and has a strong focus on primary prevention of hazards. We will use regulation to refer to both regulation and to legislation. In the USA, extremely high rates of injury and occupational diseases led to the conviction that regulation was needed to control hazards for health and safety at work. The introduction of the Occupational Safety and Health Act in 1970 in the US was meant to "assure as far as possible every working man and woman in the nation safe and healthful working conditions" (Viscusi 2005). However, it is not easy to translate this aim into operational terms. In reality, a workplace entirely free of risk is an illusion, as there may always be a very small risk, and risks are inherently connected to human behaviour. Therefore, it is more sensible to view the aim of occupational safety and health (OSH) regulation as inducing desired management behaviours so that companies have policies in place to optimally control health and safety risks at work. Acceptable level of risk and optimal control are concepts whose definitions vary from country to country. It is obvious that it is not only technical possibilities that will define what is deemed optimal. In the end, one must balance the costs of controlling risks against the benefits of preventing serious risks or fatalities (Shapiro 1997; Viscusi 2005).

Despite reform of OSH regulation, the general idea behind enforcement of OSH laws has remained more or less unchanged for decades. Governments introduce regulation to ensure health and safety at work. Legislation gives a legal basis for enforcement, to obtain compliance and to change the way employees and employers conduct themselves in relation to occupational health and safety. This is exemplified in the Stresa declaration on workers' health, signed by the Advisory Committee of the Global Network of World Health Organization (WHO) Collaborating Centres for Occupational Health. In the declaration, legislation is described only in combination with its enforcement, and weak legislation is seen as one reason for hazardous workplaces (WHO 2006). Also, international strategies as formulated by the International Labour Organization and the Sixtieth World Health Assembly call for regulations, occupational health standards, collaboration and appropriate level of enforcement, as well as workplace inspections to protect and promote health (ILO 2004; WHO 2007). Legislation and its enforcement is claimed to "provide good opportunities for improving the health of workers and promote a culture of health and safety at work" (WHO 2006). This suggests that regulation, closely linked to enforcement, continues to have an important role to play.

There is, however, little evidence that regulation enforcement tools reduce the incidence of occupational diseases or injuries. Viscusi 2005 states that the introduction in 1970 of the Occupational Safety and Health Act in the US with its related enforcement did not change the trend of injury rates that already had been declining for decades. In a recent review of the introduction of OSH regulation in the construction industry, three US studies found neither an effect on injury rates immediately after the introduction of the regulation, nor a beneficial change in the trend of injury rates over time (van der Molen 2007). On the other hand, rates of specific exposure to chemical substances have also been declining in recent decades. Other authors maintain that OSH regulation is responsible for this decline (Kromhout 2000). It remains difficult to disentangle the effects of introducing regulation and enforcing it. The interpretation of changes in trends over time remains difficult if there are no disruptions in the trend associated with specific policy measures.

One systematic review of incentives for improving occupational health and safety found evidence that inspections and actual citations and penalties reduced injuries (Tompa 2007). A recent report of the US Government Accountability Office reviewed OSHA's monitoring policies and concluded that there was insufficient evaluation of the effectiveness of its enforcement tools (US GAO 2013). On the whole, then, the existing evidence is at best somewhat tentative.

Description of the condition

The burden of occupational injuries and fatal work-related diseases is still large worldwide (Concha-Barrientos 2005; Hämäläinen 2009). Dust-related lung diseases and injuries like falls from heights continue to cause many fatalities every year. Especially in emerging economies like China, India and Brazil, rates of injuries and occupational diseases are still unacceptably high. More than 350,000 workers die annually due to unintentional occupational injuries, more than 90% of these deaths occur among men, and more than half of those men work in the WHO South-East Asia and Western Pacific regions (WHO 2009).

Description of the intervention

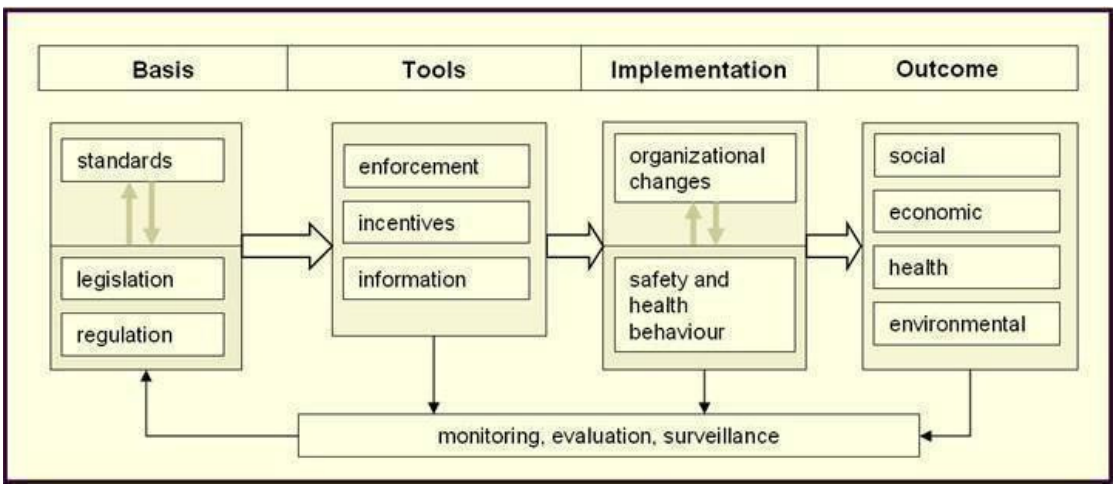
Regulation is used by governments around the world to protect workers against health and safety risks at work. A government can propose new legislation which needs to be passed as law by a legislative body, for example a parliament. Governments can also issue more detailed regulations to enforce the legislation, although these do not always have the status of law. Enforcement of OSH regulation is conducted in most countries by special government enforcement agencies. Additionally, there are many agencies which help to implement OSH standards and regulations, such as social insurance agencies, private insurers or certifiers of management systems. Only specific agencies have the power to enforce compliance, however. In many European countries the so-called 'labour inspectorate' is responsible for OSH legislation enforcement. In the United Kingdom, for example, it is the responsibility of the

Health and Safety Executive. In the US it is the Occupational Safety and Health Administration's responsibility. Even though the agencies' names vary, their duties and instruments are essentially the same. The agencies monitor the implementation of regulation at existing and future workplaces, for example when planning a new factory. Agencies arrange workplace inspections, and audits of companies' health and safety policies. The inspectors should encourage compliance with OSH regulation and enforce compliance if needed. The type and scope of inspections and penalty in case of non-compliance depends on the power which is given by law to the inspector and its agency. This can vary across countries and so the effect of the intervention might vary across countries too. In most cases, however, when inspectors find violations of the law, they can punish the violator immediately by issuing a warning, an order to comply with the law, a citation or a monetary penalty. In addition, some agencies can commence prosecution in court. If indicated it is also possible to (temporarily) close down machinery, departments or the whole company to prevent further non-compliance and to immediately remove workers from the identified hazard(s). In general all penalties like warnings, citations or firm closure can only be imposed after an inspection has taken place. Inspections can also result in giving information or consultation with the aim of resolving the deviation from the law. In addition to these negative incentives, most enforcement agencies also use positive incentives to induce compliance, such as rewarding excellent compliance, or exempting companies from inspections when they engage experts or consultants to inspect their workplaces (Gray 1993; Gunningham 2007).

How the intervention might work

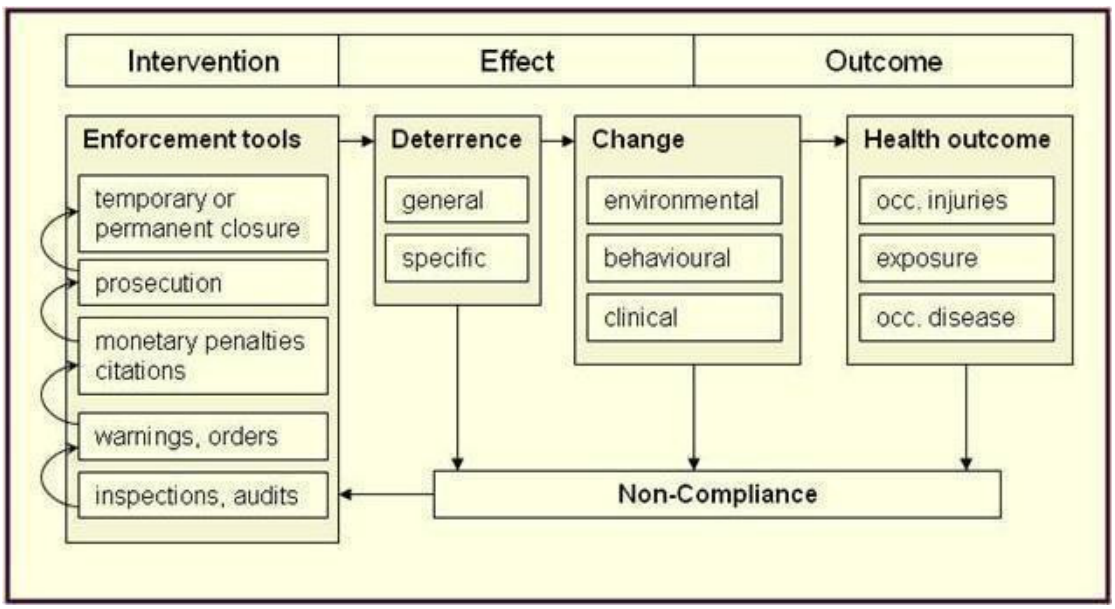
OSH regulation aims to promote safety and health at the workplace. Solutions to improve compliance with regulations are various, and enforcement is only one. We present a description and conceptual configuration of the content of applicable interventions in Figure 1. It is an adapted variation of the framework for policy implementation to promote diet and physical activity (WHO 2008). The model shown in Figure 1 helps to explain how enforcement is related to similar interventions and how social, economic, health and environmental benefits are thought to be achieved. It shows that OSH regulations and standards aim to influence employees' and employers' safety and health behaviour using several tools. The behaviour change takes place on an individual and organizational level. Companies might change their safety policies by putting up signs in hazardous workplaces, or by investing in safety equipment. Individually, the worker and employer might change their safety and health behaviour by wearing hearing protection more frequently, or by following work-rest schedule guidelines. It is worth noting that this process is influenced by the workers' and employers' attitudes and beliefs, which are often understood as being integral parts of safety and health behaviour. Incentives, enforcement and information, also known as 'carrots, sticks and sermons', are all considered by governments to ensure compliance with occupational regulations. Monitoring, evaluation and surveillance are recommended during the whole process to provide possibilities for necessary modifications. We have to keep in mind that the framework at Figure 1 shows only legislative and regulatory interventions. Other solutions to improve health and safety at work are not included, such as market forces. Even so, leaving workplace health and safety risks to be mitigated by market forces alone is not regarded as realistic by economic experts (Viscusi 2005).

Figure 1. Effects of occupational safety and health regulation enforcement tools



We present a theoretical model of how enforcement works and influences health and safety at the workplace in Figure 2. Enforcement is thought to have two slightly different effects. Most importantly, enforcement should lead to general deterrence, or what can be called a primary preventive effect. Another effect of enforcement is that it should lead to specific deterrence; a decrease in the recurrence of violations among those found violating the law and consequently punished. These effects are supposed to be much smaller than the general effect of deterrence (Shapiro 1997). For overall deterrence to be effective, the risk of punishment should be sufficiently severe for employers to infer that it pays to comply with the law or, as Shapiro 1997 puts it, "today's temptation is outweighed by tomorrow's punishment". Costs of violating OSH regulations which are below the cost of compliance might not lead to compliant behaviour. On the other hand, enforcement which is too strict or perceived to be unfair can lead to resentment and a dismissive attitude towards the regulator; this can be regarded as a serious side-effect of this intervention. To prevent this, Ayres and Braithwaite have recommended a "responsive regulatory approach" based on a judicious mix of co-operation and enforcement applied appropriately to the specific situation (Braithwaite 2006; Braithwaite 2007a; Braithwaite 2007b; Braithwaite 2011).

Figure 2. Effects of specific occupational safety and health regulation enforcement tools



A pyramidal approach to increase the likelihood of enforcement succeeding has also been recommended (Shapiro 1997). Regulatory actions would begin at the base of the pyramid by attempting to coax compliance by persuasion. Giving information and consultation following an inspection fits into this pyramidal approach. The next step is to issue a warning letter and, if this fails, impose administrative penalties. Further up the pyramid the regulator could employ criminal prosecution or temporarily shut down the entity. The ultimate sanction would be permanent shutdown of the entity.

One of the major problems with enforcement is the magnitude of the task. Inspections can never cover all workplaces. It has been calculated that in the US the rate of inspection per entity is about once in 100 years (Viscusi 2005). Targeting of inspections on companies that are likely to violate the law is hampered by a lack of prognostic data to enable their identification (Alper 2009). Other authors have argued that to increase enforcement effectiveness, labour unions should play a larger role in enforcement (Lierman 2010).

Why it is important to do this review

OSH enforcement tools such as penalties and prosecutions are common in all countries. It is important to know to what degree

monetary penalties, inspections or other enforcing activities influence workers' health and safety. However, there is considerable uncertainty as to what is the most effective and efficient approach. The one systematic review that has evaluated the effectiveness of OSH enforcement tools (Tompa 2007) is already several years old and did not use Cochrane methodology to locate and synthesise studies.

OBJECTIVES

To assess the effects of occupational safety and health regulation enforcement tools for preventing occupational diseases and injuries.

METHODS

Criteria for considering studies for this review

Types of studies

When enforcing regulation it is usually not feasible to randomise study participants, even though technically it would be possible. Legal and practical constraints will probably prevent randomisation but it is conceivable that these constraints could be overcome by using a cluster-randomised design. We included studies that randomise participants (randomised controlled trials (RCTs)) as well as the following non-randomised study designs: controlled before-after studies (CBAs) and interrupted time series (ITS). Given the general decrease of injury and exposure rates over the past 50 or so years, we believe that the study designs included should be able to control for trends over time that cannot be ascribed to the interventions. Without such caution, it would be difficult to make inferences from studies.

CBAs, otherwise known as prospective cohort studies or quasi-experimental studies, are easier to perform than randomised controlled trials, taking into account that the intervention is carried out at the group level, and they still have reasonable validity. We defined CBA studies as prospective or retrospective studies in which measurements of the outcome are available both before and after the implementation of the intervention for both the intervention and control group, and in which the outcome is measured at the same moment in time for both intervention and control group.

ITS 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 or 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 compensate for the lack of a control group to a certain extent (Ramsay 2003).

Further, we included study designs that are popular in economics, called panel studies. Usually they are based on data that are available in an existing database. We included panel studies using longitudinal outcome data for multiple entities (more than one firm or workplace) measured at least at two points in time. This could be either a so-called balanced panel or unbalanced panel. A balanced panel includes longitudinal data where the outcome is measured for every participant at every point in time. An unbalanced panel includes longitudinal data where the outcome is measured for every participant at least at two points in time and not necessarily at every point in time. That means an unbalanced panel has missing observations. The data are then analysed using regression analysis. The result of the regression analysis of the panel study shows the outcome as the difference between a control and an intervention group (Stock 2007). We further required panel studies to measure the outcome as a change over time. This meant the regression analysis had to include a so-called time lag variable. If this was not the case we excluded the study. The type of regression analysis used in panel studies can vary across studies (e.g. logistic regression, logit model or negative binomial regression). We included studies regardless of the type of regression analysis used. Study authors could also include various additional variables in the regression model beside the outcome and time lag variable. We treated those variables as adjustment for confounders which was not an inclusion or exclusion criteria but part of the 'Risk of bias' assessment.

We also included studies that reported employers' and employees' attitudes, opinion or beliefs on enforcement tools in order to be able to better explain the results of the review. We included studies regardless of the methods used and participants included as long as the opinions of workers or employers were reported or analysed separately from the other groups (as e.g. labour inspectors). In the context of this review we refer to all studies reporting opinions, attitudes or beliefs as qualitative studies. We excluded publications with the opinion of only one participant such as opinion papers.

Types of participants

We included studies in which the intervention has been targeted either at whole companies or at individual workplaces.

For the qualitative studies, we included workers and employers or supervisors.

Types of interventions

We included all types of enforcement activities by any agency officially assigned by the government to enforce compliance with OSH regulation, not connected to the actual company that is inspected. We categorised regulation enforcement interventions to consist of one or more of the following components.

- Inspections and audits which were defined as any kind of monitoring activity to check a company's compliance with OSH law or regulations. These inspections could be random, programmed, a follow-up or take place following an event. We excluded studies on the effects of voluntary consultations.
- Warnings or orders intended to change work practices, management policies, worker behaviour, equipment etc. in order to comply with law or regulations. These could be spoken or written.
- Citations or monetary penalties.
- Prosecution.
- Closure of the firm either temporary or permanent.

Types of outcome measures

We included studies that measured the effect on either exposures to health or safety hazards or on rates of occupational diseases and injuries. We included studies only if the effect was measured at the level of the workplace or firm. We excluded studies that measured the effects of workplace inspections at an aggregated level of an industry or a state. We excluded such studies as they cannot differentiate the effect of the intervention from other changes in the population. This is known as an ecological fallacy.

Primary outcomes

- The degree of exposure to health or safety hazards. This could be measured as being compliant with regulation or as a change in exposure after the enforcement activity. Both were considered to be equally valid.
- Incidence rates of injuries or occupational diseases.

To measure the effect of enforcement on prevention of relapses we used:

- the recurrence of the measures above.

For qualitative studies we used:

- negative attitudes of employers or workers towards one or more enforcement tools.

Secondary outcomes

We also included indicators of preventive activity such as investment in health and safety, training and education, and observable policy changes.

Search methods for identification of studies

Based on the inclusion criteria, we developed a search strategy for the various electronic databases. We took the following essential concepts of the inclusion criteria to develop the search string: firms and workplaces on the one hand and enforcement tools on the other. We used only those two concepts to ensure that the search would be sensitive enough to identify all relevant studies regardless of study design. We set no restrictions on language, publication year or publication status. The date of the last search was 1 January 2013.

Electronic searches

We searched the following electronic databases from the first day of entries:

- Cochrane Central Register of Controlled Trials (CENTRAL, *The Cochrane Library*);
- MEDLINE (PubMed);
- EMBASE (embase.com);
- CINAHL (EBSCO);
- PsycINFO (Ovid);
- OSH update (www.oshupdate.com);
- HeinOnline (www.heinonline.org);
- Westlaw International (www.westlaw.com);
- EconLit (EBSCO); and
- Scopus (scopus.com).

We present the search strategy for MEDLINE (PubMed) in Appendix 1. We translated this strategy for use in the other databases (Appendix 2).

Searching other resources

We searched also the following web sites: European Union Senior Labour Inspectors Committee (http://ec.europa.eu/social/main.jsp?catId=148&langId=en&intPageId=685), Occupational Safety and Health Administration (USA) (www.osha.gov/) and Health and Safety Executive (UK) (www.hse.gov.uk). We screened the reference lists of the included articles for additional studies and contacted researchers in the field for further published or unpublished studies.

Data collection and analysis

Selection of studies

We divided the references that we retrieved among the authors in such a way that each reference was assessed in duplicate. If the two assessors indicated the eligibility of a reference, we retrieved the full-text article for further assessment. If only one assessor was of the opinion that a reference should be included, we consulted a third author (JV or CM) before ordering the full-text article. Two authors (CM and JV) independently checked full-text articles for eligibility. We resolved disagreements by consensus or by involving a third author (RP, SC or TM).

Data extraction and management

Two persons of the team of review authors extracted data independently from each of the included studies using a standard form. We extracted the following information from RCTs, CBAs, ITS and panel studies and where appropriate also from qualitative studies:

1. design and country of the study;
2. characteristics of participants (number of participants, inclusion and exclusion criteria and other domains according to study eligibility, and 'Risk of bias' assessment criteria);
3. type and time of intervention in control and intervention group;
4. outcomes (outcome measures, data sources, follow-up time, adverse events and results);
5. funding source and conflict of interests; and
6. for studies using regression analysis, the type of regression analysis, the number of variables included in the regression model, and the definition and measurement method of four confounder variables (firm size, type of work, inspections prior to the intervention, baseline injury rates).

We extracted the outcome from all quantitative studies as reported by the authors and requested additional data when necessary. If studies reported similar outcomes measured in different ways, we only used the one we deemed the most valid. For example, for injuries we chose lost-time injuries over overall injuries because the risk of under-reporting for lost-time injuries is smaller than for overall injury claims (Azaroff 2002).

From qualitative studies we extracted attitudes, opinions and beliefs of employers and employees towards the intervention by extracting themes and citations as reported in the studies.

Where possible, we resolved discrepancies in the data extraction by consensus. Otherwise we involved a third author (JV).

Assessment of risk of bias in included studies

Two review authors (CM and JV) assessed the risk of bias of all included studies independently. We used a consensus method when disagreements occurred. The Characteristics of included studies table includes the 'Risk of bias' assessments for quantitative studies.

We evaluated the risk of bias in RCTs, CBAs and panel studies with the checklist developed by Downs and Black (Downs 1998). We only used the items on internal validity of the checklist and not those on reporting quality or external validity. The instrument has been shown to have good reliability, internal consistency and validity. The 13 items of the checklist include the domains of the 'Risk of bias' tool recommended in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011): random sequence generation, allocation concealment, blinding of participants, blinding of assessors, incomplete outcome data and selective reporting. We modified the answers to the questions of the checklist so that they fit the 'Risk of bias' tool as implemented in RevMan 2011 by using 'high', 'low' or 'unclear' instead of 1 or 0 as proposed by the checklist authors. We specified the criteria of the checklist for our review according to the following scheme.

1. We considered self reports of non-fatal injuries or occupational diseases to have a high risk of bias in outcome measurement. We based this decision on evidence of under-reporting as presented in Azaroff 2002.
2. We considered the risk of bias due to participants or assessors not being blinded as high if the outcome was exposure or self reported.
3. We judged studies to have a low risk of bias for confounders if a minimum of three out of the following four baseline characteristics were similar in intervention and control group or were adjusted for in the analyses: firm size (small, i.e. fewer than 250 employees and big firms, i.e. 250 employees or more), type of work (physical or mental effort), pre-intervention inspections and injury rates.
4. In our opinion just inspecting a workplace does not ensure compliance with standards. We therefore judged participants to be compliant with the intervention or inspection only if the study authors specifically described that participants followed the instructions or the orders given by the inspector or if the fines were paid or if a firm was closed.

We judged the overall quality of an RCT, CBA or panel study at low risk of bias if all following seven items were rated as at low risk of bias: blinding of outcome assessor, follow-up, outcome measure, selection bias (population), selection bias (time), adjustment for confounding and incomplete outcome data.

With ITS studies we used the eight 'Risk of bias' criteria presented by Ramsay 2003, that are based on earlier work of the Cochrane Effective Practice and Organization of Care (EPOC) Review group. We used the data from one CBA study to perform an ITS analysis and we used the 'Risk of bias' checklist for ITS for this study (Chen 2008).

With studies using companies' self reports of injuries or occupational diseases, we intended to assess the possibility of under-reporting as a result of inspections or announcement of inspections (Gray 1993; Haviland 2012; McQuiston 1998; Robertson 1983) but there was no possibility of finding out if this was the case, so we refrained from doing so.

With qualitative studies, we used the supplemental handbook guidance available online from the Cochrane Qualitative Research Methods Group (Hannes 2011). We adapted a critical appraisal tool from the JBI QARI checklist, originally created by the Joanna Briggs Institute (JBI 2011) and the checklist developed by Verbeek 2004 (Appendix 3). We assessed the risk of bias of the qualitative studies in three domains.

1. Consistency and neutrality of method and reporting
2. Credibility of method and subjects
3. Transferability of analysis and conclusions

We answered the questions either with Yes, No or Unclear, with a Yes indicating low risk of bias and a No indicating high risk of bias. Three review authors (AA, CM, RP) independently assessed

the overall risk of bias of the qualitative studies based on a judgement of whether the items could have influenced the outcome. We considered the questions 4, 5, 7, 8 and 9 more likely to influence the outcome than questions number 1, 2, 3 and 6. We used the following rating system to judge the overall quality of a study.

- **high** quality if at least four YES in the first group and two YES in the second group;
- **moderate** quality if at least two YES in the first group and three YES in the second group; and
- **low** quality if less than one YES in the first group.

Measures of treatment effect

For all study types, we scored diseases and injuries as unfavourable and compliance as a favourable outcome. Thus, an increase in injuries is unfavourable but an increase in compliance is favourable. We used risk ratios (RR) as measures of treatment effect for dichotomous outcomes and mean differences (MD) and their standard deviations (SD) for continuous outcomes. With CBA studies, we planned to put the outcome measurements in the data tables both at baseline and follow-up to ensure that baseline imbalances were taken into account. However, all CBAs had been analysed with regression analyses and we could only put the resulting RRs or ORs into the data tables using the generic inverse variance method.

With panel studies that used regression analysis and presented their results as a beta-value, we transformed the beta-coefficients into relative risks if the analyses involved a log transformation (Burstyn 2010; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Kniesner 2004; Levine 2012; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001) and into a mean difference if there was no such transformation (Robertson 1983) according to the methods described by Austin 2011. If studies used compliance rates as an outcome and provided odds ratios as the effect measure, we transformed the odds ratios (ORs) into risk ratios (RRs) because due to the high prevalence of compliance (40% to 75%) ORs would overestimate the RRs. We used the formula provided by Zhang 1998 for the adjustment.

With ITS studies, we extracted data from original papers and reanalysed them according to recommended methods for analysis of ITS designs for inclusion in systematic reviews (Ramsay 2003) and as also recommended for evaluation of law studies by 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 auto-correlation between individual observations. For each study, we fitted a first-order autoregressive time-series model to the data using a modification of the parameterization of Wagner 2002. Details of the mode specification are as follows: Y= ß0 + ß1time + ß2 (time-p) I (time > p) + ß3 I (time > p) + E, E ~ N (0, s2). 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 auto regressive process (AR1). The parameters ß have the following interpretation: ß1 is the pre-intervention slope, ß2 is the difference between post and pre-intervention slopes, and ß3 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 then standardised the data from ITS studies in order to obtain effect sizes by dividing the outcome and standard error by the pre-intervention standard deviation, as recommended by Ramsay 2003. Thus we have two separate outcomes for an ITS study: the effect size for the short-term change in the level of outcome due to the intervention which can be interpreted as an additive effect, and the effect size for the long-term change in the trend in time or change of slope indicating an increasing effect of the intervention. Chen 2008 reported time series on eight different groups of firms. Three of these had sufficient data points. We first analysed these different time series as described above and then combined them in a meta-analysis. We finally put the results of the meta-analyses in the data tables as one pooled result.

Unit of analysis issues

No study included in this review employed a cluster-randomised design and no study compared several active interventions with one control intervention. Thus there were no unit of analysis issues.

Dealing with missing data

We contacted authors of seven studies to try to obtain missing data (Burstyn 2010; Foley 2012; Geminiani 2008; Haviland 2012; Kniesner 2004; Levine 2012) and we did obtain additional data for three (Foley 2012; Kniesner 2004; Levine 2012). We did not succeed in contacting the author from one study (Kemmlert 1994). We obtained risk ratios and standard errors for one study (Foley 2012) and standard errors for two (Kniesner 2004; Levine 2012). We could not obtain data from the authors of the other studies. For one study (Haviland 2012), we calculated the standard error from the P values given in the article according to the methods described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

Assessment of heterogeneity

We considered the risk of bias and the effects of RCTs, CBAs, ITS and panel studies to be different.

We assessed similarity between studies by assessing whether the intervention could reasonably be expected to yield similar effects or to work similarly in the various populations, control conditions, follow-up times and outcomes. To this end, we considered the following major and minor sources of heterogeneity:

1) Major sources of heterogeneity: type of intervention, type of outcome and follow-up time.

We considered all types of inspections (e.g. random, programmed) to be similar. Similarly, we deemed all types of penalty to be similar interventions. All control conditions with no intervention were deemed similar. We further considered the following categories of outcomes to be different: exposure, occupational diseases and injuries. We divided the type of injuries into fatal and non-fatal and occupational diseases into acute or chronic. We assumed that it would take a considerable time before inspections would lead to a change in outcome. Therefore, we categorised follow-up as short-term up to one year, medium-term from one to three years and long-term with a follow-up longer than three years.

2) Minor sources of heterogeneity: inspection and penalty type, type of work, company size and previous inspections.

We considered the various types of inspections and penalties, such as inspections after complaints with citations or unannounced inspections without penalties, as a minor source of heterogeneity and analysed these interventions in subgroups. We also made subgroups if workers had mostly physical work, such as construction workers, or when tasks involved mostly mental effort, such as in office workers. We also made subgroups according to firm size as small (fewer than 250 employees) and big firms (250 employees or more) (Gray 2005a; Haviland 2012). Further, we made subgroups if study participants had been subjected to previous inspections, as pre-intervention experiences can reduce the effects of enforcement (Levine 2012).

We assessed statistical heterogeneity by means of the Chi^2 test, as implemented in the forest plots in RevMan 5 (RevMan 2011). We used a significance level of $P < 0.10$ to indicate if there is a problem with heterogeneity. In addition, we quantified the degree of heterogeneity using the I^2 statistic where an I^2 value less than 40% indicates heterogeneity that is unimportant, 30% to 60% indicates a moderate degree of heterogeneity, between 50% and 90% indicates substantial heterogeneity and 75% to 100% considerable heterogeneity.

Assessment of reporting biases

We reduced the effects of reporting bias by including studies and not articles. When articles reported on the same study, we included data only once or from several articles as far as it was necessary. We considered publications of panel studies as the same study if the data sets used were from the same time and the same source with similar inclusion criteria and interventions. Table 1 gives an overview of the characteristics of all articles reporting on panel studies included in this review.

We prevented location bias by searching multiple databases and we prevented language bias by not excluding articles based on language. We checked for outcome reporting bias as part of the 'Risk of bias' assessment.

Data synthesis

We present results separately for different study designs (RCTs, CBAs, ITS, panel studies and studies with qualitative outcomes). We pooled data from quantitative studies we judged to be clinically homogeneous with RevMan 5 software (RevMan 2011). To combine effect sizes, we used the general inverse variance method in RevMan 5 (RevMan 2011).

Most of the studies were clinically heterogenous and, therefore, we applied a random-effects model for the meta-analysis. All estimates include a 95% confidence interval (CI).

We used the GRADE approach as described in the *Cochrane Handbook for Systematic Reviews of Interventions* and as implemented in the GRADEPro 3.2 software (GRADEpro 2008) to present the quality of evidence and 'Summary of findings' tables.

We present qualitative results separately from results of quantitative studies. We used a narrative summary to present the results of the qualitative studies.

Subgroup analysis and investigation of heterogeneity

We could not include more than 15 studies per comparison, therefore we could not explore heterogeneity in a meta-regression. We explored heterogeneity using subgroups based on the factors considered to be minor sources of heterogeneity. Those were type of inspection and penalty, type of labour, company size and previous inspections.

Sensitivity analysis

We could not conduct a sensitivity analysis because we judged all included studies to have a high risk of bias.

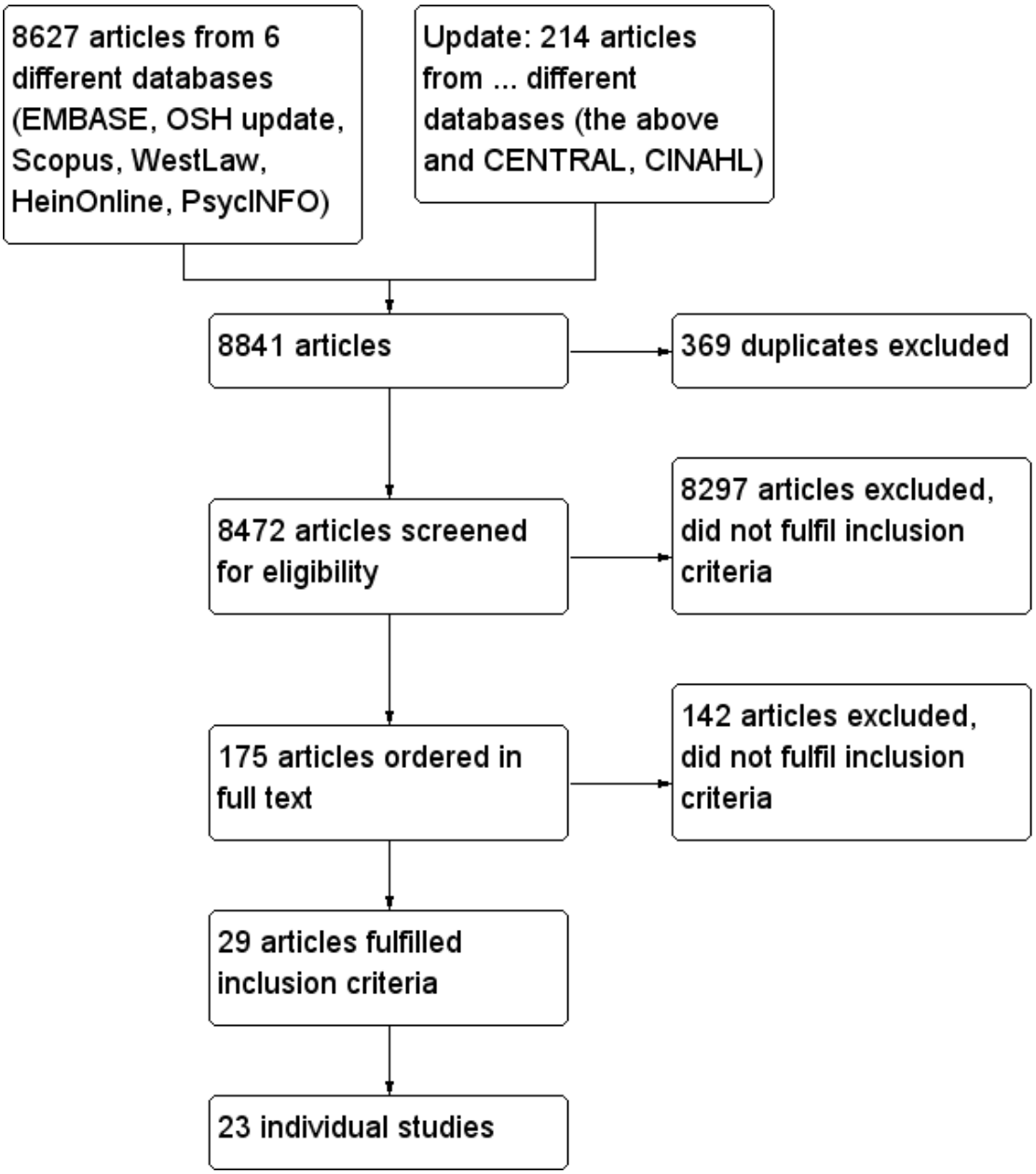
RESULTS

Description of studies

Results of the search

Our search yielded 8841 references. After deletion of duplicates 8472 references remained. Screening for eligibility resulted in 175 references to be assessed in full text. After full-text assessment, 28 articles fulfilled the inclusion criteria. Of those, 18 articles reported single studies, one article reported three studies, one article reported two study arms and eight articles described already included studies. This resulted in 23 studies being included in this review (see Figure 3).

Figure 3. Flow diagram.



Included studies

Study designs

We included two RCTs (Hogg-Johnson 2011; Kemmlert 1994), two CBAs (Levine 2012; Nelson 1997), one ITS (Chen 2008) and 12 panel studies (Burstyn 2010; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Kniesner 2004; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001).

We also included six studies presenting opinions, attitudes or beliefs towards one or more of the enforcement tools. Because there is no regular place for these types of studies in a Cochrane Review, it was most convenient to put the references under the heading of studies awaiting classification. Two qualitative studies reported the results of observations (Bordas 2001; Gray 2006). Two studies reported survey outcomes (Geminiani 2008; Guidotti 1996). One study used a focus group with semi-structured interviews and observations (Gillen 2004) and one study used interviews (Mayhew 1999).

Interventions

The interventions were poorly described in all included studies. Studies merely reported a type of inspection and a type of penalty following the inspection but none of them described the process and focus of the interventions. It was unclear how the inspections were performed, what the inspectors' instructions were and if or what kind of change process in the inspected firms was set in motion by the inspections.

Quantitative studies

The following characteristics of inspections were available in the included studies.

Announcement of inspections: The announcement in advance of inspections could change its effects because firms can prepare themselves for the coming inspection. One study reported that inspections were without prior warning to the workplace or firm (Hogg-Johnson 2011) and another study reported on inspections that were announced in advance (Kemmlert 1994). Fifteen studies did not specify if the inspection was announced but for five studies from the USA we inferred that the intervention included announced inspections because the intervention included inspections after complaint and follow-up or programmed inspections (Gray 2005a; Gray 2005b; Gray 2005c; Weil 1996; Weil 2001).

Procedure for choosing firms: Firms can be randomly chosen for inspections or inspections can be targeted at high-risk firms using work injuries or injury claim rates for selection. The effect of a targeted inspection triggered by a recent injury might be different compared to an inspection of a randomly chosen workplace. Two studies specified that workplaces were randomly chosen for an inspection from a pool of high-risk firms (Hogg-Johnson 2011; Levine 2012). We inferred from the labelling of the inspections as programmed inspections, inspections after complaints, and accident and fatality investigations that these were also targeted inspections in four studies (Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012).

Inspection type: Eight studies included any type of inspection (Foley 2012; Kniesner 2004; Levine 2012; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001) but in two of these studies the authors studied also follow-up or complaint inspections (Weil 1996; Weil 2001). One study included almost any inspection, specified as programmed, referral, fatality and accident investigations, inspections after complaints and other unscheduled evaluations (Nelson 1997). Four studies included only programmed and complaint inspections (Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012). In one study, the inspection is described as either comprehensive or focused on particular occupational health and safety hazards (Hogg-Johnson 2011) and another study describes it as focused on musculoskeletal stress factors (Kemmlert 1994). In one study (Chen 2008), truck companies are inspected as part of a compliance review for motor carriers. Burstyn 2010 studied inspections according to the profile of the inspectors that were following a proactive approach.

Inspection intensity: We assumed that the number of inspections, the hours spent by the inspectors at the workplace and the number of penalties issued would influence the effect, where we would expect a bigger effect with more intense inspections. Three studies described the intensity of the inspection as the number of penalties following an inspection (Kniesner 2004; Weil 1996; Weil 2001), the number of inspections (Weil 1996; Weil 2001) or the duration of the inspection (Weil 1996; Weil 2001).

Inspector characteristics: Another aspect of the interventions are the characteristics of the inspector performing the inspection. One study described the style of the inspector and divided inspectors either into following a proactive approach or following a reactive approach (Burstyn 2010).

Inspection penalties: Two studies evaluated the effect of more versus less penalties or if the inspections were only with penalties (Weil 1996; Weil 2001). Four studies included only inspections with penalties (Haviland 2012; Hogg-Johnson 2011; Kniesner 2004; Nelson 1997). Eleven studies included inspections regardless of whether penalties were issued (Burstyn 2010; Chen 2008; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Kemmlert 1994; Levine 2012; Robertson 1983; Smith 1979a; Smith 1979b).

Type of penalty: None of the studies included prosecution of firms and none included *persons* violating standards, legislation or regulation. Eight studies described either fines, warnings, orders or citations and one study also included closure orders (Kniesner

2004). Two studies included only citations (Foley 2012; Robertson 1983). One study included citations and fines (Nelson 1997). One study included fines or closure orders (Kniesner 2004). Two studies include only orders (Burstyn 2010; Hogg-Johnson 2011). Two studies (Chen 2008; Kemmlert 1994) included warnings and orders. Nine studies did not name the type of penalty included in the study (Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Levine 2012; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001).

Reasons for penalty: Three studies reported the reasons why penalties were given. In Kemmlert 1994 the penalties were given if the workplace showed musculoskeletal stress factors and in Weil 1996 if violations occurred with machine-guarding and hand-held tools safety standards. The other studies did not specify the violations of occupational health and safety regulations.

Intensity of the penalty: Only one study specified that the orders were both voluntary and formal compliance orders (Burstyn 2010). None of the included studies described the intensity of the penalty in terms of amount of money, duration of closure etc.

Process

Only one study described the process of the intervention (Kemmlert 1994). The labour inspectors followed an ergonomic workplace checklist to identify musculoskeletal stress factors and received special training beforehand. The penalties could be warnings or orders in case of observed insufficiencies. The authors used the intervention also to measure the outcome at the end of the study, so the control group received the same intervention at the end of the study.

Co-interventions

One study (Hogg-Johnson 2011) reported co-interventions. Both groups, the control group and the enforcement group, received consulting activities (described as consulting, technical, general, contact or calls and other). The control group received also inspections and investigations from the ministry of labour (22% of participants). One study (Foley 2012) stated having excluded firms that had had both consultation and enforcement visits. With 16 studies, we assumed that co-interventions such as consultation or training or both were possible but were not reported (Burstyn 2010; Chen 2008; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Kemmlert 1994; Kniesner 2004; Levine 2012; Nelson 1997; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001).

Control group intervention

In nine studies the control group received no intervention (Chen 2008; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Hogg-Johnson 2011; Kemmlert 1994; Levine 2012; Nelson 1997). Two studies compared firms receiving an inspection early in the year to firms receiving the intervention later in the year (Smith 1979a; Smith 1979b). The idea was that the inspections early in the year should have resulted in an effect on the outcome whereas the inspections later in the year would have had no such effect. One study used firms receiving no intervention or an intervention without penalties as a control group (Haviland 2012). Two studies used either firms receiving any other type of inspection with or without further penalties, or firms receiving fewer inspections, or firms receiving a lower number of penalties as a control group (Weil 1996; Weil 2001). In Kniesner 2004 the control group received fewer inspections with penalties (either fines or closure orders). Another study compared inspections with citations to inspections without citations (Robertson 1983). One study compared inspections from reactive inspectors to inspections from inspectors following a proactive approach (Burstyn 2010).

Qualitative studies

None of the six included qualitative studies described the enforcement intervention of interest in detail (Table 2). Studies neither reported if the data collection took place before, during or after a specific intervention. Three of six qualitative studies included any occupational health and safety regulatory enforcement activity (Bordas 2001; Gillen 2004; Guidotti 1996). Two studies focused on any type of inspection (Gray 2006; Mayhew 1999). Mayhew 1999 reported that the inspections were with or without penalties but did not further specify the type of penalties. In one study (Geminiani 2008) the inspectors themselves were the intervention of interest with any enforcing activity.

Outcomes

Quantitative studies

Thirteen studies evaluated injuries or disease. One of those reported accidents which included injuries, and one reported days away from work. Eleven studies reported non-fatal injuries, and two studies included fatal and non-fatal injuries (Chen 2008; Levine 2012). Four studies reported exposure as outcome.

Nine of the 13 studies measuring injuries used lost-workday injuries. These injuries were measured as the number of registered lost-workday injury claims per person (full-time equivalent) one year after the intervention (Foley 2012; Hogg-Johnson 2011; Smith 1979a; Smith 1979b), per firm per year (Gray 2005a; Gray 2005b; Gray 2005c; Robertson 1983) or per firm and quarter (Kniesner 2004).

One study used fall injury claim rates per firm including only compensable claims with a minimum of four days of lost work time because of a fall (from elevation, platform or ladder, fall from piled matter, fall on stairs, fall into openings, fall from roof or fall to lower level) (Nelson 1997).

One study reported days away from work caused by disease or injury where we inferred that these were per firm per year (Haviland

2012).

One study reported the total number of motor carrier crashes per group per year (Chen 2008). The accidents involved a truck or a bus of motor carriers operating in the United States and resulted in at least one fatality, injury or vehicle towed away from the scene as a result of disabling crash damage.

One study used all types of injury claims per year per firm (Levine 2012).

Four studies reported exposure measured as compliance with a standard. Burstyn 2010 predicted the number of inspector visits needed to resolve non-compliance. Two studies predicted the change in probability of having zero violations with any safety standard (Weil 2001) or with machine-guarding and hand-held tools safety standards (Weil 1996). One study measured exposure as reduced workload which was achieved if the harmful situation reported in the injury report on musculoskeletal injuries did not exist any more (Kemmlert 1994).

Qualitative studies

Qualitative outcomes were workers' and employers' opinions and beliefs regarding inspections (Geminiani 2008), regarding OSHA enforcement (Bordas 2001), managers' opinions and beliefs about OSHA enforcement (Gillen 2004), observed reaction by workforce towards planned inspections (Gray 2006), opinions on enforcement of occupational health and safety standards and if it increases cost or time to complete a job (Guidotti 1996), and opinions and beliefs about the impact of inspections on OSH (Mayhew 1999).

Measure of treatment effect

Most of the quantitative studies measured the effect as change of the outcome per year and either per firm (Chen 2008; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Levine 2012; Robertson 1983) or per full-time equivalent (Foley 2012; Hogg-Johnson 2011; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001). In one study (Weil 2001) the effect is also measured as change from the 6th inspection to the mean value of the following inspections (7th and more) per firm. One study measured the change for each of four quarters and the authors estimated the equilibrium multiplier effects of one year per firm (Kniesner 2004).

One study reported the total number of events before and after the intervention (Kemmlert 1994). Another study predicted the number of interventions needed to achieve zero violations per firm (Burstyn 2010).

Time period

Seven studies included data from after 2000 but most of the studies analysed older data. Studies used data from the 1970s up to 2008 with two to 10-year time periods. Five studies analysed data sets starting in the 1970s and covering three to seven years (Gray 2005a; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996).

Four studies included data sets from the 1980s with two to 15-year coverage (Gray 2005b; Kemmlert 1994; Kniesner 2004; Weil 2001). Six quantitative studies used data from the 1990s with two to 10-year time coverage (Chen 2008; Foley 2012; Gray 2005c; Haviland 2012; Levine 2012; Nelson 1997). Two studies used data collected after 2000 and covering three (Burstyn 2010) and six-year time periods (Hogg-Johnson 2011).

Three qualitative studies analysed data from the 1990s (Bordas 2001; Guidotti 1996; Mayhew 1999). One study collected data in 2000 (Gillen 2004) and two studies did not report the time of the study (Geminiani 2008; Gray 2006).

Participants

Quantitative studies

Fifteen quantitative studies included in total 146,004 firms. The number of firms included ranged from 3 to 113,441 with a median value of 1219 firms. Six quantitative studies did not report the number of firms or participants (Burstyn 2010; Gray 2005a; Gray 2005b; Gray 2005c; Smith 1979a; Smith 1979b).

The type of industry reported in quantitative studies was the manufacturing industry (Burstyn 2010; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Hogg-Johnson 2011; Kniesner 2004; Robertson 1983; Smith 1979a; Smith 1979b), construction industry (Nelson 1997; Weil 2001) and woodworking industry (Weil 1996). Three quantitative studies included participants from unknown types of industries (Foley 2012; Levine 2012; Kemmlert 1994).

All quantitative studies included firms except for one study which included individual workplaces (Kemmlert 1994). Most of the studies included participants engaged in mostly physical work (Burstyn 2010; Chen 2008; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Hogg-Johnson 2011; Kniesner 2004; Levine 2012; Nelson 1997; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001). One study included mixed type of work (Foley 2012) and one study did not report the type of work included (Kemmlert 1994).

The firm size included was not reported for nine studies (Burstyn 2010; Chen 2008; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Hogg-Johnson 2011; Kemmlert 1994; Kniesner 2004). Three studies included any firm size (Nelson 1997; Smith 1979a; Smith 1979b), two studies included only big firms (Robertson 1983; Weil 2001), one study included only small firms (Haviland 2012) and two studies included mostly small firms (Levine 2012; Weil 1996). Most of the studies did not report if the participants had inspections prior the study intervention. Three studies reported that the participants included did not have inspections two years (Hogg-Johnson 2011; Levine 2012) or one year prior the study intervention (Foley 2012).

Qualitative studies

The number of participants reported in qualitative studies ranged between 22 and 150 with a mean of 77.5 participants per study (Table 2). Two qualitative studies did not report the number of participants (Bordas 2001; Gray 2006).

Qualitative studies included participants from the manufacturing industry (Gray 2006), construction industry (Geminiani 2008; Gillen 2004; Mayhew 1999), logging industry (Bordas 2001) and sand/oil industry (Guidotti 1996).

All qualitative studies included worker and employers except for two studies including workers only (Guidotti 1996; Mayhew 1999). The type of work was mostly physical in all but one study, which recruited office workers (Geminiani 2008). The firm size was not reported for four studies. One study included only small firms (Bordas 2001) and another study included any firm size (Gillen 2004). Whether participants experienced inspections was only reported in one study (Bordas 2001) and the experience was mixed (some yes some no).

Countries

All studies were from high-income countries (Australia, Canada, Sweden, USA), as defined by World Bank 2012, except one that was from South Africa, which is considered upper middle-income.

We included 17 quantitative studies from three countries. Most of the quantitative studies were conducted in the USA (Chen 2008; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Haviland 2012; Kniesner 2004; Levine 2012; Nelson 1997; Robertson 1983; Smith 1979a; Smith 1979b; Weil 1996; Weil 2001). Two studies were conducted in Canada (Burstyn 2010; Hogg-Johnson 2011) and one in Sweden (Kemmlert 1994).

We included six studies reporting qualitative outcomes from four countries. Two studies are from Canada (Gray 2006; Guidotti 1996), two from the USA (Bordas 2001; Gillen 2004), one from Australia (Mayhew 1999) and one from South Africa (Geminiani 2008).

Excluded studies

We did not include qualitative or quantitative studies which analysed only the implementation of a directive or new law without assessing the enforcing intervention (e.g. Adams 2007; Attfield 1992; Lissner 2011). We further excluded studies where the outcome did not allow for conclusions about the enforcement intervention. We excluded one study which measured the effect of a campaign with multiple interventions, including incentives or other non-enforcing actions, as a whole (Mancini 2005). We also excluded studies if the enforcement of regulations was not focused on occupational health and safety but concerned public health, for example the enforcement of a smoking ban (Baron-Epel 2012) or tattoo regulations (Raymond 2003). We excluded cross-sectional studies and studies using panel data without including a time lag variable in the regression analysis (Boden 1985; Ko 2010; Smitha 2001). We excluded studies if the unit of analysis was not the individual workplace or firm but on an aggregate level, either sub-industry, industry or national level (Auld 2001; Viscusi 1979). We excluded opinion papers or qualitative studies with only one participant (e.g. Brown 2003) and studies not presenting qualitative outcomes from employees or employers (as e.g. opinions from inspectorates only) (Niskanen 2013).

Risk of bias in included studies

The risk of bias in RCTs, CBAs and panel studies was as follows (Figure 4; Figure 5):

Figure 4. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

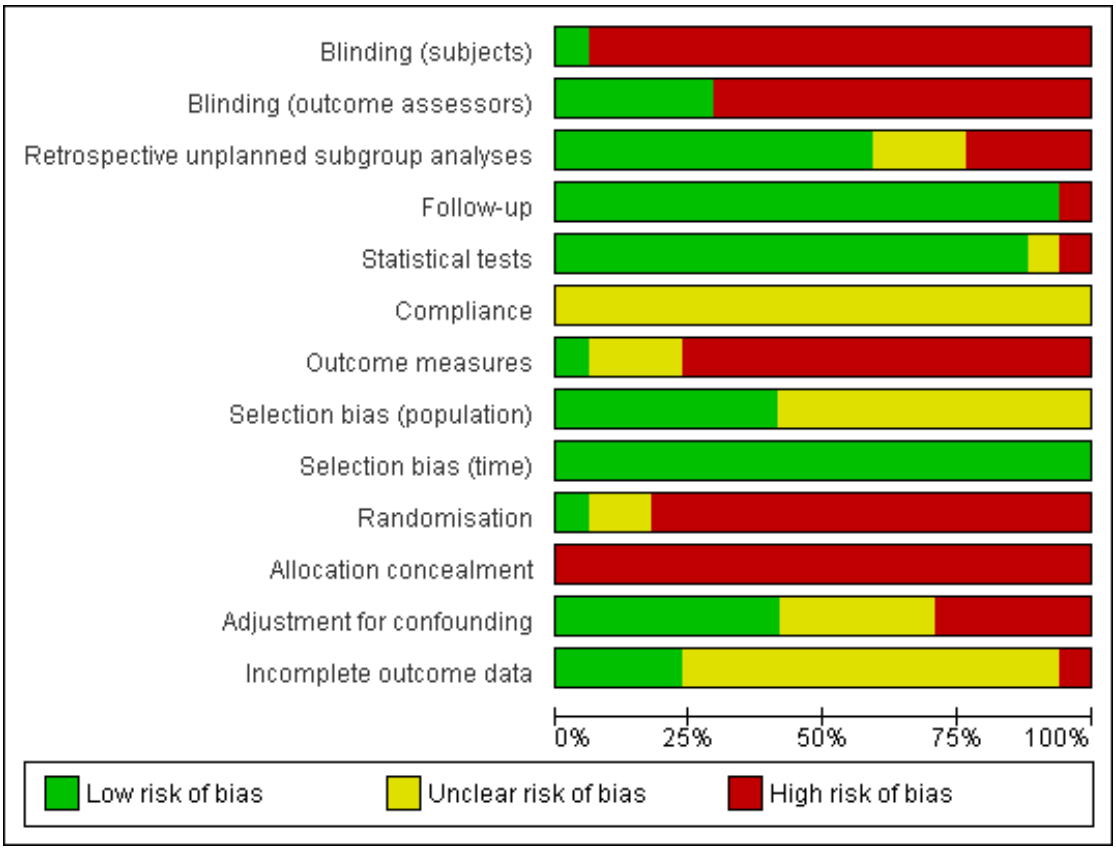
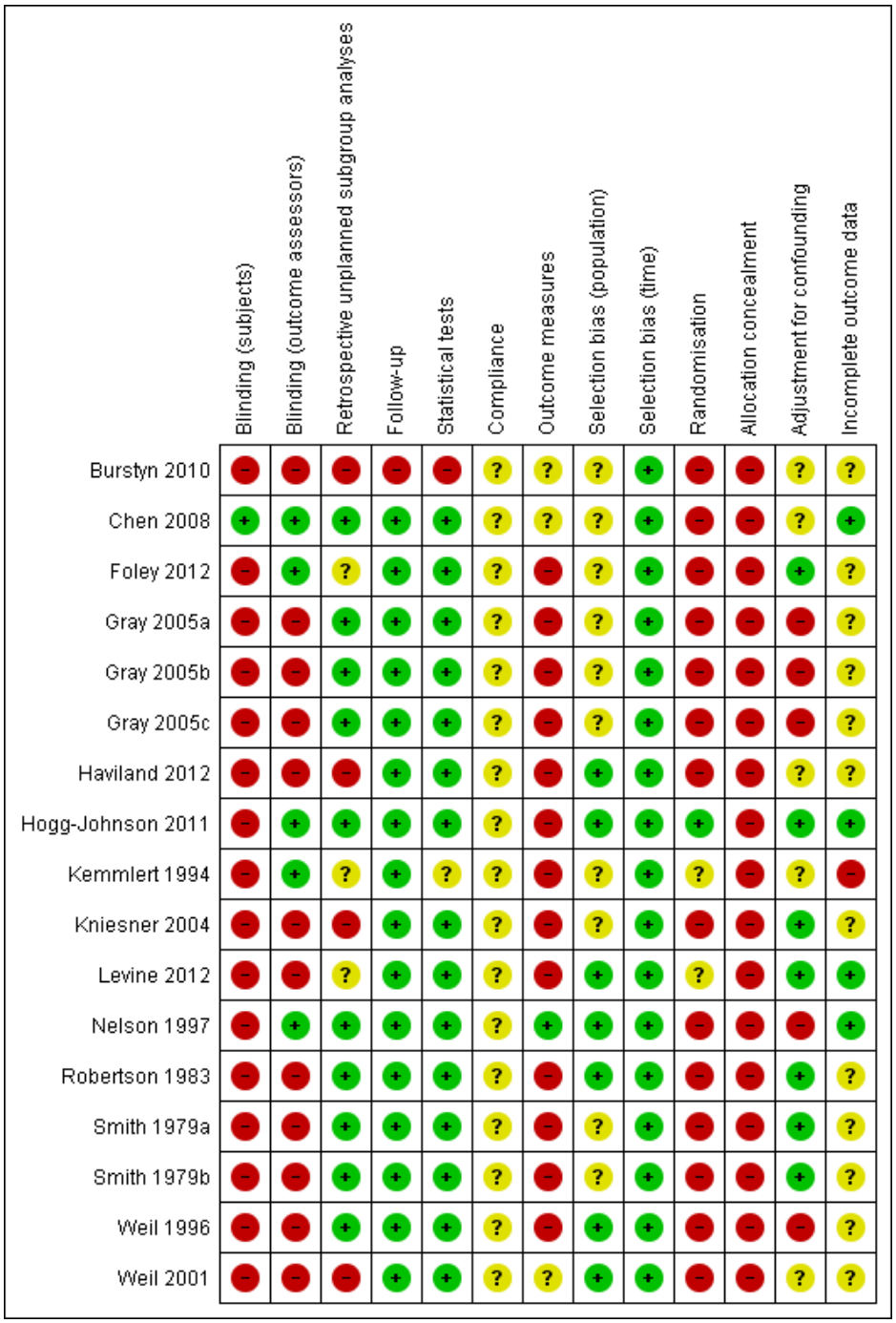


Figure 5. 'Risk of bias' summary: review authors' judgements about each risk of bias item for each included study.



Allocation

We judged the risk of selection bias in three different domains: participants, time and allocation concealment.

Population

We considered studies including participants from the same population as low risk of bias. If studies compared participants from different regions or industries the outcome can be biased as different standards apply or other circumstances differ. We judged the risk of bias low for one RCT (Hogg-Johnson 2011), two CBAs (Levine 2012; Nelson 1997) and four panel studies (Haviland 2012; Robertson 1983; Weil 1996; Weil 2001). One RCT (Kemmlert 1994) and eight panel studies (Burstyn 2010; Foley 2012; Gray 2005a; Gray 2005b; Gray 2005c; Kniesner 2004; Smith 1979a; Smith 1979b) did not report details and we judged the risk of bias as unclear.

Time period

All studies compared participants selected from the same time period. We judged the risk of bias low for all included studies.

Allocation concealment

We judged the risk of bias high for all studies. In both RCTs the allocation of the intervention was not concealed and the inspectors could influence which workplaces to inspect. In all non-randomised studies the inspectors could always freely choose. We consider inspectors more likely to inspect workplaces that they would consider to be more hazardous or to leave out workplaces they remember to have visited earlier. The decision by the inspector can also be motivated by other unknown factors.

Blinding

We assessed the blinding of participants and outcome assessors in two domains: outcome assessor and participants.

Outcome assessor

The outcome assessors did not know which participants received the intervention in two RCTs (Hogg-Johnson 2011; Kemmlert 1994), one CBA (Nelson 1997) and one panel study (Foley 2012). For all other non-randomised studies the risk of bias was high.

Participants

Participants were not blinded to the intervention, either in RCTs nor CBAs nor panel studies. This is caused by the fact that the intervention takes place at the workplace and includes the observation of the workers and processes. It is impossible to hide the intervention from workers and employers. In one study (Chen 2008), we judged the risk of bias as unclear because the outcome was measured as the number of accidents reported in police crash reports. The outcome would not have changed regardless of whether the participants would have known if they received the intervention. In case of injuries, the outcome measure in all studies was based on injury self reports. It is possible that participants not blinded to the intervention were less or more likely to report or claim injuries after inspections. This would not apply for fatal injuries but none of the included studies measured the effect on fatal injuries only. Exposure was measured as violation of occupational health and safety regulations. The outcome is most likely biased by the knowledge of the participants that an inspection takes place. The risk of bias would be less if the inspections were unannounced, but none of the included exposure studies analysed the effect of inspections without prior warning to the workplace.

Incomplete outcome data

It was difficult to assess if outcome data were complete. In the non-randomised studies, authors used data sets and did not report if there was any loss of data.

Selective reporting

We did not assess selective reporting as this was not part of our checklist and all studies were non-randomised and none of the studies reported having published a protocol.

Other potential sources of bias

Compliance

We considered that, for example, the visit of an inspector at a workplace does not ensure compliance with the intervention. Therefore, we could only judge this if the authors reported information about whether or not the participants followed the instructions or the orders given by the inspector, if the fines got paid or whether or not the firms actually closed down. None of the studies assessed the compliance with the intervention and we judged the risk of bias unclear for all included studies.

Outcome measure

Studies relying on self reports of non-fatal injuries or occupational diseases are considered to have a high risk of bias. We considered lost-time injuries as more reliable outcome measure when based on self reports than all injuries or injuries without days away from work (Azaroff 2002). We judged all but one study unclear or high risk of bias.

Adjustment for confounders

For 60% of the studies it was either unclear if there were baseline differences or if the intervention effect was adjusted for confounders in the analysis or it was clear that there were unadjusted differences.

Risk of bias in ITS

For Chen 2008, we analysed the data as an ITS. We judged that it was unclear if the intervention occurred independent of other changes. The study had sufficient data points and the tests performed were correct. We judged that it was unlikely that the intervention had affected data collection and that the outcomes were assessed blind for the intervention. It was unclear if the data covered the complete data set and how reliable the outcome data were.

Risk of bias in qualitative studies

Consistency and neutrality

Connection of methods

We judged half of the included studies of high quality as three studies had a clear connection between philosophical perspective, methodology, objectives, data collection, representation and analyses of data (Gillen 2004; Gray 2006; Mayhew 1999) (Table 3). The other three studies were of unclear or poor quality. Two of those studies did not report the methods applied and did not present the themes derived from the data (Bordas 2001; Guidotti 1996). One study presented a review of the literature but did not use the theoretical background to explain the attitudes of contractors or the differences with inspectors' attitudes (Geminiani 2008).

Other potential bias

Only one study clearly reported the context of the study and discussed the influence in the conclusion (Gray 2006). The other five studies were of poor or unclear quality. One study rarely described the context and the inclusion in the discussion was missing (Geminiani 2008). We judged the other four studies unclear, as a description of the context was missing.

Reporting

Only one study was of high quality (Gillen 2004). The reporting was clear and coherent. The other five studies did not report sampling methods, recruitment conditions, inclusion and exclusion criteria, the method of data collection, or the description of the derivation of themes (Bordas 2001; Geminiani 2008; Gray 2006; Guidotti 1996; Mayhew 1999) and we judged them as of poor quality.

Credibility

Recruitment

Participation in the studies was voluntary. We judged the quality as high for two studies because they applied various methods of outreach including monetary incentives (Gillen 2004) or chose people from the national phone register (Guidotti 1996). The four other studies did not report the recruitment process (Bordas 2001; Gray 2006) or did not explain drop-out and exclusion criteria (Geminiani 2008; Mayhew 1999).

Participants

None of the included studies described both participants and workplace characteristics and no study considered these characteristics in the discussion of findings and their implications. We judged the quality as low.

Ethical approval

We considered two studies of high quality (Gillen 2004; Gray 2006). The other studies did not report ethical approval. We could not judge if they fulfil ethical standards and we assessed the quality as unclear for these four studies.

Transferability

Intervention specific

Only two studies reported what the intervention of interest was (Geminiani 2008; Gray 2006). We judged the quality poor if it was not specified what type of enforcement tool was evaluated or if the authors analysed different tools in the same category (Bordas 2001; Gillen 2004; Guidotti 1996; Mayhew 1999).

Triangulation

None of the studies was of high quality. Two studies used a range of methods but one study did not report if both sources were used for the data analysis and if similar conclusions could be drawn (Bordas 2001). One study did not report the conclusions (Gillen 2004). The other four studies did not apply more than one method and the quality is rated as low (Geminiani 2008; Gray 2006; Guidotti 1996; Mayhew 1999).

Reliable data and conclusion

We judged only one study as high quality as the authors reported that two researchers independently open coded the interviews and derived themes by consensus (Gillen 2004). We judged one study as low quality as the data were interpreted by one author only and the findings were not validated by the participants (Gray 2006). Four studies provided too little information and the quality is unclear.

Overall rating of risk of bias

We judged all RCTs, CBAs and panel studies as being at serious risk of bias (Figure 5). We assessed the ITS study as being of low risk of bias.

Based on the criteria listed in Table 3, we considered one qualitative study (Gillen 2004) as of moderate quality and the remaining five studies as low-quality studies (Bordas 2001; Geminiani 2008; Gray 2006; Guidotti 1996; Mayhew 1999).

Effects of interventions

See: **Summary of findings for the main comparison** Inspection compared to no intervention for preventing occupational diseases and injuries

1. Inspection versus no intervention

1.1 Outcome injuries, short-term follow-up

At short-term follow-up, the results of seven studies provide no evidence for or against an effect of inspections on fatal or non-fatal injuries.

Randomised controlled trials (RCTs)

One study (Hogg-Johnson 2011) used a randomised design and found no effect of unannounced inspections versus no intervention in workplaces with mostly physical work with a risk ratio (RR) of 1.04 (95% confidence interval (CI) 0.90 to 1.21) for non-fatal injuries. The workplaces had not been inspected up to two years prior to the current inspection but the size of the workplaces was not reported (Analysis 1.1). The authors also analysed if there was a negative effect of firms going out of business but there was no difference between study arms

Controlled before-after (CBA) studies

Levine 2012 found no effect of inspections versus no inspections on both fatal and non-fatal injuries (RR 0.98, 95% CI 0.88 to 1.09) at short-term follow-up with a CBA design. The inspected firms were of small size, had mostly physical work and had not been inspected up to two years prior to the current inspections (Analysis 1.2).

Another CBA study compared inspections versus no inspections but reported insufficient data to be included in the meta-analysis (Nelson 1997). The authors reported fall injury rates before and after inspections in the intervention group and fall injury rates at similar times for the control group. The change in the incidence before-after was 1.02 fall injuries/100 person-years for the intervention group and 0.03 for the control group. The likelihood of experiencing a reduction in injury claim rate yielded an odds ratio (OR) of 2.3 with $P < 0.0001$ in the authors' logistic regression analysis.

Panel studies using regression analysis

Three studies with a panel study design evaluated the effect of inspections on non-fatal injuries versus no intervention (Foley 2012; Smith 1979a; Smith 1979b). The results in Foley 2012 were presented separately for the fixed and non-fixed firms and the authors used different control groups for each comparison. In the meta-analysis of the studies inspections decreased the injury rate by 8% (RR 0.92, 95% CI 0.89 to 0.95; $I^2 = 0\%$). None of the studies clearly reported the type of inspections nor the penalties. The type of work was mostly physical in two studies and mixed in the other. Workplaces had not been inspected one year prior to the current inspection in one study (Foley 2012) and in the other studies this was unknown (Analysis 1.3).

1.2 Outcome injuries, medium-term follow-up

At medium-term follow-up, in four studies, there was no evidence of an effect of inspections on fatal or non-fatal injuries.

CBA studies

Levine 2012 also did not find an effect of inspections versus no inspections at medium-term follow-up (RR 0.87, 95% CI 0.75 to 1.02) (Analysis 1.4).

Panel studies

Analysis of various time-series of injury data found a decreasing effect of inspections versus no inspections at medium-term follow-up over different time periods (Gray 2005a; Gray 2005b; Gray

2005c). In a meta-analysis of these panel studies, there was a non-significant 3% decrease at the medium-term follow-up (RR 0.97, 95% CI 0.94 to 1.01) (Analysis 1.5).

1.3 Outcome injuries, long-term follow-up

At long-term follow-up, two studies provide evidence of a substantial decrease in injury and accident rates after firms have been inspected.

CBA studies

In the same study of Levine 2012, there was an effect of inspections at long-term follow-up with a RR of 0.77 (95% CI 0.64 to 0.92) (Analysis 1.6). The authors also analysed whether firms that were inspected had shorter survival time, less employment or fewer sales but this was not the case.

Interrupted time series (ITS) studies

Chen 2008 found a significant decrease of both the level (effect size (ES) -2.42, 95% CI -2.88 to -1.96) and slope (ES -0.89, 95% CI -0.98 to -0.80) of accident rates per year over time after firms had been inspected compared to firms that had not been inspected (Analysis 1.7; Analysis 1.8).

1.4 Outcome exposure, medium-term follow-up

At medium-term follow-up one study showed no effect of inspections on exposure.

RCTs

There were no studies that measured the effect of inspections on workplace exposures at the short or long-term follow-up. Kemmlert 1994, in a RCT, found no effect of inspections on reduction of physical workload at the medium-term follow-up (RR 0.88, 95% CI 0.59 to 1.32). The inspections were announced in advance and orders were given to remedy violations. The type of work was not reported (Analysis 1.9).

2. Specific inspections versus any other type of inspections

Outcome exposure, short-term follow-up

Panel studies

Two studies reported on the effect of four specific types of inspections versus any other type of inspection on compliance with orders at follow-up in a panel study design (Weil 1996; Weil 2001). There was a significant effect of an increase in compliance after follow-up inspections (RR 2.55, 95% CI 2.43 to 2.68), after complaint inspections (RR 1.18 95% CI 1.07 to 1.30) and after accident investigations (RR 1.24, 95% CI 1.12 to 1.37). The effect of complaint inspections was smaller and non-significant in small firms, however (Analysis 2.1).

3. Inspection with citations versus inspection without citations

Outcome injuries, short and medium-term follow-up

Panel studies

Robertson 1983 evaluated the effect of inspections with citations versus those without in a panel study. He found an effect in the short term (mean difference (MD) -23.6 injuries, 95% CI -41.7 to -5.5) (Analysis 3.1) but not at medium-term follow-up (MD -2.8, 95% CI -23.9 to 18.3) (Analysis 3.2).

4. Inspection with more penalties versus inspections with fewer penalties

Outcome injuries, short and medium-term follow-up

Panel studies

Haviland 2012 found that inspections with penalties led to a 7% decrease in injury rates in the short term compared to no inspections or inspections without penalties, but there was no effect at medium-term follow-up (OR 1.00, 95% CI 0.99 to 1.00) (Analysis 4.1; Analysis 4.2).

Kniesner 2004 examined the effects of inspections and reported no effect of inspections but, given the complicated analysis, we could not extract data to be used in meta-analysis.

Outcome exposure, short-term follow-up

Panel studies

Weil 1996 found an increase in compliance with standards with higher penalties compared to lower penalties in small firms but they did not find this in another study in big firms (Weil 2001) (Analysis 5.1).

Based on four studies, there could be an effect of the amount of penalties in the short term and in small firms, but this is not sustained in the longer term nor in big firms.

6. More inspections versus fewer inspections

Outcome exposure, short-term follow-up

Panel studies

Two studies evaluated the effect of the number and the order of inspections on the compliance with regulations (Weil 1996; Weil 2001). Both for big (RR 1.21, 95% CI 1.03 to 1.41) and small firms (RR 2.82, 95% CI 2.11 to 3.77), the first inspection had a bigger impact on compliance than subsequent inspections. Weil 2001 also compared the effect of inspections up to the sixth versus more than six and did not find a difference (RR 1.00, 95% CI 1.00 to 1.00) (Analysis 6.1; Analysis 7.1).

The same studies also evaluated inspections that were of longer duration compared to shorter inspections and found a non-significant increase in compliance in small firms but a decrease in compliance in big firms, which resulted in a 4% decrease in compliance when the inspections were longer (RR 0.96, 95% CI 0.94 to 0.99) (Analysis 8.1).

7. Attitude of inspectors

Outcome exposure, long-term follow-up

Burstyn 2010 studied whether the attitude or work style of inspectors influenced the outcome of inspections. There was no difference in injury rates after inspections by inspectors who were autonomy oriented versus inspectors that were coercive oriented (Analysis 9.1).

Grading of the evidence

All but two studies were observational and would thus be rated as low quality as a point of departure in the GRADE approach. Given the diverse nature of study designs that we included, we refrained from further refining the level of evidence as it was difficult to assess the precise risk of bias in the panel studies. For all comparisons and outcomes the evidence was based, almost exclusively, on observational studies and therefore we assessed the quality of evidence as low for all.

Results from qualitative studies

The results from studies reporting workers' or employers' opinions, beliefs or attitudes towards enforcement of occupational health and safety regulation shows positive opinions and beliefs as well as negative opinions and attitudes. Even though all included studies focused on different phenomena than our question of interest, all reported results important to our review.

Positive opinions and beliefs

Support for inspections

Two studies report positive opinions towards enforcement of occupational health and safety regulations (Bordas 2001; Gillen 2004). Both studies took place in the USA around the same time period (1998 and 2000). Bordas 2001 found positive beliefs towards inspections from one participant. He stated that the inspectors could help with safety. Nevertheless, this was connected to the opinion that the person actually preferred inspectors to stay away:

"prefer they stay away but they could possibly help with safety"

One study reported positive opinions of workers from the construction industry towards enforcement in general (Gillen 2004). The participants pointed out that the enforcement should be consistent as well as apply to every firm. Negative opinions of workers were reported if uniform enforcement was lacking but the study did not present citations.

"If they are going to make a regulation, and a good one, it should be enforced", " uniform enforcement. You don't feel like that being safe you are putting yourself to a disadvantage to your competitors"

The presence of Cal/OSHA was reported as difficult but if all firms were subject to uniform enforcement it was believed to be a positive occurrence. The authors conclude that many were in support of more effective enforcement but did not provide citations or themes for this result (Gillen 2004).

Usefulness of inspections and penalties

One Australian interview study from 1997 showed positive opinions of construction workers towards workplace audit or inspector visit. Other workers replied to find inspector visits, other type of visits, phone calls, audits and letters from the inspectorate only of some use. No citations were given. The authors report that a number of workers wanted the jurisdiction to crack down more stringently on unsafe demolition jobs (Mayhew 1999).

Enforcement good even if it increases the time or cost for the firms

Eighty per cent of the respondents to a Canadian telephone survey in 1992 expressed strong advocacy of vigorous enforcement of occupational health and safety standards even if it would increase the cost or the time to complete a job, or both (Guidotti 1996).

Negative opinions and attitudes

Effectiveness of inspectors and lack of inspections and penalties

Three studies found negative opinions about the effectiveness of inspectors conducting their duties and enforcing regulation. The

study participants reported that inspectors show a lack of presence and oversight and cause additional administrative burden. Study participants rarely experienced the inspectors visiting the work sites or following complaints, except if an accident occurs. Two of those studies included construction workers. One study included loggers.

"have never seen them", "only come if there is a death", "prefer they stay away" "yet another additional administrative requirement" (Bordas 2001, USA, 1998)

"We have had only one inspection in 5 years", "As far as the DoL (labour inspector) is concerned, they have only visited my site once in 2005", "I have phoned them in the past and gave them addresses of dangerous building sites, but nothing was ever done about it", "No contact at all", "Inspectors have never visited my premises", "Never seen inspectors", "Do they exist?", "They only show up in country areas when there is a serious accident" (Geminiani 2008, South Africa, time of study unknown)

"no one doing the right thing - no enforcement's … could give you 12 names where it is going wrong … Workplace Health and Safety just another government department taking money off us and not doing anything for us." (Mayhew 1999, Australia, 1997)

Mayhew 1999 reported further that the lack of enforcement activity in general, but also warning letters with subsequent workplace audit without further penalties, was evaluated as useless by construction workers. Geminiani 2008 also analysed the opinions regarding inspectors' knowledge and skills in the construction industry and the appropriateness of the checklist used during the inspections, but no themes and citations were reported in the study. We contacted the authors but did not receive additional information.

Workers' reaction towards inspections

One study reports the workers' reactions when health and safety inspections take place (Gray 2006, Canada, time of study unknown). The workforce was observed to create "Potemkin villages" which is used as a term to describe the creation of an illusion for the inspectors. Workers and supervisors behaved and reported to be in line with every occupational health and safety requirement as long as the inspection lasted. The creation of a local culture ("Potemkin villages") is seen as a negative outcome, but the "cause" of this outcome is not clearly expressed. It is unclear if any inspection would cause the behaviour or if it shows the lack of surprise inspections. The authors did not present conclusions on workers' beliefs and opinions about the intervention or its effectiveness.

DISCUSSION

Summary of main results

We found low-quality evidence in seven studies that inspections had inconsistent results at one to three years follow-up. They can lead to a decrease or they can result in a similar level of injury rates or exposure to health hazards as no intervention. At more than three years follow-up, two studies showed low-quality evidence of a substantial decrease in injuries and accidents after firms had been inspected compared to firms that were not inspected.

Compared to any inspection, first inspections, follow-up inspections, complaint and accident inspections resulted in higher compliance after the inspections.

Inspections with citations and penalties or with higher penalties could result in fewer injuries and more compliance in the short term but not in the long term, nor in big firms.

Longer inspections and more frequent inspections probably do not result in more compliance.

Even though one study included work stoppage as an enforcement tool, neither firm closure nor prosecution of firms was evaluated in studies. In studies that found a decrease in injury rates, the effect was usually small: up to about a 10% decrease in the injury rate.

We judged all studies to be at high risk of bias and thus the overall quality of the evidence presented is low to very low. The reporting of injuries or the assessment of compliance used in most studies are based on self report and the reporting could be influenced by the inspections.

Qualitative studies show that there is support for enforcement among workers. However, workers doubt if the inspections are effective because they see that inspections are rare and they observe that violations of health and safety standards would be temporarily fixed to mislead the inspectors.

Overall completeness and applicability of evidence

We put considerable effort in locating studies even though many of them were quite old. Studies were from various fields, such as econometrics, law and occupational safety, which made it difficult to locate them. It could be possible that we missed some older studies published in areas outside occupational health. However, given the low quality of most studies, the overlap of data sources and the small effect sizes, it is doubtful if this would have changed our results.

We felt that the various study designs could not be combined in a single meta-analysis. Studies also used different data and participants that made a comparison of the effects difficult, such as studying only small firms or only the construction industry. This resulted in a rather fragmented picture of the available evidence.

It is evident that randomised controlled trials are possible: even though we had anticipated that there might be none, we found two - one older and one more recent. Compared to the panel studies that use existing data and given the number of possible confounders, we believe that there is a strong case for using the RCT design to provide evidence for the effectiveness of enforcement

tools. Since many firms are randomly chosen to be inspected, it should not be too difficult to use a randomised study design.

The evidence that we found was for the most part from the US. There were hardly any studies from Europe and no studies from Asia or Latin America. Given the different cultures in regulation and safety issues, we believe that the evidence is especially applicable to North America. Due to the lack of description of the inspection process and the lack of knowledge of which factors in a work organisation especially set a process of prevention in motion, it is difficult to apply the results of studies to practice. It is difficult to say if inspections have a specific effect or that they have a more general deterrent effect. Apparently, given the differences in effect found for various types of inspections, it does make a difference how and why an inspection is carried out. More focused inspections yielded a better result in terms of compliance and injury prevention. It might be valuable to concentrate efforts in these types of inspections.

In the qualitative studies, we found support for enforcement but there were also opinions that this was not effective because inspections were too infrequent or the likelihood of being inspected was too low. Also, the possibility of setting up temporarily improved safety measures to mislead inspectors was an explanation for the lack of effects of enforcement.

The majority of the studies were from before the year 2000 which raises the question of whether their results are still applicable. Gray 2005a observed a decline in effect of inspections in more recent times. From the available evidence, it is unclear if there really is such a decline in effectiveness. More recent studies still report a beneficial effect of enforcement (Levine 2012). We did not observe structural differences between the older and newer studies. It is conceivable that the general attitude towards regulation has changed and that enforcement would thus be more or less difficult. However, the results of the review do not allow any conclusions here.

We felt that making changes in safety and health at work requires a long-term commitment of a work organisation and investment in workplace improvements or safety culture. Therefore, we expected enforcement to be more effective in the long term. Even though we do not know if this assumption about long-term investment holds, the results of studies included in this review show more convincing effects in the long term (after four years follow-up).

We evaluated the effect of inspections. They form only one element in a long chain that finally leads to fewer occupational injuries and diseases: the development of effective preventive measures, the implementation of these measures through legislation or regulation, and then the subsequent enforcement of these measures with inspections. It could be that inspections work especially well for some occupational safety and health problems, such as decreasing noise exposure or chemical exposure, but not for others, such as physical workload, where the preventive measures are less clear cut. Some authors ascribe the decrease in chemical exposure levels over time in the US and Western Europe to the combined effect of legislation and enforcement (Creely 2007). However, we did not find studies that evaluated the effects of inspections which focused on chemical exposures.

Quality of the evidence

We assessed the quality of the evidence as low to very low. Most studies were observational with a lack of standardised adjustment for confounders or prognostic factors. From the RCTs that we did find, it can be inferred that randomisation is possible. We believe that this substantially increases the quality of the evidence. The quality of the panel studies was especially low and the studies were in general badly reported. It was difficult to judge which variables had been entered into regression analyses and how they were coded (0 or 1) which made the interpretation of the studies very difficult. Authors also varied substantially in theoretical points of departure and in their preference for a certain regression model (probit, tobit, OLS). We believe that many of these problems could be overcome by applying a pragmatic RCT design.

Another major reason for the lack of quality was that study results could be biased because participants were not blinded and the injury rates were self reported. It would not be too difficult to manipulate the reporting of smaller non-fatal injuries without lost time. Most studies differentiated between fatal, non-fatal and non-fatal with lost time injuries. We judged studies as having a high risk of bias if the outcome was non-fatal injuries. Only for fatal injuries did we judge that there would be no such risk of bias, because it would be almost impossible to not report a fatal injury. Even though we intended to adjust for under-reporting or to take this into account, we had no means to do so and the only way to take this into account was in the 'Risk of bias' assessment.

Potential biases in the review process

A large number of the included studies are based on the panel study design. We had difficulties in assessing the effects of the studies because usually only the beta-coefficients were reported with P values. Because these are difficult to interpret, we transformed them into rate ratios. Even though we scrutinised studies for information about the regression models used, we were not always completely sure our interpretation was correct. We therefore refrained from transforming the results of one study (Kniesner 2004) which used Arellano and Bond regression. Panel studies neither reported if a balanced (complete follow-up of all participants) or an unbalanced (varying number of participants over time periods) design was used. It could be that the potential bias in these studies has been underestimated: we would have judged balanced studies at lower risk of bias then unbalanced studies.

Most of the studies were older but we used the results as if they still would be applicable. We do not know if and how the effects of enforcement have changed over time.

The follow-up of most studies was, given a probably lengthy process of preventive measures, relatively short with one-year follow-up. It could be that the non-significant results of studies can be explained by this relatively short follow-up time.

Agreements and disagreements with other studies or reviews

Tompa also reviewed the effects of regulation and its enforcement and concluded that general deterrence is less effective in reducing injury incidence and severity, whereas specific deterrence with regard to citations and penalties does indeed have an impact (Tompa 2007). However, general and specific deterrence were not well defined. It seems that the authors took a cross-sectional relation of inspections and injury rates as a specific deterrence and a later effect of inspections on injuries as a general deterrence effect. They also included different studies than we did and the method for study synthesis relied on a 'best evidence synthesis' and did not include a quantification of the study effects so that they could be combined. We believe that the results in our review are more realistic and are based on better qualification and quantification of the intervention effect. Nevertheless, Tompa's conclusion that there is limited evidence that inspections as such are effective is similar to ours. The US Government Accountability Office's report described a scenario of insufficient evaluation of the effectiveness of the US OSHA enforcement tools and offered practical recommendations for assessment of effectiveness and the monitoring of federal and state enforcement efforts (US GAO 2013). The key recommendations involve including outcomes in OSHA's own assessments of its enforcement initiatives, and making better use of data from its audits. This concurs with our observation that it is unclear how the inspections work or how they finally affect injury rates, occupational disease rates or exposures. The report mentions three studies with beneficial outcomes that are included in this review (Foley 2012; Haviland 2012; Levine 2012).

Fines or a higher level of fines can lead to lower injury rates in the short term but not in the long term and not in big firms. There were no studies on prosecution of firms. The studies were in most part from the US. All evidence was rated as of low quality. Qualitative research shows support among workers for enforcement but also skepticism about its effectiveness. However, most studies were old and possibly do not represent current opinions.

Implications for research

To better understand the effect of occupational safety and health regulation enforcement interventions, better evaluation studies such as pragmatic randomised controlled trials, are needed in which firms or workplaces are randomised to specific enforcement tools or to regular inspections. Instead of using existing observational data, there is a need for experimental studies. It is important to ensure that prognostic factors such as previous inspections at the same workplace, firm size, baseline injury rates and type of work are equally distributed in intervention and control group, as can be achieved through randomisation.

Given that enforcement agencies like OSHA in the US or the Labour Inspectorates in Europe usually work on a national scale, it should be possible to randomise a sufficiently large group of firms. The specific enforcement approach should be clearly defined and described, especially the type and scope of the inspection, as well as the resulting prevention measures taken by firms. The control group should consist of regular inspections and there it would also be necessary to monitor or survey how these are conducted and what happens at the workplaces. The outcome should be measured at sufficiently long follow-up, such as three years after the intervention has been carried out. The outcome should preferably be based on objective injury or exposure data, such as those collected by insurance firms or for reasons other than because of the inspections. In addition, as secondary outcomes, data should be collected about productivity and firm lifespan, because these are believed to be adverse effects of inspections and often used as arguments to counter enforcement policies.

AUTHORS' CONCLUSIONS

Implications for practice

Inspections as an enforcement tool have inconsistent effects in the short term but they do decrease injury rates after more than three years follow-up. Specific types of inspections result in higher compliance rates than was achieved on average with inspections.

ACKNOWLEDGEMENTS

We thank WorkSafeBC from Canada for proposing the review title and generously providing funding for this review. We also thank Jani Ruotsalainen from the Cochrane Occupational Safety and Health Group for his helpful comments and edits and Megan Prictor for copy editing the text.

REFERENCES

References to studies included in this review

Burstyn 2010 *{published data only}*

Burstyn I, Jonasi L, Wild TC. Obtaining compliance with occupational health and safety regulations: a multilevel study using self-determination theory. *International Journal of Environmental Health Research* 2010;**20**(4):271–87.

Chen 2008 *{published data only}*

Chen GX. Impact of federal compliance reviews of trucking companies in reducing highway truck crashes. *Accident Analysis and Prevention* 2008;**40**:238–45.

Foley 2012 *{published data only}*

Baggs J, Silverstein B, Foley M. Workplace health and safety regulations: impact of enforcement and consultation on workers' compensation claims rates in Washington State. *American Journal of Industrial Medicine* 2003;**43**(5):483–94.

Foley M, Fan ZJ, Rauser E, Silverstein B. The impact of regulatory enforcement and consultation visits on workers' compensation claim incidence rates and costs, 1999-2008. *American Journal of Industrial Medicine* 2012;**55**(11): 976–90.

Gray 2005a *{published data only}*

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries. NBER Working Paper No. 9119 August 2002:1–31.

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries, 1979-1998. Industrial and Labor Relations Review 2005; Vol. 58, issue 4:571–87.

Scholz JT, Gray WB. Can government facilitate cooperation? An informational model of OSHA enforcement. American Journal of Political Science 1997; Vol. 41, issue 3:693–717.

Scholz JT, Gray WB. OSHA enforcement and workplace injuries: a behavioral approach to risk assessment. *Journal of Risk and Uncertainty* 1990;**3**(3):283–305.

Gray 2005b *{published data only}*

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries. NBER Working Paper No. 9119 August 2002:1–31.

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries, 1979-1998. Industrial and Labor Relations Review 2005; Vol. 58, issue 4:571–87.

Gray 2005c *{published data only}*

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries. NBER Working Paper No. 9119 August 2002:1–31.

Gray WB, Mendeloff JM. The declining effects of OSHA inspections on manufacturing injuries, 1979-1998. Industrial and Labor Relations Review 2005; Vol. 58, issue 4:571–87.

Mendeloff J, Gray WB. Inside the black box: how do OSHA inspections lead to reductions in workplace injuries? . *Law and Policy* 2005;**27**(2):219–37.

Haviland 2012 *{published data only}*

Haviland A, Burns R, Gray W, Ruder T, Mendeloff J. What kinds of injuries do OSHA inspections prevent?. *Journal of Safety Research* 2010;**41**(4):339–45.

Haviland AM, Burns RM, Gray WB, Ruder T, Mendeloff J. A new estimate of the impact of OSHA inspections on manufacturing injury rates, 1998-2005. American Journal of Industrial Medicine 2012 May 7 [Epub ahead of print]: 1–12.

Haviland AM, Burns RM, Gray WB, Ruder T, Mendeloff J. The Impact of OSHA Inspections on Lost Time Injuries in Manufacturing: Pennsylvania Manufacturing, 1998-2005. Santa Monica, CA: RAND Corporation, 2008. http://www.rand.org/pubs/working˙papers/WR592.

Hogg-Johnson 2011 *{published data only}*

Hogg-Johnson S, Robson L, Cole D, Subrata P, Amick B, Tompa E, et al.A randomised controlled study to evaluate the effectiveness of targeted occupational health and safety consultation or inspection in Ontario manufacturing workplaces.. *Occupational and Environmental Medicine* 2012;**69**(12):1–11.

Kemmlert 1994 *{published data only}*

Kemmlert K. Prevention of occupational musculoskeletal injuries. Labour Inspectorate investigation. *Scandinavian Journal of Rehabilitation Medicine. Supplement* 1996;**35**: 1–34.

Kemmlert K. Preventive effects of work place investigations in connection with musculoskeletal occupational injuries. *Scandinavian Journal of Rehabilitation Medicine* 1994;**26**: 21–6.

Kniesner 2004 *{published data only}*

Kniesner TJ, Leeth JD. Data mining mining data: MSHA enforcement efforts, underground coal mine safety, and new health policy implications. *Journal of Risk and Uncertainty* 2004;**29**(2):83–111.

Levine 2012 *{published data only}*

Levine DI, Toffel MW, Johnson MS. Randomized government safety inspections reduce worker injuries with no detectable job loss. *Science* 2012;**336**(6083):907–11.

Nelson 1997 *{published data only}*

Nelson NA, Kaufman J, Kalat J, Silverstein B. Falls in construction: injury rates for OSHA-inspected employers before and after citation for violating the Washington State Fall Protection Standard. *American Journal of Industrial Medicine* 1997;**31**(3):296–302.

Robertson 1983 *{published data only}*

Robertson LS, Keeve JP. Worker injuries: the effects of Workers' Compensation and OSHA inspections. *Journal of Health Politics, Policy and Law* 1983;**8**(3):581–97.

Smith 1979a *{published data only}*

Smith RS. The impact of OSHA inspections on manufacturing injury rates. *Journal of Human Resources* 1979;**14**(2):145–70.

Smith 1979b *{published data only}*
Smith RS. The impact of OSHA inspections on manufacturing injury rates. *Journal of Human Resources* 1979;**14**(2):145–70.

Weil 1996 *{published data only}*
Weil D. If OSHA is so bad, why is compliance so good?. *RAND Journal of Economics* 1996;**27**(3):618–40.

Weil 2001 *{published data only}*
Morantz AD. Has devolution injured American workers? State and federal enforcement of construction safety. *Journal of Law Economics & Organization* 2009;**25**(1):183–210.
Weil D. Assessing OSHA performance: new evidence from the construction industry. *Journal of Policy Analysis and Management* 2001;**20**(4):651–74.

References to studies excluded from this review

Adams 2007 *{published data only}*
Adams Jones M, Adams Jones D, Powell C. Towards reality and away from blind acceptance: changes in the attitude of public house bar staff to their exposure to second hand tobacco smoke in the work-place since the announcement of legislative measures. *Health Education Journal* 2007;**66**(4):323–34.

Attfield 1992 *{published data only}*
Attfield MD, Althouse RB. Surveillance data on US coal miners' pneumoconiosis, 1970 to 1986. *American Journal of Public Health* 1992;**82**(7):971–7.

Auld 2001 *{published data only}*
Auld MC, Emery JCH, Gordon DV, McClintock D. The efficacy of construction site safety inspections. *Journal of Labor Economics* 2001;**19**(4):900–21.

Baron-Epel 2012 *{published data only}*
Baron-Epel O, Satran C, Cohen V, Drach-Zehavi A, Hovell M F. Challenges for the smoking ban in Israeli pubs and bars: analysis guided by the behavioral ecological model. *Israel Journal of Health Policy Research* 2012;**1**(1):28.

Boden 1985 *{published data only}*
Boden LI. Government regulation of occupational safety: underground coal mine accidents 1973-75. *American Journal of Public Health* 1985;**75**(5):497–501.

Brown 2003 *{published data only}*
Brown GD. China's factory floors: an industrial hygienist's view. *International Journal of Occupational and Environmental Health* 2003;**9**(4):326–39.

Joy 2007 *{published data only}*
Joy JJ, Middendorf PJ. Noise exposure and hearing conservation in U.S. coal mines-a surveillance report. *Journal of Occupational and Environmental Hygiene* 2007;**4**(1):26–35.

Ko 2010 *{published data only}*
Ko K, Mendeloff J, Gray W. The role of inspection sequence in compliance with the US Occupational Safety and Health Administration's (OSHA) standards: Interpretations and implications. *Regulation and Governance* 2010;**4**(1):48–70.

Lissner 2011 *{published data only}*
Lissner L, Zayzon R. Is the European directive 98/24/EC on chemical agents effective?: Evaluation of its practical implementation at workplaces. *Gefahrstoffe Reinhaltung der Luft* 2011;**71**(6):247–54.

Mancini 2005 *{published data only}*
Mancini G, Baldasseroni A, Laffi G, Curti S, Mattioli S, Violante FS. Prevention of work related eye injuries: long term assessment of the effectiveness of a multicomponent intervention among metal workers. *Occupational and Environmental Medicine* 2005;**62**(12):830–5.

Morantz 2009 *{published data only}*
Morantz AD. Has devolution injured American workers? State and federal enforcement of construction safety. *Journal of Law Economics & Organization* 2009;**25**(1):183–210.

Niskanen 2013 *{published data only}*
Niskanen T. The effects of the enforcement legislation in the Finnish occupational safety and health inspectorate. *Safety Science* 2013;**55**:133–48.

Raymond 2003 *{published data only}*
Raymond MJ, Halcón LL, Pirie PL. Regulation of tattooing in Minneapolis and St. Paul, Minnesota: tattooists' attitudes and relationship between regulation and practice. *Public Health Reports* 2003;**118**(2):154–61.

Smitha 2001 *{published data only}*
Smitha MW, Kirk KA, Oestenstad KR, Brown KC, Lee SD. Effect of state workplace safety laws on occupational injury rates. *Journal of Occupational and Environmental Medicine* 2001;**43**(12):1001–10.

Viscusi 1979 *{published data only}*
Viscusi WK. The impact of occupational safety and health regulation. *The Bell Journal of Economics* 1979;**10**(1):117–40.

References to studies awaiting assessment

Bordas 2001 *{published data only}*
Bordas RM, Davis GA, Hopkins BL, Thomas RE, Rummer RB. Documentation of hazards and safety perceptions for mechanized logging operations in East Central Alabama. *Journal of Agricultural Safety and Health* 2001;**7**(2):113–23.

Geminiani 2008 *{published data only}*
Geminiani F L, Smallwood J J, Van Wyk J J. The effectiveness of the Department of Labour (DoL) occupational health and safety (OH&S) Inspectorate relative to the construction industry in South Africa. COBRA 2008 - Construction and Building Research Conference of the Royal Institution of Chartered Surveyors. 2008.
Geminiani F L, Smallwood J J, Van Wyk J J. The effectiveness of the occupational health and safety (OH&S) inspectorate in South African construction. Association of Researchers in Construction Management, ARCOM 2008 - Proceedings of the 24th Annual Conference2. 2008; Vol. 2:1113–21.
Geminiani FL, Smallwood JJ, Fee S. A comparative analysis between contractors' and inspectors' perceptions of the

department of labour occupational health and safety inspectorate relative to South African construction. *Safety Science* 2013;**53**:186–92.

Gillen 2004 ***{published data only}***
Gillen M, Kools S, McCall C, Sum J, Moulden K. Construction managers' perceptions of construction safety practices in small and large firms: a qualitative investigation. *Work* 2004;**23**(3):233–43.

Gray 2006 ***{published data only}***
Gray Garry C. The regulation of corporate violations: punishment, compliance, and the blurring of responsibility. *British Journal of Criminology* 2006;**46**(5):875–92.

Guidotti 1996 ***{published data only}***
Guidotti TL, Watson L, Wheeler M, Jhangri GS. The Fort McMurray Demonstration Project in social marketing: health- and safety-related behaviour among oil sands workers. *Occupational Medicine* 1996;**46**(4):265–74.

Mayhew 1999 ***{published data only}***
Mayhew C. The impact of auditing on the OHS performance of demolishers. *Journal of Occupational Health and Safety - Australia and New Zealand* 1999;**15**(5):441–7.

Additional references

Alper 2009
Alper SJ, Karsh BT. A systematic review of safety violations in industry. *Accident Analysis and Prevention* 2009;**41**: 739–54.

Austin 2011
Austin PC, Laupacis A. A tutorial on methods to estimating clinically and policy-meaningful measures of treatment effects in prospective observational studies: a review. *International Journal of Biostatistics* 2011;**7**(1):i–32.

Azaroff 2002
Azaroff LS, Levenstein C, Wegman DH. Occupational injury and illness surveillance: conceptual filters explain underreporting. *American Journal of Public Health* 2002;**92** (9):1421–9.

Braithwaite 2006
Braithwaite V. Ten things you always wanted to know about regulation and never wanted to ask. *Australian Law Librarian* 2006;**14**(3):19–28.

Braithwaite 2007a
Braithwaite V. Responsive regulation and taxation: introduction. *Law & Policy* 2007;**29**(1):3–10.

Braithwaite 2007b
Braithwaite V, Murphy K, Reinhart M. Taxation threat, motivational postures, and responsive regulation. *Law & Policy* 2007;**29**(1):137–58.

Braithwaite 2011
Braithwaite V. Motivations, attitudes, perceptions, and skills: pathways to safe work. SafeWork Australia, Canberra, Australia. SafeWork Australia, 2011.

Concha-Barrientos 2005
Concha-Barrientos M, Nelson DI, Fingerhut M, Driscoll T, Leigh J. The global burden due to occupational injury. *American Journal of Industrial Medicine* 2005;**48**(6):470–81.

Creely 2007
Creely KS, Cowie H, Van TM, Kromhout H, Tickner J, Cherrie JW. Trends in inhalation exposure-a review of the data in the published scientific literature. *Annals of Occupational Hygiene* 2007;**51**:665–78.

Downs 1998
Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health* 1998;**52**(6):377–84.

GRADEpro 2008
Brozek J, Oxman A, Schünemann H. GRADEpro. 3.2 for Windows. GRADE working group, 2008.

Gray 1993
Gray WB, Scholz JT. Does regulatory enforcement work? A panel analysis of OSHA enforcement. *Law & Society Review* 1993;**27**(1):177–213.

Gunningham 2007
Gunningham N. Prosecution for OHS offences: deterrent or disincentive?. *Sydney Law Review* 2007;**29**:359–90.

Hannes 2011
Hannes K. Chapter 4: Critical appraisal of qualitative research. In: Noyes J, Booth A, Hannes K, Harden A, Harris J, Lewin S, Lockwood C (editors). Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions. Version 1 (updated August 2011). Cochrane Collaboration Qualitative Methods Group, 2011. Available from URL http://cqrmg.cochrane.org/supplemental–handbook–guidance.

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 www.cochrane-handbook.org.

Hämäläinen 2009
Hämäläinen P, Saarela KL, Takala J. Global trend according to estimated number of occupational accidents and fatal work-related diseases at region and country level. *Journal of Safety Research* 2009;**40**(2):125–39.

ILO 2004
International Labour Organization. Conclusions adopted by the International Labour Conference at its 91st Session, 2003. Global Strategy on Occupational Safety and Health 2004:1–20.

JBI 2011
The Joanna Briggs Institute. Reviewers' Manual. Available from http://www.joannabriggs.edu.au/Documents/sumari/Reviewers%20Manual-2011.pdf 2011.

Kromhout 2000
Kromhout H, Vermeulen R. Long-term trends in occupational exposure: Are they real? What causes them? What shall we do with them?. *Annals of Occupational Hygiene* 2000;**44**(5):325–7.

Lierman 2010
Lierman BE. To assure safe and healthful working conditions: taking lessons from labor unions to fulfil OSHA's promises. *Loyola Journal of Public Interest Law* 2010;**12**(2010):1–38.

McQuiston 1998
McQuiston TH, Zakocs RC, Loomis D. The case for stronger OSHA enforcement - evidence from evaluation research. *American Journal of Public Health* 1998;**88**(7): 1022–4.

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 2011
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

Shapiro 1997
Shapiro SA, Rabinowitz RS. Punishment versus cooperation in regulatory enforcement: a case study of OSHA. *Administrative Law Review* 1997;**49**:713–62.

Stock 2007
Stock JH, Watson MW. *Introduction to Econometrics*. 2nd Edition. Boston, MA: Pearson/Addison Wesley, 2007.

Tompa 2007
Tompa E, Trevithick S, McLeod C. Systematic review of the prevention incentives of insurance and regulatory mechanisms for occupational health and safety. *Scandinavian Journal of Work, Environment and Health* 2007;**33**(2):85–95.

US GAO 2013
US Government Accountability Office. Workplace safety and health; further steps by OSHA would enhance monitoring of enforcement and effectiveness [Report GOA-13-61]. US Government Accountability Office, Washington, DC, USA 2013:1–34.

van der Molen 2007
van der Molen HF, Lehtola MM, Lappalainen J, Hoonakker PL, Hsiao H, Haslam R, et al.Interventions for preventing injuries in the construction industry. *Cochrane Database of Systematic Reviews* 2007, Issue 4. [DOI: 10.1002/14651858.CD006251.pub2]

Verbeek 2004
Verbeek J, Sengers MJ, Riemens L, Haafkens J. Patient expectations of treatment for back pain: a systematic review of qualitative and quantitative studies. *Spine* 2004;**29**(20): 2309–18. [PUBMED: 15480147]

Viscusi 2005
Viscusi WK, Harrington JE, Vernon JM. Regulation of workplace health and safety. In: Viscusi WK, Harrington JE, Vernon JM editor(s). *Economics of Regulation and Anti-trust*. 4th Edition. Cambridge (MA): Massachusetts Institute of Technology, 2005:827–63.

Wagner 2002
Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. *Journal of Clinical Pharmacy and Therapeutics* 2002;**27**(4):299–309.

WHO 2006
WHO Collaborating Centres in Occupational Health. Declaration on Workers Health. Approved at the Seventh Meeting of the WHO Collaborating Centres for Occupational Health. Stresa, Italy: WHO, 8–9 June 2006: 1–4.

WHO 2007
World Health Assembly (WHA). Workers' health: global plan of action. WHA 60.26. World Health Organization, 23 May 2007:1–12.

WHO 2008
World Health Organization. *WHO Global Strategy on Diet, Physical Activity and Health: A Framework to Monitor and Evaluate Implementation*. Geneva, Switzerland: World Health Organization, 2008.

WHO 2009
World Health Organization. *Global Health Risk: Mortality and Burden of Disease Attributable to Selected Major Risks*. Geneva, Switzerland: World Health Organization, 2009.

World Bank 2012
World Bank. World data bank. World Development Indicators.. http://data.worldbank.org/data-catalog/world-development-indicators 21 December 2012.

Zhang 1998
Zhang J, Yu KF. What is the relative risk? A method for correcting the odds ratio in cohort studies of common outcomes. *JAMA* 1998;**280**(19):1690–1.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

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

Burstyn 2010

Methods	Panel study
Participants	Canada, Alberta Firms, manufacturing industry N = ??
Interventions	Intervention group: inspection with or without penalty (proactive inspector) • Inspection: OSHA • Proactive inspector Control group: inspection with or without penalty (reactive inspector) • Inspection: OSHA • Reactive inspector
Outcomes	Primary outcome, exposure: As # of compliance orders needed to resolve non-compliance
Notes	Time of the intervention: 2002-2006 Firm size: not reported Type of work: not reported Previous inspections: not reported Baseline injury rates: not reported Funding: not reported Conflict of interest: no

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (subjects)	High risk	Outcome compliance, participants not blinded
Blinding (outcome assessors)	High risk	Outcome assessors not blinded
Retrospective unplanned subgroup analyses	High risk	No prespecified model
Follow-up	High risk	No adjustment
Statistical tests	High risk	Poisson regression model, survival analysis missing
Compliance	Unclear risk	Not reported
Outcome measures	Unclear risk	Outcome was time to compliance, unclear

Burstyn 2010 (*Continued*)

Selection bias (population)	Unclear risk	Type of industry not reported
Selection bias (time)	Low risk	Recruited over the same time period
Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Unclear risk	Type of work not reported, pre-intervention inspection not assessed, unclear adjustment for baseline injury rates
Incomplete outcome data	Unclear risk	Not reported

Chen 2008

Methods	CBA; we used presented data to perform an ITS analysis
Participants	USA Firms, trucking industry N = 113,441 between 1999-2001 Firm size: number of employees Type of work: motor carrier driver Previous inspections: unknown Baseline injury rates: median over years 2.99 crashes per 100 trucks
Interventions	Intervention group: inspection, warnings and orders (N = 3705) • Inspection (compliance review (CR)): Safety ratings to determine whether a motor carrier meets the Section 385.5 Safety Fitness standards (FMCSA 2006), (1) satisfactory, (2) conditional satisfactory or (3) unsatisfactory, a follow-up review may be conducted to ensure that all necessary corrective actions have been taken • Warnings or orders: carriers receiving a conditional satisfactory or unsatisfactory rating are required to undertake corrective actions within 30 days or the carrier's operating authority will be revoked and an operating out of service will be imposed prohibiting the carrier from operating any motor vehicle in the USA Control group: no compliance review (N = 109,736) (Co-interventions not reported)
Outcomes	Primary outcome, exposure: Number of crashes, involving a truck or a bus of motor carriers operating in the United States with at least 1 fatality, injury or vehicle towed away from the scene as a result of disabling crash damage
Notes	Time of the intervention: 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003; 1999-2001 used as years with interruption for ITS analysis Funding: not reported Conflict of interest: no

Chen 2008 (*Continued*)

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	Low risk	No blinding of participants, but knowing about receiving the inspection does not change the number of reported accidents
Blinding (outcome assessors)	Low risk	Measurements used from crash file "contains data from State Police crash reports"
Retrospective unplanned subgroup analyses	Low risk	No data dredging, objectives of this study: "whether the reduction occurred in every sub-group of reviewed trucking companies"
Follow-up	Low risk	Similar follow-up for cases and controls
Statistical tests	Low risk	Appropriate
Compliance	Unclear risk	Not reported
Outcome measures	Unclear risk	"some states did not report all eligible crashes"
Selection bias (population)	Unclear risk	Outcome on company level, but one carrier could be employed in more than one company
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	No randomisation
Allocation concealment	High risk	Not randomised
Adjustment for confounding	Unclear risk	Firm size, previous inspection for control and intervention group unknown
Incomplete outcome data	Low risk	No loss

Foley 2012

Methods	Panel study
Participants	USA, Washington Firms, mixed types of industry N = 8752
Interventions	Intervention group: inspections: (N = 440) • Programmed or complaint inspection Control group: no intervention (N = 8312)
Outcomes	Primary outcome, injuries: Number of registered lost-workday claims per 100 FTEs, the change in the claim rate 1 year after intervention
Notes	Time of the intervention: 1999 Firm size: number of employees not reported Type of work: mixed Previous inspections: not in 1 year prior intervention Baseline injury rates: only the account's SFY 1997 compensable claims rate were controlled for in the analysis (intervention and control group 2 and 3 had consistently higher average SFY 1997 claims rates than those accounts with no activity) Funding: Washington State Department of Labor and Industries Conflict of interest: authors are employed by the Funder (Washington State Department of Labor and Industries)

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (subjects)	High risk	Based on self reported outcome
Blinding (outcome assessors)	Low risk	Registered compensable cases, based on self reporting
Retrospective unplanned subgroup analyses	Unclear risk	Predefined model not presented
Follow-up	Low risk	Same time period
Statistical tests	Low risk	Univariate and multivariate analyses were undertaken
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Based on self reporting
Selection bias (population)	Unclear risk	Not reported
Selection bias (time)	Low risk	Same time period

Foley 2012 (*Continued*)

Randomisation	High risk	No randomisation
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Low risk	3 out of 4, adjusted for average size and baseline claim-injury rates, no intervention 2 years prior evaluation, type of work unclear
Incomplete outcome data	Unclear risk	Not reported

Gray 2005a

Methods	Panel study
Participants	USA Firms, manufacturing industry N = not reported
Interventions	Intervention group: inspection with or without penalty • Inspection: OSHA (programmed inspections and inspections after complaint) Control group: no inspection
Outcomes	Primary outcome, injuries: Total number of lost work days during the year/per firm
Notes	Time of the intervention: 1979-1985 (7 years) Firm size: not reported Type of work: mostly physical Previous inspections: not reported Baseline injury rates: not reported Funding: not reported Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, lost time injuries self reported
Blinding (outcome assessors)	High risk	Data from data base but relying on self reports
Retrospective unplanned subgroup analyses	Low risk	Prespecified model
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	t-test and maximum likelihood estimates

Gray 2005a (*Continued*)

Compliance	Unclear risk	Not reported
Outcome measures	High risk	Relying on self reports
Selection bias (population)	Unclear risk	Not reported if firms from different states (USA)
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	High risk	Not adjusted for different size, injury rates, previous interventions, type of work similar
Incomplete outcome data	Unclear risk	Not reported

Gray 2005b

Methods	Panel study
Participants	USA Firms, manufacturing industry N = not reported
Interventions	Intervention group: inspection with or without penalty • Inspection: OSHA (programmed inspections and inspections after complaint) Control group: no inspection
Outcomes	Primary outcome, injuries: Total number of lost work days during the year/per firm
Notes	Time of the intervention: 1987-1991 (4 years) Firm size: not reported Type of work: mostly physical Previous inspections: not reported Baseline injury rates: not reported Funding: not reported Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, lost time injuries self reported

Gray 2005b (*Continued*)

Blinding (outcome assessors)	High risk	Data from database but relying on self reports
Retrospective unplanned subgroup analyses	Low risk	Prespecified model
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	t-test and maximum likelihood estimates
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Relying on self reports
Selection bias (population)	Unclear risk	Not reported if firms from different states (USA)
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	High risk	Not adjusted for different size, injury rates, previous interventions, type of work similar
Incomplete outcome data	Unclear risk	Not reported

Gray 2005c

Methods	Panel study
Participants	USA Firms, manufacturing industry N = not reported
Interventions	Intervention group: inspection with or without penalty • Inspection: OSHA (programmed inspections and inspections after complaint) Control group: no inspection
Outcomes	Primary outcome, injuries: Total number of lost work days during the year/per firm
Notes	Time of the intervention: 1992-1998 (7 years) Firm size: not reported Type of work: mostly physical Previous inspections: not reported Baseline injury rates: not reported Funding: not reported

Gray 2005c (*Continued*)

	Conflict of interest: no	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, lost time injuries self reported
Blinding (outcome assessors)	High risk	Data from database but relying on self reports
Retrospective unplanned subgroup analyses	Low risk	Prespecified model
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	t-test and maximum likelihood estimates
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Relying on self reports
Selection bias (population)	Unclear risk	Not reported if firms from different states (USA)
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	High risk	Not adjusted for different size, injury rates, previous interventions, type of work similar
Incomplete outcome data	Unclear risk	Not reported

Haviland 2012

Methods	Panel study, regression analysis
Participants	USA, Pennsylvania Firms, manufacturing industry N = 8645
Interventions	Intervention group: programmed inspection with penalty (no detail) Control group: no intervention or inspection without penalty
Outcomes	Injuries: days away from work (DAW), including disease and injury, DAW per 100 person/year

Haviland 2012 (*Continued*)

Notes	Time of the intervention: 1998-2005 (8 years) Firm size: only small firms (20 to 250 employees) Type of work: mostly physical, manufacturing industry Previous inspections: not reported Baseline injury rates: not reported Funding: Commonwealth Pennsylvania Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, outcome self reported (lost time injury data)
Blinding (outcome assessors)	High risk	Workers are assessors, data from registry (Worker Compensation data) but based on self report
Retrospective unplanned subgroup analyses	High risk	No predefined model, firms with 10 to 20 employees excluded
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Claims of lost time injuries, self reported
Selection bias (population)	Low risk	Same population (manufacturing industry and adjustment for SIC, all in Pennsylvania)
Selection bias (time)	Low risk	Same time
Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Unclear risk	Same type of work and firm size, pre-inspections and baseline injury rate difference not reported/adjusted
Incomplete outcome data	Unclear risk	Not reported

Hogg-Johnson 2011

Methods	RCT
Participants	Canada Firms, manufacturing industry N = 1219 Firm size: not reported, described as similar in intervention and control groups Type of work: manufacturing industry, mostly physical Previous inspections: not in 2 years prior study intervention
Interventions	Intervention group: inspection and orders (N = 619) • Inspection: once or twice during intervention year, without prior warning • Orders: upon inspection, inspectors wrote orders based on non-compliance with legislative and regulatory OSH requirements comprehensive or focused on particular hazard Control group: no intervention (N = 600)
Outcomes	Primary outcome, injuries: Annual mean and median claim rates per year as: 1) overall injury claim rate (yearly rates of claims registered per 100 FTE), 2) lost time injury claim rate (yearly rates of claims registered per 100 FTE) and 3) disability day rate (measured as number of calendar days of 100% wage replacement within 2 years of date of accident for all claims filed within the year per 100 FTE) Adverse outcome, firm closure: mean and median firm closure rates (whether a firm went out of business in a given year)
Notes	Time of the intervention: 1 April 2006 to 31 March 2007 Baseline injury rates: exact data not presented, significantly different with higher rates in the intervention than control group 2 Funding: Ontario Workplace Safety & Insurance Board Research Advisory Council, Canadian Institute of Health Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, plausible to bias claim rates (self reported)
Blinding (outcome assessors)	Low risk	Extracted from official administrative records (Ontario Workplace Safety and Insurance Board)
Retrospective unplanned subgroup analyses	Low risk	No data dredging
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	Generalised estimating equation models and Pearson square test, Wilcoxon rank sum test

Hogg-Johnson 2011 (*Continued*)

Compliance	Unclear risk	Not reported
Outcome measures	High risk	Self reported (claims)
Selection bias (population)	Low risk	Different intervention groups
Selection bias (time)	Low risk	Different intervention groups
Randomisation	Low risk	Conducted using SAS 9.1, afterwards further exclusion but same criteria and similar percentage excluded across study groups
Allocation concealment	High risk	Not concealed, participants received in addition non-assigned interventions during intervention period (enforcing activities ca. 21%, consulting activities ca. 16%)
Adjustment for confounding	Low risk	Firm size and previous inspections comparable; all firms in manufacturing industry
Incomplete outcome data	Low risk	Less than 20%

Kemmlert 1994

Methods	RCT
Participants	Sweden Individual workplaces, type of industry not reported N = 195
Interventions	Intervention group: inspection with or without penalty (N = 98) • Inspection: by labour inspectorate, announced inspection, assessment following ergonomic workplace checklist to identify musculoskeletal stress factors, inspector received special training • Warning or orders: inspector notices in case of insufficiencies and to express demands Control group: no intervention (N = 97)
Outcomes	Primary outcome, exposure: Reduced workload, achieved if the harmful situation reported in the injury report on musculoskeletal injuries did not exist anymore Adverse outcome, active employment: Employment status (whether a worker went out of employment in a given workplace after 3 years)
Notes	Time of the intervention: 1985 Firm size: not reported Type of work: not reported Previous inspections: not reported

Kemmlert 1994 (*Continued*)

	Baseline injury rates: occupational musculoskeletal injury report was inclusion criteria for control and intervention group, outcome workplace specific Funding: no funding Conflict of interest: no

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (subjects)	High risk	Subjects not blinded
Blinding (outcome assessors)	Low risk	Assessor blinded
Retrospective unplanned subgroup analyses	Unclear risk	No protocol available
Follow-up	Low risk	Similar follow-up
Statistical tests	Unclear risk	Not reported
Compliance	Unclear risk	Not assessed
Outcome measures	High risk	Outcome improvement relies on self reported data from baseline
Selection bias (population)	Unclear risk	Not reported, outcome for individual workplaces
Selection bias (time)	Low risk	Same time period
Randomisation	Unclear risk	Method not described
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Unclear risk	Firm size and previous inspections unknown
Incomplete outcome data	High risk	> 50% loss of follow-up

Kniesner 2004

Methods	Panel study
Participants	USA Firms, mining industry N = 292
Interventions	Intervention: inspection with penalty or inspection with closure order • Inspection: MSHA • Penalty: fine • Closure order: mine closure

Kniesner 2004 (*Continued*)

	Control group: fewer inspections with penalty or fewer inspections with closure order • Inspection: MSHA • Penalty: fine • Closure order: mine closure
Outcomes	Primary outcome, injuries: Number of lost workday injuries including fatalities per firm per quarter
Notes	Time of the intervention: 1983-1997 (15 years) Firm size: not reported Type of work: mostly physical Previous inspections: not reported Baseline injury rates: not reported Funding: not reported Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Self report
Blinding (outcome assessors)	High risk	Worker assessed, self reported
Retrospective unplanned subgroup analyses	High risk	+200 models
Follow-up	Low risk	Same time
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Self reported
Selection bias (population)	Unclear risk	Adjustment for different location unclear
Selection bias (time)	Low risk	Over same time
Randomisation	High risk	No randomisation
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Low risk	Adjustments for firm size and injury rates, same type of work, previous inspections unknown
Incomplete outcome data	Unclear risk	Not reported

Levine 2012

Methods	CBA Comment: intervention subjects randomly chosen, controls matched (when fulfilling inclusion criteria matched according to same industry, same region; the firm with the most similar numbers of employees got chosen)
Participants	USA, California Firms, mixed industries N = 818
Interventions	Intervention group: inspection and if indicated further penalties • Random inspection by Cal/OSHA inspectors for industries with high injury rates • No details about further penalties Control group: no random inspection
Outcomes	Primary outcome, injuries: Injury rates Adverse outcome: Firm closure (survival), sales, employment, payroll
Notes	Time of the intervention:1996-2006 Firm size: mostly small firms, mean 34.28 (36.3) (# 0-570) Type of work: mostly physical work Previous inspections: no OSHA inspection in 2 years prior intervention Baseline injury rates: exact data not presented, pre-trends showed no statistical difference (14% decline in intervention and 12% decline in control group) Funding: Commission on Health and Safety and Workers' Compensation; Harvard Business School's Division of Research and Faculty Development; Kauffman Foundation; University of California at Berkeley's Institute for Research on Labor and Employment; University of California's Labor and Employment Relations Fund Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Outcome self reported to the workers' compensation system
Blinding (outcome assessors)	High risk	Injury data from the workers' compensation system but self reported
Retrospective unplanned subgroup analyses	Unclear risk	No protocol available
Follow-up	Low risk	Similar
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported

Levine 2012 (*Continued*)

Outcome measures	High risk	Based compensation claims
Selection bias (population)	Low risk	Same population
Selection bias (time)	Low risk	Same time
Randomisation	Unclear risk	Randomisation procedure not reported
Allocation concealment	High risk	Inspectors could make a choice
Adjustment for confounding	Low risk	Control and intervention group similar at baseline in 3 of 4 confounders (included mostly small firms, mostly physical work, included only if no inspections 2 years prior intervention)
Incomplete outcome data	Low risk	6% loss to follow-up

Nelson 1997

Methods	CBA
Participants	USA, Washington Firms, construction industry N = 9085 (I = 784, C = 8,301) Firm size (number of employees): any size, I mean = 22 (1 to 542), C mean = 7 (1 to 404) (number of employees were identified as number of hours worked by employees assuming that each full-time employee works 2000 hours per year (40 hours per week for 50 weeks per year)) Type of work: mostly physical, same type of industry Previous inspections: not reported Baseline injury rates: as number of fall claims I = 0.399, C = 0.071; fall injury rate per 200,000 hrs worked I = 1.78 C = 1.04
Interventions	Intervention group: inspection, citations and monetary penalties (N = 784): • Visited by state plan safety inspectors, reasons for the inspections included programmed (scheduled) evaluations (83.2%), employee complaints (8.2%), referral (1.8%), inspection of a fatality or accident (0.8%), and other unscheduled evaluations (5.9%) • Cited for violating the falls in construction standard • Monetary penalties no details Control group: no intervention (N = 8301)
Outcomes	Primary outcome, injuries: Fall injury claims with min. 4 days of lost work time because of fall (coded as injury event or exposure codes including fall from elevation, platform or ladder; fall from piled matter; fall on stairs; fall into openings; fall from roof and fall to lower level)

Nelson 1997 (*Continued*)

Notes	Time of the intervention: 1991-1992, median date of inspection was October 1991 Funding: not reported Conflict of interest: unknown

Risk of bias

Bias	Authors' judgement	Support for judgement
Blinding (subjects)	High risk	Not blinded, outcome (compensable fall claims with work time loss) likely to be influenced
Blinding (outcome assessors)	Low risk	Claims and employment data were obtained from the Washington state department of labour and industries files
Retrospective unplanned subgroup analyses	Low risk	No data dredging
Follow-up	Low risk	Similar follow-up
Statistical tests	Low risk	Appropriate
Compliance	Unclear risk	No evidence is reported on whether follow-up inspections were done
Outcome measures	Low risk	Self reported claims from register but unlikely biased (compensable fall claims with work time loss)
Selection bias (population)	Low risk	Same population
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	No randomisation
Allocation concealment	High risk	Not concealed, allocation to intervention by inspector
Adjustment for confounding	High risk	Inspections prior intervention not assessed, fall injury rate per 200,000 hours worked higher in inspection group than in control group (I = 1.78; C = 1.04), not adjusted for different firm size
Incomplete outcome data	Low risk	No loss

Robertson 1983

Methods	Panel study
Participants	USA Firms, manufacturing industry N = 3 plants, total 2700 workers
Interventions	Intervention: inspection with penalty by OSHA • Any type of inspection • Penalties include citation Control group: • Inspection without citation
Outcomes	Primary outcome, injuries: • Lost time injuries per firm per year
Notes	Time of the intervention: 1973-1980 Firm size: big firms Type of work: mostly physical work Previous inspections: not reported Baseline injury rates: included in analyses as $^{observed}/_{expected}$ injuries, rates not reported Funding: Yale University by Atlantic Richfield Corporation Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Self reported injuries
Blinding (outcome assessors)	High risk	Not blinded
Retrospective unplanned subgroup analyses	Low risk	One analysis
Follow-up	Low risk	Same follow-up time
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Self reported
Selection bias (population)	Low risk	Firms from same industry
Selection bias (time)	Low risk	Same time
Randomisation	High risk	No randomisation

Robertson 1983 (*Continued*)

Allocation concealment	High risk	No allocation concealment
Adjustment for confounding	Low risk	Adjusted except for pre-intervention inspections
Incomplete outcome data	Unclear risk	Not reported

Smith 1979a

Methods	Panel study
Participants	USA Firms, manufacturing industry N = unknown
Interventions	Intervention: inspection with or without penalty Control group: inspection later in that year
Outcomes	Primary outcome, injuries: • Lost workday injuries rate (# injuries per 100 full-time workers)
Notes	Time of the intervention:between 1972-1973 Firm size: any size, average not reported, separate analysis for firms < 100, 100 to 249 and > 250 workers Type of work: mostly physical work Previous inspections: not reported Baseline injury rates: not reported, included in regression model as injury rate in year of inspection Funding: Department of Labor Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, outcome self reported
Blinding (outcome assessors)	High risk	Not blinded
Retrospective unplanned subgroup analyses	Low risk	One model
Follow-up	Low risk	Same time of follow-up
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported

Smith 1979a (*Continued*)

Outcome measures	High risk	Self reported
Selection bias (population)	Unclear risk	All from manufacturing industry, geographical region unclear
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	No randomisation
Allocation concealment	High risk	No allocation concealment
Adjustment for confounding	Low risk	3 out of 4 (same type of industry, adjusted for injury rates, separate analysis for different firm size, previous inspection unknown)
Incomplete outcome data	Unclear risk	Not reported

Smith 1979b

Methods	Panel study
Participants	USA Firms, manufacturing industry N = unknown
Interventions	Intervention: inspection with or without penalty Control group: inspection later in that year
Outcomes	Primary outcome, injuries: • Lost workday injuries rate (# injuries per 100 full-time workers)
Notes	Time of the intervention:between 1973-1974 Firm size: any size, average not reported, separate analysis for firms < 100, 100 to 249 and > 250 workers Type of work: mostly physical work Previous inspections: not reported Baseline injury rates: not reported, included in regression model as injury rate in year of inspection Funding: Department of Labor Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded, outcome self reported
Blinding (outcome assessors)	High risk	Not blinded

Smith 1979b (*Continued*)

Retrospective unplanned subgroup analyses	Low risk	One model
Follow-up	Low risk	Same time of follow-up
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Self reported
Selection bias (population)	Unclear risk	All from manufacturing industry, geographical region unclear
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	No randomisation
Allocation concealment	High risk	No allocation concealment
Adjustment for confounding	Low risk	3 out of 4 (same type of industry, adjusted for injury rates, separate analysis for different firm size, previous inspection unknown)
Incomplete outcome data	Unclear risk	Not reported

Weil 1996

Methods	Panel study
Participants	USA Firms, custom woodworking industry N = 250
Interventions	Inspection with and without penalties by OSHA • Any type of inspection including complaint and follow-up inspections • Penalties include fines and citation Comparisons: • More inspections versus fewer inspections • Complaint versus any inspection • Follow-up versus any inspection • Higher inspection intensity versus lower intensity, as in: length of inspection and amount of fine
Outcomes	Primary outcome, exposure: Violation of safety standard (machine guarding)

Notes	Time of the intervention: 1972-1991 Firm size: small, average 52 employees Type of work: mostly physical work Previous inspections: not reported Baseline injury rates: not reported Funding: National Science Foundation, Boston University Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Worker not blinded, influence on outcome possible
Blinding (outcome assessors)	High risk	Not blinded, inspectors measuring outcome themselves (violation of safety standard)
Retrospective unplanned subgroup analyses	Low risk	Only one regression model
Follow-up	Low risk	Same length
Statistical tests	Low risk	Logistic regression
Compliance	Unclear risk	Not reported
Outcome measures	High risk	Self reported
Selection bias (population)	Low risk	Same population
Selection bias (time)	Low risk	Same time period
Randomisation	High risk	No randomisation
Allocation concealment	High risk	Not concealed, not randomised
Adjustment for confounding	High risk	Not adjusted, differences not measured for pre-intervention inspections and injury rates
Incomplete outcome data	Unclear risk	Not reported

Weil 2001

Methods	Panel study
Participants	USA Firms, construction industry N = 2060
Interventions	Inspection with and without penalties • Any type of inspection by OSHA including complaint, accident or fatality investigation Comparisons: • Complaint inspection versus any inspection • Accident/fatality investigation versus any inspection • More inspections versus fewer inspections • Higher inspection intensity versus lower intensity, as in: length of inspection and amount of fine
Outcomes	Primary outcome, exposure: Violation of safety standard (machine guarding)
Notes	Time of the intervention: 1987-1993 Firm size: big firms Type of work: mostly physical work Previous inspections: not reported Baseline injury rates: not reported Funding: NIOSH Conflict of interest: no

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Blinding (subjects)	High risk	Not blinded
Blinding (outcome assessors)	High risk	Not blinded, inspectorate measure compliance
Retrospective unplanned subgroup analyses	High risk	Many models
Follow-up	Low risk	Same follow-up for all
Statistical tests	Low risk	Regression analysis
Compliance	Unclear risk	Not measured
Outcome measures	Unclear risk	Measured by different inspectorates
Selection bias (population)	Low risk	Same type of industry
Selection bias (time)	Low risk	Same time

Weil 2001 (*Continued*)

Randomisation	High risk	Not randomised
Allocation concealment	High risk	Not concealed
Adjustment for confounding	Unclear risk	Not adjusted, not measured for pre-intervention inspections and injury rates
Incomplete outcome data	Unclear risk	Not reported

- C: control
- CBA: controlled before-after study
- FMCSA: Federal Motor Carrier Safety Administration
- FTE: full-time equivalent
- I: intervention
- MSHA: Mine Safety and Health Administration
- ITS: interrupted time series study
- NIOSH: National Institute for Occupational Safety and Health
- OSH: occupational safety and health
- OSHA: Occupational Safety and Health Administration
- SIC: Standard Industrial Classification
- SFY: state fiscal year

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

Study	Reason for exclusion
Adams 2007	Not about enforcement
Attfield 1992	Not about enforcement, but legislation only
Auld 2001	Analysis on sub-industry not firm or workplace level
Baron-Epel 2012	Qualitative study, not about enforcement of occupational health and safety but of a smoking ban legislation
Boden 1985	Panel data, no time variable included in regression analysis, excluded as cross-sectional study
Brown 2003	Opinion paper, single person
Joy 2007	Evaluation of stricter regulation, not a study of variation of enforcement tools
Ko 2010	Panel study, no time lag variable, excluded as cross-sectional study
Lissner 2011	Qualitative study, not about enforcement but legislation only

(*Continued*)

Mancini 2005	Enforcement of only part of the assessed intervention
Morantz 2009	Panel study that used the same data as Weil 2001. The analysis included state versus federal inspectors. This is not an intervention that could be easily applied and neither is this a factor that is easily explained. We decided therefore to exclude this study to prevent counting studies twice
Niskanen 2013	Survey, only inspectors' opinions not employers' or employees' opinions
Raymond 2003	Not about occupational health and safety
Smitha 2001	Missing time lag variable
Viscusi 1979	Panel data, outcome measured at aggregate industry level not at firm or individual workplace level

Characteristics of studies awaiting assessment *[ordered by study ID]*

Bordas 2001

Methods	Technique of analyses: triangulation Data collection: observations (behaviour, at the workplace) and interviews (formal, informal and structured)
Participants	USA, east central Alabama, 1998 Worker and employer, logging industry N = unknown (5 crews with 2 to 15 workers) Firm size: small Type of work: mostly physical Previous inspections: not reported
Interventions	Any by OSHA
Outcomes	Main outcome: hazard and safety perception Including workers' and employers' perception of OSHA enforcement
Notes	Funding: United States forest servings Conflict of interest: no

Geminiani 2008

Methods	Data collection: survey, questionnaire Theory driven Technique of analysis: descriptive analysis
Participants	Republic of South Africa, time of study unclear Civil and building constructors, construction industry N = 626 included, 107 respondents Firm size: not reported

Geminiani 2008 (*Continued*)

	Type of work: mostly office work Previous inspections: not reported
Interventions	Inspectorate in general
Outcomes	Main outcome: effectiveness of labour inspections Including: opinions and beliefs regarding inspectorates
Notes	Funding: not reported Conflict of interest: no

Gillen 2004

Methods	Technique of analyses: content analysis, thematic Data collection: focus group with semi-structured interviews and behavioural observations
Participants	USA, California, 2000 Safety managers, construction industry N = 22 Firm size: any, average not reported Type of work: mostly physical Previous inspections: not reported
Interventions	Any enforcement activity by Cal/OSHA
Outcomes	Main outcome: perception of construction safety practices Including: managers' opinion and beliefs about OSHA enforcement
Notes	Funding: California Department of Health Services Conflict of interest: no

Gray 2006

Methods	Technique of analyses: narrative, grounded ethnographic themes Data collection: in depth participative observation (5 months)
Participants	Canada, time of study not reported Workers and employer, manufacturing industry N = 1 firm Firm size: not reported Type of work: mostly physical Previous inspections: not reported
Interventions	Any enforcement activity of occupational health and safety standards eligible, analysed for inspections
Outcomes	Main outcome: the role of worker agencies in regulatory enforcement Including: reaction by workforce towards planned inspections

Gray 2006 (*Continued*)

Notes	Funding: not reported Conflict of interest: no

Guidotti 1996

Methods	Technique of analyses: descriptive (cross-tabulation) Data collection: telephone survey
Participants	Canada, Fort Mc Murray region, 1992 Workers, sand oil industry (predominantly 30 to 44 years old (55%), 96% male) N = 150 Firm size: not reported Type of work: mostly physical Previous inspections: not reported
Interventions	Any enforcement activity of occupational health and safety standards
Outcomes	Main outcome: health- and safety-related behaviour among oil sands workers Including: opinion towards enforcement of occupational health and safety standard even if it increases cost or time to complete a job
Notes	Funding: Occupational Health and Safety Heritage Grant Program of Labour Conflict of interest: no

Mayhew 1999

Methods	Technique of analyses: content analyses, quantitative Data collection: semi-structured face-to-face interview plus questionnaire with closed and open-ended questions
Participants	Australia, Queensland, 1997 Demolishers, construction industry N = 31 (18.4% of 168 included workers) Firm size: not reported Type of work: mostly physical Previous inspections: some yes some no, type and time not reported
Interventions	Inspection/audit with or without further enforcement activity by jurisdiction
Outcomes	Main outcome: impact on OSH performance Including opinion and beliefs about inspection with or without further enforcement
Notes	Funding: DETIR and National Occupational Health and Safety Commission Conflict of interest: No

- OSHA: Occupational Safety and Health Administration

DATA AND ANALYSES

Comparison 1. Inspection versus no intervention

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Non-fatal injuries, short-term, RCT	1		Rate Ratio (Random, 95% CI)	Totals not selected
1.1 Unannounced inspection with orders, mostly physical work, unknown firm size, no inspections 2 years prior	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
2 Fatal and non-fatal injuries, short-term, CBA	1		Risk Ratio (Random, 95% CI)	Totals not selected
2.1 Random inspection with or without penalty (type of penalty unknown), mostly physical work, small firms, no inspections 2 years prior intervention	1		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
3 Non-fatal injuries, short-term, panel study	3		Risk Ratio (Random, 95% CI)	0.92 [0.89, 0.95]
3.1 Inspection (type unknown) with or without penalty (type unknown), mostly physical work, any firm size, prior inspections unknown	2		Risk Ratio (Random, 95% CI)	0.92 [0.87, 0.97]
3.2 Inspection (type unknown) with or without citations, mixed type of work (non-fixed site), unknown firm size, no inspections 1 year prior intervention	1		Risk Ratio (Random, 95% CI)	0.92 [0.86, 0.99]
3.3 Inspection (type unknown) with or without citations, mixed type of work (fixed site), unknown firm size, no inspections 1 year prior intervention	1		Risk Ratio (Random, 95% CI)	0.92 [0.86, 0.99]
4 Fatal and non-fatal injuries, medium-term, CBA	1		Risk Ratio (Random, 95% CI)	Totals not selected
4.1 Random inspection with or without penalty (type unknown), mostly physical work, small firms, no inspections 2 years prior intervention	1		Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
5 Non-fatal injuries, medium-term, panel study	3		Risk Ratio (Random, 95% CI)	0.97 [0.94, 1.01]

5.1 Programmed and complaint inspection with or without penalty (type unknown), mixed type of work, unknown firm size, prior inspections unknown	3	Risk Ratio (Random, 95% CI)	0.97 [0.94, 1.01]
6 Fatal and non-fatal injuries, long-term, CBA	1	Risk Ratio (Random, 95% CI)	Totals not selected
6.1 Random inspection with or without penalty (type unknown), mostly physical work, small firms, no inspections 2 years prior intervention	1	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
7 Fatal and non-fatal injuries, long-term, ITS-level	1	Std. Mean Difference (Random, 95% CI)	Totals not selected
7.1 Inspections with or without warning and orders, mostly physical work, unknown firm size, prior inspections unknown	1	Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
8 Fatal and non-fatal injuries, long-term, ITS-slope	1	Std. Mean Difference (Random, 95% CI)	Totals not selected
8.1 Inspection with or without warning and orders, mostly physical work, unknown firm size, prior inspections unknown	1	Std. Mean Difference (Random, 95% CI)	0.0 [0.0, 0.0]
9 Reduced Exposure, medium-term, RCT	1	Risk Ratio (Random, 95% CI)	Totals not selected
9.1 Announced inspections with or without warning and orders, type of work not reported	1	Risk Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 2. Inspection type versus any other type of Inspection with or without penalties

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, exposure (compliance; < 1 violation), panel study	2		Rate Ratio (Random, 95% CI)	Totals not selected
1.1 Follow-up inspection with or without penalty (type unknown), mostly physical work, small firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

1.2 Complaint inspection with or without penalty (type unknown), mostly physical work, small firms, inspections prior unknown	1	Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
1.3 Complaint inspection with or without penalty (type unknown), mostly physical work, big firms, inspections prior unknown	1	Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
1.4 Accident investigation with or without penalty (type unknown), mostly physical work, big firms, inspections prior unknown	1	Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 3. Inspection with citation versus inspection without citation

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, non-fatal injuries, panel study	1		Mean Difference (Fixed, 95% CI)	Totals not selected
1.1 Any type of inspection with citation, mostly physical work, big firms, inspections prior unknown	1		Mean Difference (Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Medium-term, non-fatal injuries, panel studies	1		Mean Difference (Fixed, 95% CI)	Totals not selected
2.1 Any type of inspection with citation, mostly physical work, big firms, inspections prior unknown	1		Mean Difference (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 4. Inspection with penalty versus no intervention or inspection only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, non-fatal injury, panel study	1		Odds Ratio (Random, 95% CI)	Totals not selected
1.1 Programmed or complaint inspection with penalty (type unknown), mostly physical work, small firms, prior inspections unknown	1		Odds Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

2 Medium-term, non-fatal injury, panel study	1	Odds Ratio (Random, 95% CI)	Totals not selected
2.1 Programmed or complaint inspection with penalty (type unknown), mostly physical work, small firms, prior inspections unknown	1	Odds Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 5. More penalties versus fewer penalties

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, exposure compliance (< 1 violation), panel study	2		Rate Ratio (Fixed, 95% CI)	Totals not selected
1.1 Inspection (type unknown) with penalty (type unknown), mostly physical work, small firms, inspections prior unknown	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Inspection (type unknown) with penalty (type unknown), mostly physical work, big firms, inspections prior unknown	1		Rate Ratio (Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 6. First inspection versus more than one inspection

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, exposure compliance (< 1 violation), panel study	2		Rate Ratio (Random, 95% CI)	Totals not selected
1.1 Any inspection with or without penalty (type unknown), mostly physical work, small firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]
1.2 Any inspection with or without penalty (type unknown), mostly physical work, big firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 7. Six inspections versus more than six inspections

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, exposure compliance (< 1 violation), panel study	1		Rate Ratio (Random, 95% CI)	Totals not selected
1.1 Any inspection with or without penalties (type unknown), mostly physical work, big firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Comparison 8. More inspection hours versus fewer hours

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Short-term, exposure compliance (< 1 violation), panel study	2		Rate Ratio (Random, 95% CI)	0.96 [0.94, 0.99]
1.1 Any type of inspection with or without penalty (type unknown), mostly physical work, small firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	1.06 [0.87, 1.28]
1.2 Any type of inspection with or without penalty (type unknown), mostly physical work, big firms, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.96 [0.93, 0.99]

Comparison 9. Autonomy oriented versus coercive oriented inspectors

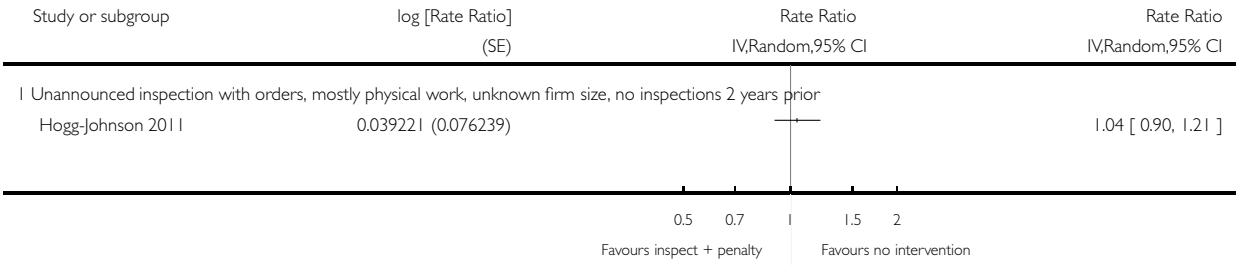
Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Long-term, exposure (number of visits needed to resolve non-compliance), panel study	1		Rate Ratio (Random, 95% CI)	Totals not selected
1.1 Type of inspection unknown with or without compliance orders, type of work unknown, firm size unknown, inspections prior unknown	1		Rate Ratio (Random, 95% CI)	0.0 [0.0, 0.0]

Analysis 1.1. Comparison 1 Inspection versus no intervention, Outcome 1 Non-fatal injuries, short-term, RCT.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 1 Non-fatal injuries, short-term, RCT

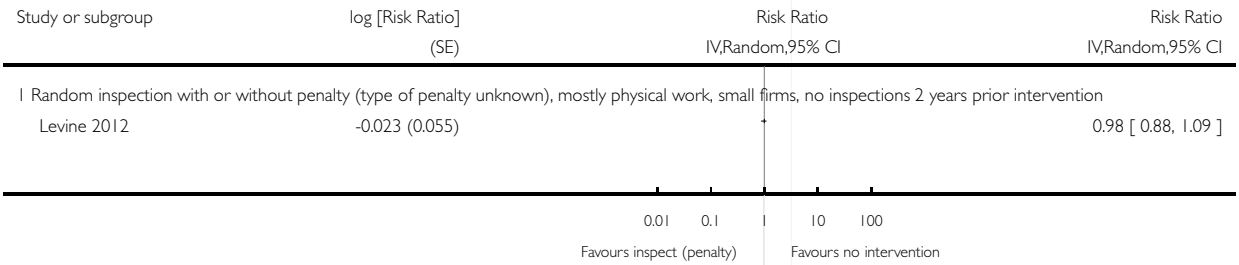


Analysis 1.2. Comparison 1 Inspection versus no intervention, Outcome 2 Fatal and non-fatal injuries, short-term, CBA.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 2 Fatal and non-fatal injuries, short-term, CBA



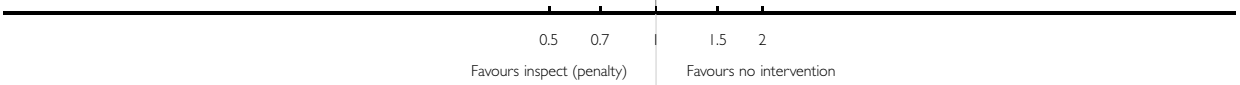
Analysis 1.3. Comparison 1 Inspection versus no intervention, Outcome 3 Non-fatal injuries, short-term, panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 3 Non-fatal injuries, short-term, panel study

Study or subgroup	log [Risk Ratio] (SE)	Risk Ratio IV,Random,95% CI	Weight	Risk Ratio IV,Random,95% CI
1 Inspection (type unknown) with or without penalty (type unknown), mostly physical work, any firm size, prior inspections unknown				
Smith 1979a	-0.07377 (0.037968)		23.1 %	0.93 [0.86, 1.00]
Smith 1979b	-0.10023 (0.037927)		23.2 %	0.90 [0.84, 0.97]
Subtotal (95% CI)			**46.3 %**	**0.92 [0.87, 0.97]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.24$, df = 1 (P = 0.62); $I^2 = 0.0\%$				
Test for overall effect: Z = 3.24 (P = 0.0012)				
2 Inspection (type unknown) with or without citations, mixed type of work (non-fixed site), unknown firm size, no inspections 1 year prior intervention				
Foley 2012	-0.078 (0.0362)		25.5 %	0.92 [0.86, 0.99]
Subtotal (95% CI)			**25.5 %**	**0.92 [0.86, 0.99]**
Heterogeneity: not applicable				
Test for overall effect: Z = 2.15 (P = 0.031)				
3 Inspection (type unknown) with or without citations, mixed type of work (fixed site), unknown firm size, no inspections 1 year prior intervention				
Foley 2012	-0.0823 (0.0344)		28.2 %	0.92 [0.86, 0.99]
Subtotal (95% CI)			**28.2 %**	**0.92 [0.86, 0.99]**
Heterogeneity: not applicable				
Test for overall effect: Z = 2.39 (P = 0.017)				
Total (95% CI)			**100.0 %**	**0.92 [0.89, 0.95]**
Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 0.28$, df = 3 (P = 0.96); $I^2 = 0.0\%$				
Test for overall effect: Z = 4.57 (P < 0.00001)				
Test for subgroup differences: $Chi^2 = 0.04$, df = 2 (P = 0.98), $I^2 = 0.0\%$				

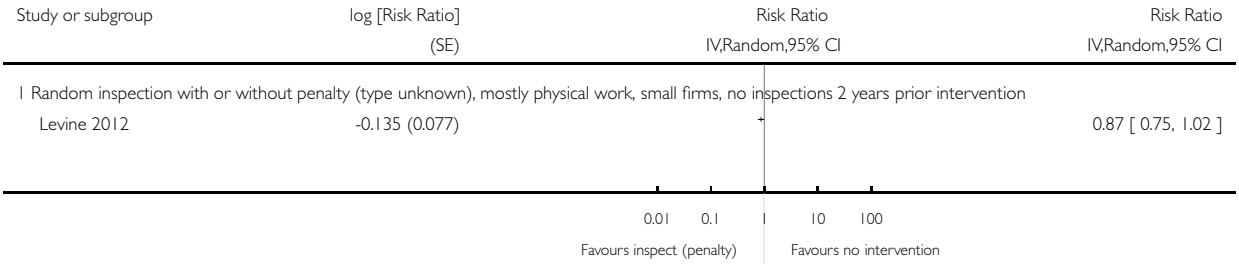


Analysis 1.4. Comparison 1 Inspection versus no intervention, Outcome 4 Fatal and non-fatal injuries, medium-term, CBA.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 4 Fatal and non-fatal injuries, medium-term, CBA

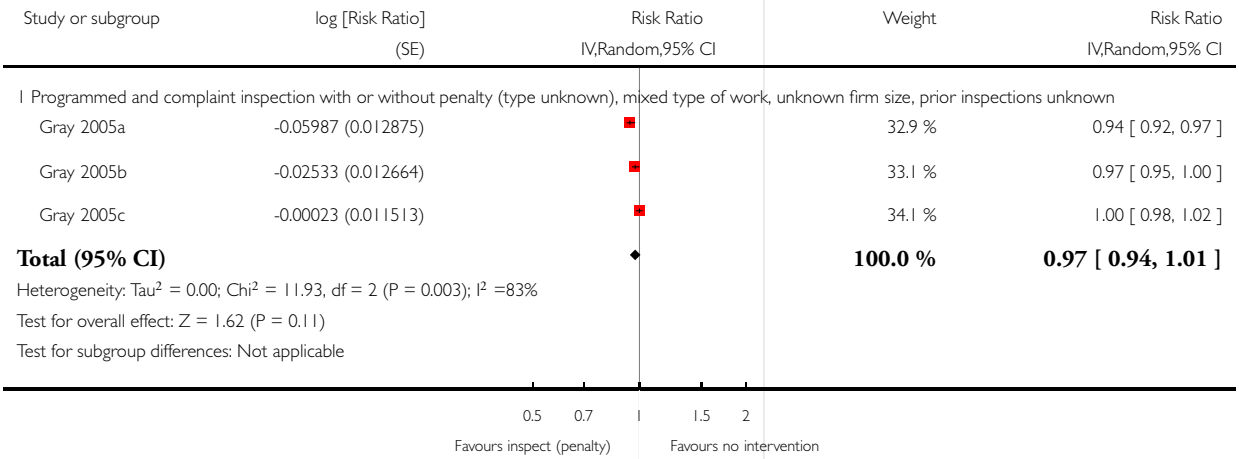


Analysis 1.5. Comparison 1 Inspection versus no intervention, Outcome 5 Non-fatal injuries, medium-term, panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 5 Non-fatal injuries, medium-term, panel study

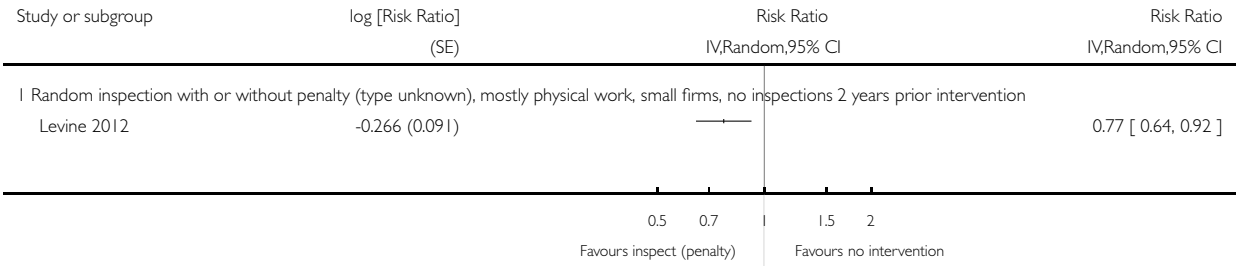


Analysis 1.6. Comparison 1 Inspection versus no intervention, Outcome 6 Fatal and non-fatal injuries, long-term, CBA.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 6 Fatal and non-fatal injuries, long-term, CBA

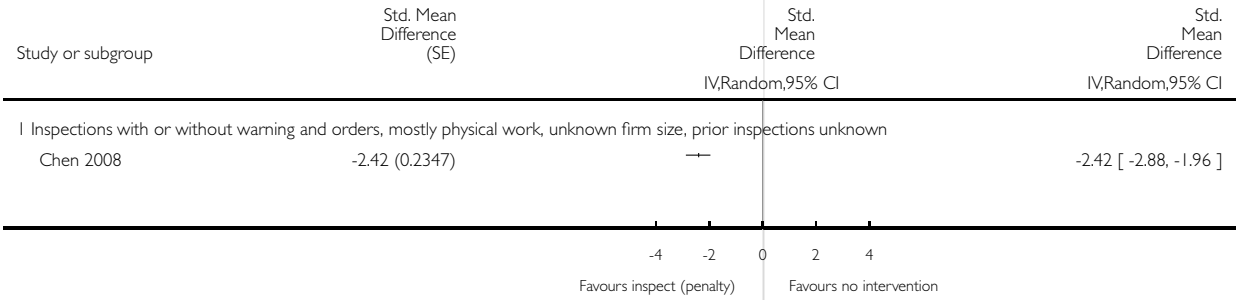


Analysis 1.7. Comparison 1 Inspection versus no intervention, Outcome 7 Fatal and non-fatal injuries, long-term, ITS-level.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 7 Fatal and non-fatal injuries, long-term, ITS-level

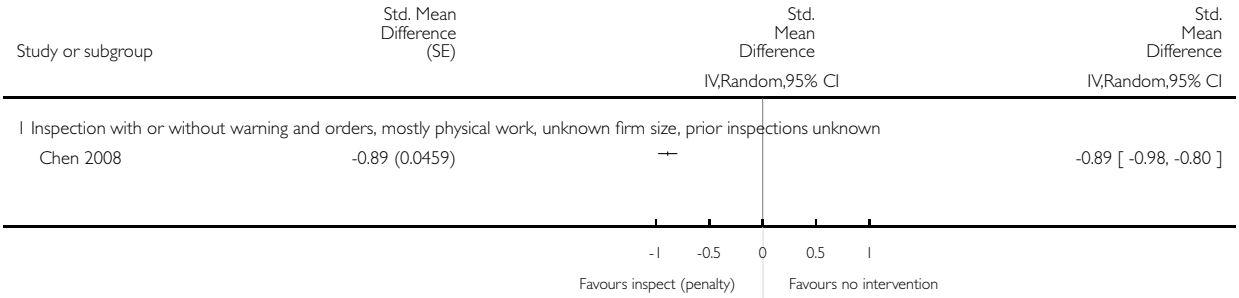


Analysis 1.8. Comparison 1 Inspection versus no intervention, Outcome 8 Fatal and non-fatal injuries, long-term, ITS-slope.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 8 Fatal and non-fatal injuries, long-term, ITS-slope

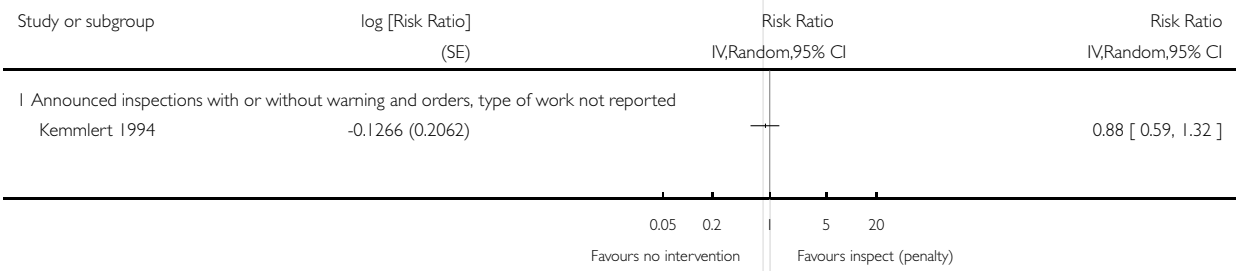


Analysis 1.9. Comparison 1 Inspection versus no intervention, Outcome 9 Reduced Exposure, medium-term, RCT.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 1 Inspection versus no intervention

Outcome: 9 Reduced Exposure, medium-term, RCT

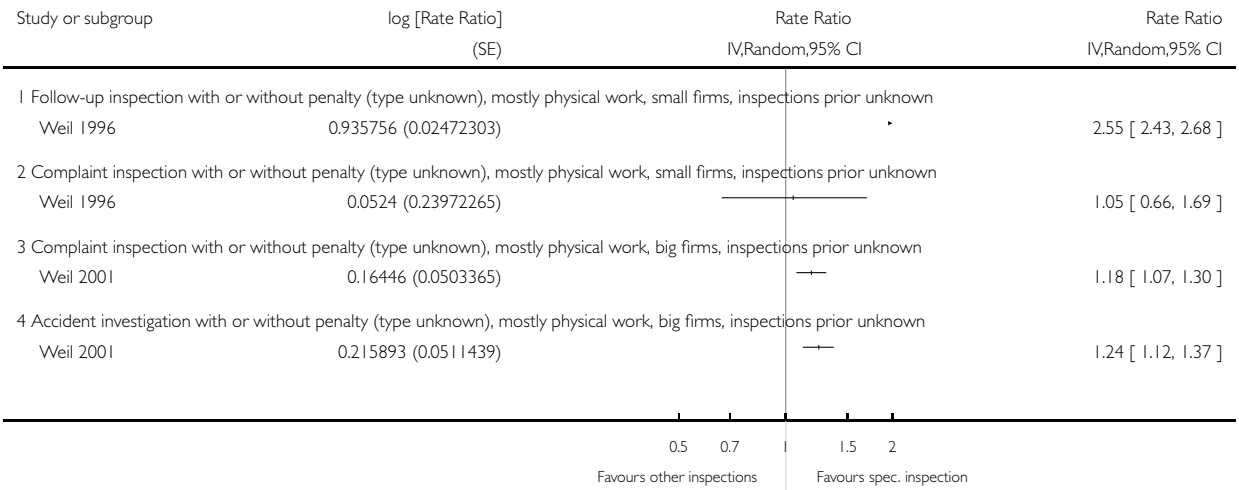


Analysis 2.1. Comparison 2 Inspection type versus any other type of Inspection with or without penalties, Outcome 1 Short-term, exposure (compliance; < 1 violation), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 2 Inspection type versus any other type of Inspection with or without penalties

Outcome: 1 Short-term, exposure (compliance; < 1 violation), panel study

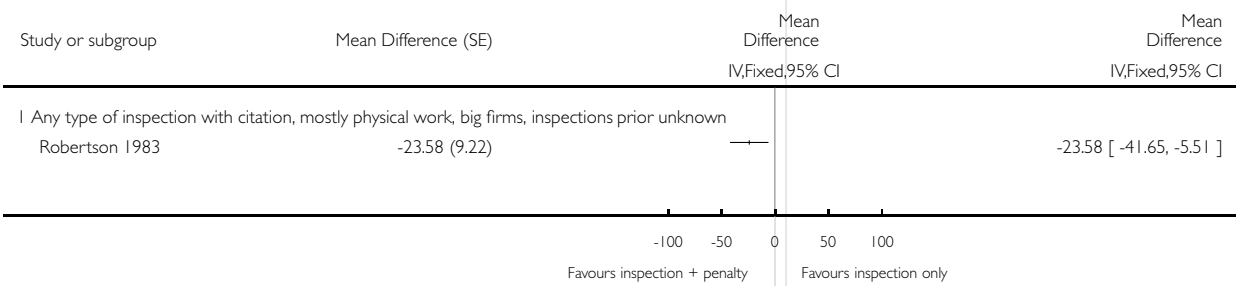


Analysis 3.1. Comparison 3 Inspection with citation versus inspection without citation, Outcome 1 Short-term, non-fatal injuries, panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 3 Inspection with citation versus inspection without citation

Outcome: 1 Short-term, non-fatal injuries, panel study

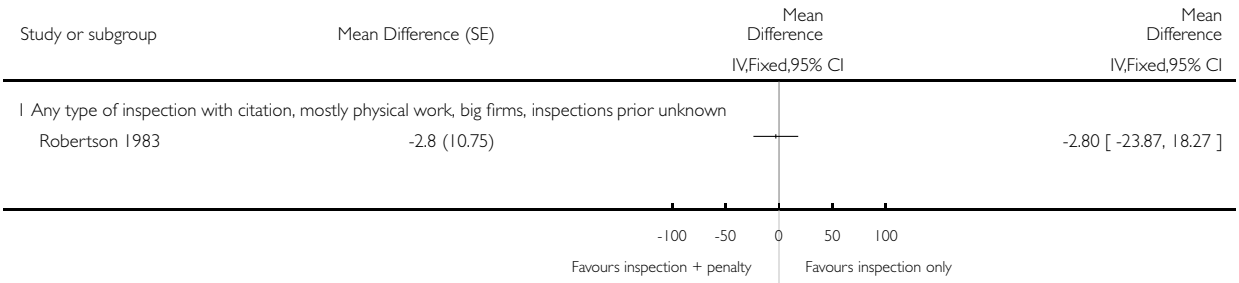


Analysis 3.2. Comparison 3 Inspection with citation versus inspection without citation, Outcome 2 Medium-term, non-fatal injuries, panel studies.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 3 Inspection with citation versus inspection without citation

Outcome: 2 Medium-term, non-fatal injuries, panel studies

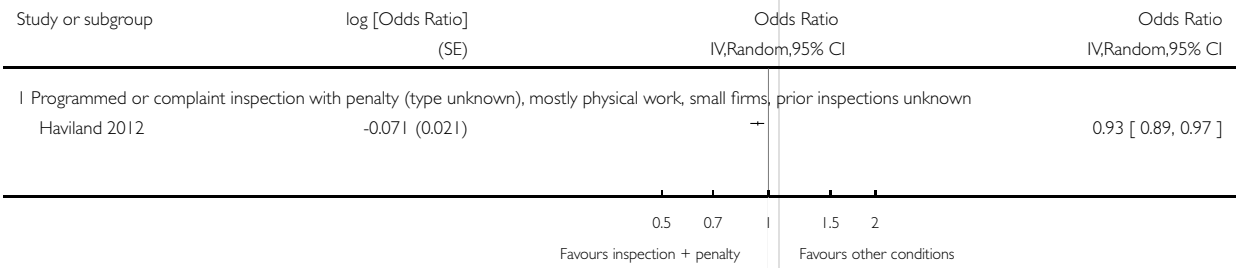


Analysis 4.1. Comparison 4 Inspection with penalty versus no intervention or inspection only, Outcome 1 Short-term, non-fatal injury, panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 4 Inspection with penalty versus no intervention or inspection only

Outcome: 1 Short-term, non-fatal injury, panel study

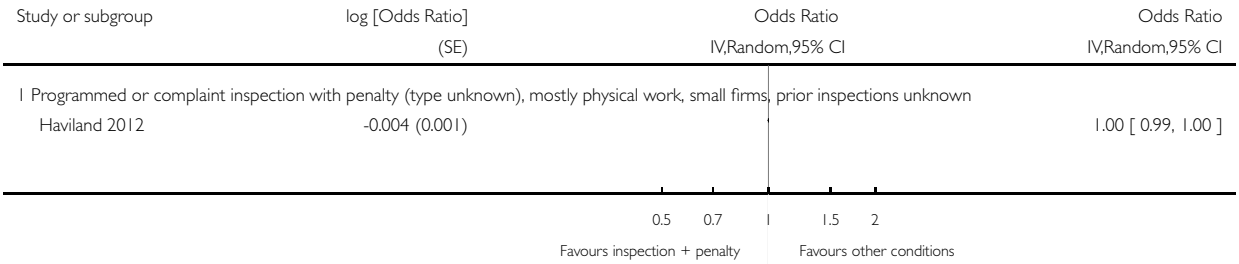


Analysis 4.2. Comparison 4 Inspection with penalty versus no intervention or inspection only, Outcome 2 Medium-term, non-fatal injury, panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 4 Inspection with penalty versus no intervention or inspection only

Outcome: 2 Medium-term, non-fatal injury, panel study

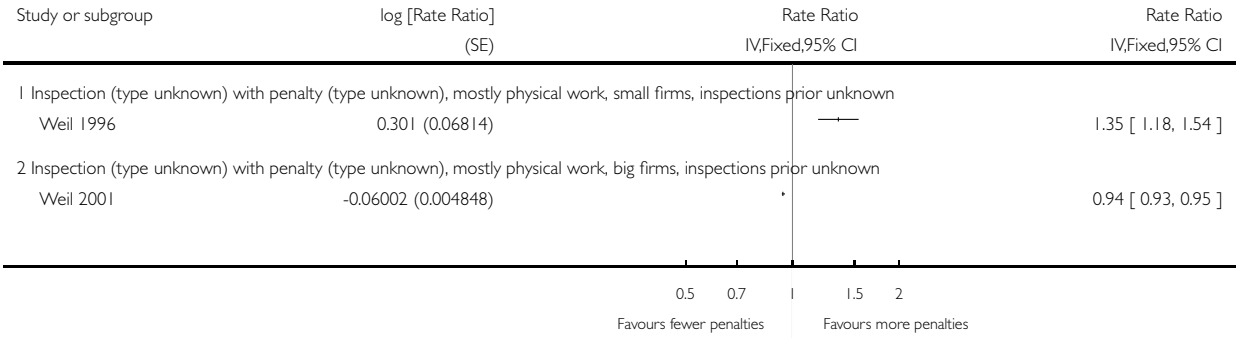


Analysis 5.1. Comparison 5 More penalties versus fewer penalties, Outcome 1 Short-term, exposure compliance (< 1 violation), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 5 More penalties versus fewer penalties

Outcome: 1 Short-term, exposure compliance (< 1 violation), panel study

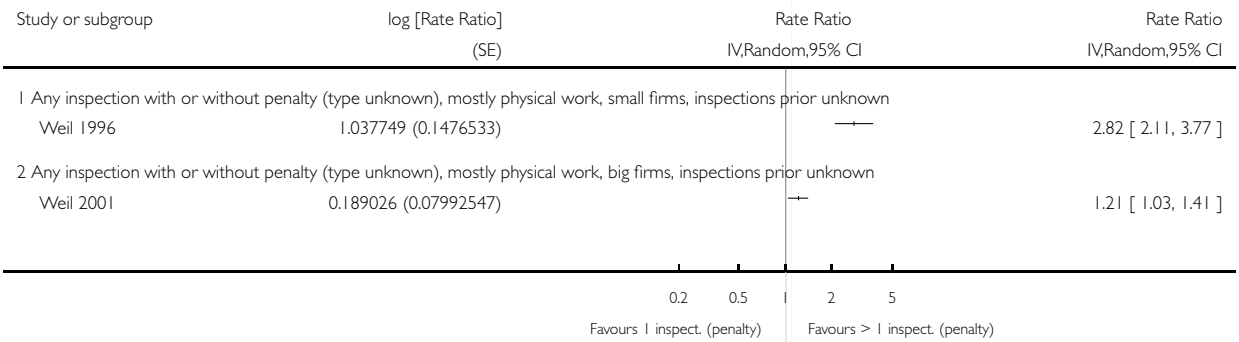


Analysis 6.1. Comparison 6 First inspection versus more than one inspection, Outcome 1 Short-term, exposure compliance (< 1 violation), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 6 First inspection versus more than one inspection

Outcome: 1 Short-term, exposure compliance (< 1 violation), panel study



Analysis 7.1. Comparison 7 Six inspections versus more than six inspections, Outcome 1 Short-term, exposure compliance (< 1 violation), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 7 Six inspections versus more than six inspections

Outcome: 1 Short-term, exposure compliance (< 1 violation), panel study

Study or subgroup	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Rate Ratio IV,Random,95% CI
1 Any inspection with or without penalties (type unknown), mostly physical work, big firms, inspections prior unknown			
Weil 2001	0.000948 (0.000688)		1.00 [1.00, 1.00]

Analysis 8.1. Comparison 8 More inspection hours versus fewer hours, Outcome 1 Short-term, exposure compliance (< 1 violation), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 8 More inspection hours versus fewer hours

Outcome: 1 Short-term, exposure compliance (< 1 violation), panel study

Study or subgroup	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Weight	Rate Ratio IV,Random,95% CI
1 Any type of inspection with or without penalty (type unknown), mostly physical work, small firms, inspections prior unknown				
Weil 1996	0.05663 (0.09732)		2.0 %	1.06 [0.87, 1.28]
Subtotal (95% CI)			**2.0 %**	**1.06 [0.87, 1.28]**
Heterogeneity: not applicable				
Test for overall effect: Z = 0.58 (P = 0.56)				
2 Any type of inspection with or without penalty (type unknown), mostly physical work, big firms, inspections prior unknown				
Weil 2001	-0.04156 (0.013853)		98.0 %	0.96 [0.93, 0.99]
Subtotal (95% CI)			**98.0 %**	**0.96 [0.93, 0.99]**
Heterogeneity: not applicable				
Test for overall effect: Z = 3.00 (P = 0.0027)				
Total (95% CI)			**100.0 %**	**0.96 [0.94, 0.99]**

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Weight	Rate Ratio IV,Random,95% CI

Heterogeneity: $Tau^2 = 0.0$; $Chi^2 = 1.00$, df = 1 (P = 0.32); $I^2 = 0.0\%$
Test for overall effect: Z = 2.89 (P = 0.0039)
Test for subgroup differences: $Chi^2 = 1.00$, df = 1 (P = 0.32), $I^2 = 0.0\%$

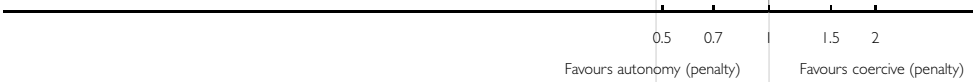
Analysis 9.1. Comparison 9 Autonomy oriented versus coercive oriented inspectors, Outcome 1 Long-term, exposure (number of visits needed to resolve non-compliance), panel study.

Review: Occupational safety and health enforcement tools for preventing occupational diseases and injuries

Comparison: 9 Autonomy oriented versus coercive oriented inspectors

Outcome: 1 Long-term, exposure (number of visits needed to resolve non-compliance), panel study

Study or subgroup	log [Rate Ratio] (SE)	Rate Ratio IV,Random,95% CI	Rate Ratio IV,Random,95% CI
1 Type of inspection unknown with or without compliance orders, type of work unknown, firm size unknown, inspections prior unknown			
Burstyn 2010	-0.0141 (0.0047)		0.99 [0.98, 1.00]



ADDITIONAL TABLES

Table 1. Characteristics of panel studies

Article ID	Industry type*	Country*	Outcome*	Time Span	Data source		N of variables	Interven-tion type*	Study ID
					IMIS	BLS			

Table 1. Characteristics of panel studies (*Continued*)

Weil 2001	Construction	USA	Exposure	87 - 93	x	x	20	Inspection with or without penalty	Weil 2001
Morantz 2007	Construction	USA	Exposure	87 - 93	x				
Weil 1996	Custom woodwork	USA	Exposure	72 - 91	x		13	Inspection with or without penalty	Weil 1996
Robertson 1983	Manufacturing	USA	Injury	73 - 80		x	5	Inspection and penalty	Robertson 1983
Scholz 1990	Manufacturing	USA	Injury	79 - 85	x	x		Inspection with or without penalty	Gray 2005a
Scholz 1997	Manufacturing	USA	Injury	79 - 85	x	x		Inspection with or without penalty	
Gray 2005a	Manufacturing	USA	Injury	79 - 85	x	x	10	Inspection with or without penalty	
Gray 2005b	Manufacturing	USA	Injury	87 - 91	x	x	10	Inspection with or without penalty	Gray 2005b
Mendelhoff 2005	Manufacturing	USA	Injury	92 - 98	x	x		Inspection with or without penalty	Gray 2005c
Gray 2005c	Manufacturing	USA	Injury	92 - 98	x	x	10	Inspection with or without penalty	
Haviland 2012	Manufacturing	USA	Injury	98 - 05	x		7	Inspection with penalty	Haviland 2012
Haviland 2010	Manufacturing	USA	Injury	98 - 05	x			Inspection with penalty	
Haviland 2008	Manufacturing	USA	Injury	98 - 05	x			Inspection with penalty	
Kniesner 2004	Mining	USA	Injury	83 - 97				Inspection with penalty	Kniesner 2004

Table 1. Characteristics of panel studies (*Continued*)

Burstyn 2010	Mixed	Canada	Exposure	03 - 06				Inspection with or without penalty	Burstyn 2010
Foley 2012	Mixed	USA	Injury	99 - 00	x			Inspection with or without penalty	Foley 2012
Smith 1979a	Mixed	USA	Injury	72 - 73		x		Inspection with or without penalty	Smith 1979a
Smith 1979b	Mixed	USA	Injury	73 - 74		x		Inspection with or without penalty	Smith 1979b

*Articles with same characteristics are considered same study.

- IMIS: Integrated Management Information System, (Occupational Safety and Health Administration, USA)
- BLS: Bureau of Labor Statistics, (USA)

Table 2. Characteristics of included qualitative studies

Bordas 2001	
Methods	Technique of analyses: triangulation Data collection: observations (behaviour, at the workplace) and interviews (formal, informal and structured)
Participants	USA, east central Alabama, 1998 Worker and employer, logging industry N = unknown (5 crews with 2 to 15 workers) Firm size: small Type of work: mostly physical Previous inspections: not reported
Intervention	Any by OSHA
Outcome	Main outcome: hazard and safety perception Including workers' and employers' perception of OSHA enforcement
Notes	Study year: 1998 Funding: United States forest servings Conflict of interest: no
Geminiani 2008	

Table 2. Characteristics of included qualitative studies (*Continued*)

Methods	Data collection: survey questionnaire Theory driven Technique of analysis: descriptive analysis
Participants	Republic of South Africa, time of study unclear Civil and building constructors, construction industry N = 626 included, 107 responses Firm size: not reported Type if work: mostly office work Previous inspections: not reported
Intervention	Inspections in general
Outcome	Main outcome: effectiveness of labour inspections Including: opinions and beliefs regarding Inspections
Notes	Funding: not reported Conflict of interest: none
Gillen 2004	
Methods	Technique of analysis: content analysis, thematic Data collection: focus group with semi-structured interviews and behavioural observations
Participants	USA, California, 2000 Safety managers, construction industry N = 22 Firm size: any, average not reported Type of work: mostly physical Previous inspections: not reported
Intervention	Any enforcement by Cal/OSHA
Outcome	Main outcome: perception of construction safety practices Including: managers' opinions and beliefs about OSHA enforcement
Gray 2006	
Methods	Technique of analyses: narrative, grounded ethnographic themes Data collection: in depth participative observation (5 months)
Participants	Canada, time of study not reported Workers and employee, manufacturing industry N = 1 firm Firm size: not reported Type of work: mostly physical Previous inspections: not reported

Table 2. Characteristics of included qualitative studies (*Continued*)

Intervention	Any enforcement activity of occupational health and safety standards eligible, analysed for inspections
Outcome	Main outcome: the role of worker agencies in regulatory enforcement Including: reaction by workforce towards planned inspections
Notes	Study year: not reported Funding: not reported Conflict of interest: no
Guidotti 1996	
Methods	Technique of analyses: descriptive (cross-tabulation) Data collection: telephone survey
Participants	Canada, Fort McMurray region, 1992 Workers, sand oil industry (predominantly 30 to 44 years old (55%), 96% male) N = 150 Firm size: not reported Type of work: mostly physical Previous inspections: not reported
Intervention	Any enforcement activity of occupational health and safety standards
Outcome	Main outcome: health- and safety-related behaviour among oil sands workers Including: opinion towards enforcement of occupational health and safety standard even if it increases cost or time to complete a job
Notes	Study year: 1992 Funding: Occupational Health and Safety Heritage Grant Program of Labour Conflict of interest: no
Mayhew 1999	
Methods	Technique of analyses: content analyses, quantitative Data collection: semi-structured face-to-face interview plus questionnaire with closed and open-ended questions
Participants	Australia, Queensland, 1997 Demolishers, construction industry N = 31 (18.4% of 168 included workers) Firm size: not reported Type of work: mostly physical Previous inspections: some yes some no, type and time not reported
Intervention	inspection/audit with or without further enforcement activity by jurisdiction
Outcome	Main outcome: impact on OSH performance Including opinion and beliefs about inspection with or without further enforcement

Table 2. Characteristics of included qualitative studies (*Continued*)

Notes	Study year: 1997 Funding: DETIR and National Occupational Health and Safety Commission Conflict of interest: no

- OSHA: Occupational Safety and Health Administration

Table 3. Critical appraisal of included qualitative studies

Qualitative Study ID:	Bordas 2001	Geminiani 2008	Gillen 2004	Gray 2006	Guidotti 1996	Mayhew 1999
Method, Reporting - consistent and neutral						
1. Is there a clear connection between philosophical perspective, methodology, objectives, methods used to collect data, representation and analyses of data?	Unclear: applied method not described, no themes derived	No: review of literature, does not address the differences in contractor- inspector-attitudes	Yes	Yes	Unclear: not reported	Yes
2. Were the researchers open about potential bias (context presented and analysed in conclusion)?	Unclear: context not described	No: context rarely described, not included in conclusion	Unclear: context not described	Yes	Unclear: context not described	Unclear: context not described
3. Is the reporting clear and coherent (sampling method, recruitment conditions, inclusion and exclusion criteria, method of data collection, description of the derivation of themes and inclusion of	No: description of derivation of themes and inclusion of citations missing	No: descriptions of the derivation of the themes and inclusion of citations missing	Yes	No: method of data collection not described, description of derivation of themes missing	No: not reported	No: description of derivation of themes and inclusion of citations missing

Table 3. Critical appraisal of included qualitative studies (*Continued*)

supporting quotations)?						
Method, Subjects - credibility						
4. Is the recruitment of the study participants free of selection bias (e.g. workers are not selected by employer)?	Unclear: selection not fully described, voluntary participation	Unclear: survey with 18% response rate	Yes (various methods of outreach, includes monetary incentive, voluntary participation)	Unclear: one firm, recruitment not reported	Yes: via telephone register	Unclear: exclusion and drop-out unclear
5. Are characteristics of subjects and workplace considered for discussion and implication of findings? (age, gender, type of work, firm size, prior inspections)	No: subjects' age etc. not described, industry described, firm size and previous inspections missing	No: except for type of industry all descriptions missing	No: age, gender, ethnicity, firm size and previous inspection experiences not assessed	No: firm size and previous inspections missing (author withheld information to keep firm identity confidential)	No: none reported	No: age, gender, ethnicity, firm size and previous inspection experiences not assessed
6. Is the research ethical according to current criteria OR for recent studies is there evidence of ethical approval by an appropriate body?	Unclear: not reported	Unclear: not reported	Yes	Yes: ethical approval not reported but anonymity addressed	Unclear: not reported	Unclear: not reported
Analyses and conclusion - transferability						
7. Is/are the specific tool(s) of enforcement described? Are different tools analysed in separate categories?	No: interventions not specified	Yes: inspection only	No: no specification of uniform enforcement tools	Yes (outcome of interest only about planned inspection)	No: enforcement in general	No: types specified as intervention but combined in analyses (inspection with or without penalty)
8. Is a range of methods used to draw similar	Unclear: questionnaire and per-	No: questionnaire only	Unclear: focus group and observation used	No: only observation	No: description of firm size and previous inspec-	No triangulation

Table 3. Critical appraisal of included qualitative studies (*Continued*)

conclusions (triangulation)?	sonal interviews done, not reported from which source the opinions derived and if similar conclusions		but conclusions not reported		tions missing	
9. Does the representation of data fit the views of the participants studied (e.g. minimum 2 researchers independently analysed the data, or outside auditors, or participants validate the findings)?	Unclear: not reported	Unclear: not reported	Yes	No: no validating of findings, interpretation of data by one researcher	Unclear: not reported	Unclear: not reported
Total score (Yes) out of 9	0	1	5	4	1	1
Total quality * **High quality**: at least 4 YES in the 1^{st} group and 2 YES in the 2^{nd} group **Moderate quality**: at least 2 YES in the 1^{st} group and 3 YES in the 2^{nd} group **Low quality**: less than 2 YES in the 1^{st} group	**Low**	**Low**	**Moderate**	**Low**	**Low**	**Low**

*Questions are categorised into 1^{st} group (no. 4, 5, 7, 8 and 9) and 2^{nd} group questions (no. 1, 2, 3 and 6) according to the likelihood of influencing the outcome.

APPENDICES

Appendix 1. MEDLINE (PubMed) search strategy

#1 ("Occupational Diseases"[MeSH] OR "Occupational Health Services"[MeSH] OR "Occupational Health Nursing"[MeSH] OR "Accidents, Occupational"[MeSH] OR "Occupational Injuries"[MeSH] OR "Occupational Health Physicians"[MeSH] OR "Occupational Exposure"[MeSH] OR "Occupational Medicine"[MeSH] OR "Occupational Health"[MeSH] OR "Asthma, Occupational"[MeSH] OR "Noise, Occupational"[MeSH] OR "Dermatitis, Occupational"[MeSH] OR "Air Pollutants, Occupational"[MeSH] OR "National Institute for Occupational Safety and Health (U.S.)"[MeSH] OR "Industry"[MeSH] OR "Hand-Arm Vibration Syndrome"[MeSH] OR "Mineral Fibers"[MeSH] OR "Fatigue Syndrome, Chronic"[MeSH] OR "Pneumoconiosis"[MeSH] OR "Occupations"[MeSH] OR "Dermatitis, Contact"[MeSH] OR "Coal Tar"[MeSH] OR "Burnout, Professional"[MeSH] OR "Air Pollutants, Occupational" [Pharmacological Action])
#2 #1 AND (legislation and jurisprudence [sh])
#3 #2 OR (inspections[tiab] OR inspection[tw] OR audits[tw] OR audit[ti] OR citations[ti] OR citation[ti] OR warning[tw] OR warnings[tw] OR penalty[tw] OR penalties[tw] OR prosecution[tw] OR closure[tw] OR court[tw] OR violation[tw] OR violations[tw] OR offence[tw] OR fines[tw] OR enforcement[tw])
#4 #3 AND (("United States Occupational Safety and Health Administration"[MeSH]) OR "workplace"[tw] OR "work place"[tw] OR "establishment"[tw] OR "manufacturing plant"[tw] OR "manufacturing plants"[tw]OR "industrial plant"[tw] OR "industrial plants"[tw] OR firms[tw] OR company[tw] OR "labour inspectorate"[tw] OR factory[tw] OR manufactory[tw] OR mill[tw] OR foundry[tw] OR mining[MeSH] OR construction industry[MeSH])
#5 #4 NOT (animals[mh] NOT humans[mh])

Appendix 2. Search strategies for other databases

EMBASE	OSH update	Westlaw International
#1 'occupational disease'/exp OR 'occupational accident'/exp OR 'occupational exposure'/exp OR 'occupational safety'/exp OR 'risk assessment'/exp **#2** 'law enforcement'/exp OR 'government'/exp OR 'government regulation'/exp **#3** # 1 AND #2 **#4** inspection:ab,ti OR inspections:ab,ti OR citation:ab,ti OR citations:ab,ti OR warning:ab,ti OR warnings:ab,ti OR 'penalty'/exp OR penalty:ab,ti OR penalties:ab,ti OR prosecution:ab,ti OR violation:ab,ti OR violations:ab,ti OR offence:ab,ti OR fines:ab,ti OR enforcement:ab,ti **#5** #3 OR #4 **#6** 'work'/exp **#7** #5 AND #6 **#8** #7 NOT ([medline]/lim NOT [embase]/lim) **#9** #8 AND 'human'/de	DC{OUBIB or OUCISC or OUHSEL or OUNIOC or OUNIOS or OURILO} **#1** AB{regulation or legislation or enforcement} **#2** AB{inspection* or audit* or citation* or warning* or penalty or penalties or prosecution or violation* or offence or fines or incentive*} **#3** #8 and #9 **#4** GW {evaluation or effectiveness or attitude* or opinion or injury or injuries or occupational disease or exposure} **#5** GW{workplace or work place or establishment or manufacturing plant or manufacturing plants or industrial plant or industrial plants or firms or company or labour inspectorate or factory or manufactory or mill or foundry or industry or mine} **#6** #10 and #11 and #12	Database: world journals #1 occupational & injuries & (inspections osha "labour inspection" "health and safety executive") & "panel study"
CENTRAL	**CINAHL**	**PsycINFO**

(*Continued*)

inspections AND (firms OR workplaces OR occupational)	Inspections AND (firms OR workplaces OR occupational)	**#1** exp Occupational Health/ or exp Occupational Safety/ or exp Working Conditions/ or exp Risk Factors/ or exp Risk Management/ or exp Accident Prevention/ or exp Occupational Exposure/ or exp Work Related Illnesses/ or exp Musculoskeletal Disorders/ or exp Occupational Stress/ or burnout.mp. or occupational health services.mp. or occupational health nursing.mp. or occupational health physicians.mp. or occupational medicine.mp. or occupational asthma.mp. or occupational noise.mp. or occupational dermatitis.mp. or dermatitis/ or eczema/ or hand arm vibration syndrome.mp. or mineral fibers.mp. or chronic fatigue syndrome/ or pneumoconiosis.mp. or contact dermatitis.mp. or coal tar.mp. or pollution/ or hazardous materials/ or noise effects/ or passive smoking/ or physiological stress/ or (national institute for occupational safety and health).mp **#2** law enforcement/ or government/ or government regulation.mp **#3** 1 and 2 **#4** (inspection* or audit* or citation* or warning* or penalty or penalties or prosecution or violation* or offence or fines).mp **#5** 3 or 4 **#6** (workplace or "work place" or establishment or "manufacturing plant" or "manufacturing plants" or "industrial plant" or "industrial plants" or firms or company or "labour inspectorate" or factory or manufactory or mill or foundry or industry or mine).mp **#7** 5 and 6 **#8** limit 7 to human
Scopus	**EconLit**	**HeinOnline**
inspections AND (firms OR workplaces OR occupational) Searched in social sciences & humanities only	inspections AND (firms OR workplaces)	inspections AND (firms OR workplaces)

Appendix 3. Critical appraisal tool - qualitative studies

Method, reporting - consistency, neutrality		**Quote from article**
1. Is there a clear connection between philosophical perspective, methodology, objectives, methods used to collect data, representation and analyses of data? *(The way the conclusion are drawn and the themes are built, e.g. if data driven Grounded Theory method applied, if theory driven clear connection to theoretical framework/existing body of knowledge.)*	Yes No Unclear	
2. Were the researchers open about potential bias? *Answer* **NO***if description about both the sending and the receiving context presented but not analysed (e.g. researchers own beliefs).*	Yes No Unclear	
3. Is the reporting clear and coherent? (In terms of 3 domains: sampling method and recruitment conditions with inclusion and exclusion criteria, method of data collection, description of the derivation of themes and inclusion of supporting quotations) *Answer* **YES***if 3 out of 3 domains.*	Yes No Unclear	
Method, subjects - credibility		Quote from Article
4. Is the recruitment of the study participants appropriate (free of selection bias)? *Answer* **NO***if e.g. study subjects recruit themselves, or workers are selected by employer.*	Yes No Unclear	
5. Are the subjects in detail described (age, gender, ethnicity)? *(minimum age and gender)* Is the work place in detail described (type of industry/work, firm size, previous inspections)? *(minimum 2 out of 3)* *Answer* **NO***if only subjects Or workplace described.*	Yes No Unclear	
6. Is the research ethical according to current criteria OR for recent studies is there evidence of ethical approval by an appropriate body?	Yes No Unclear	

(*Continued*)

Analyses and conclusion - transferability		Quote from Article
7. Is the tool(s) of enforcement described? Are different tools analysed in separate categories (e.g. when more than one tool or combined tools (e.g. incentives and enforcement) are studied) *Answer* **YES***if tool(s) are specified (e.g. inspection) and analysed in different categories. Answer* **NO***if only stated as enforcement.*	Yes No Unclear	
8. Is a range of methods used to draw similar conclusions (triangulation)? *If more than one type of data collection but no similar conclusion answer* **NO**.	Yes No Unclear	
9. Does the representation of data fit the views of the participants studied? *Answer* **YES***if 2 or more researchers independently analysed the data OR outside auditors or participants validate the findings OR similar techniques.*	Yes No Unclear	

CONTRIBUTIONS OF AUTHORS

JV, AA, JJ, TM, SC and CM conceived and designed the review. CM co-ordinates the review process. JV wrote the protocol along with CM. KN, JV and CM developed the search strategy. RP, TM, KN, AA, JJ and SC provided comments.

DECLARATIONS OF INTEREST

None known.

SOURCES OF SUPPORT

Internal sources

- Finnish Institute of Occupational Health, Finland.

Provided facilities and IT support for Christina Mischke and Jos Verbeek

External sources

- WorkSafeBC, Canada.

Provided grant which paid for Christina Mischke's salary

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We addressed heterogeneity of studies as stated in the protocol, except for the type of inspection or penalty and the type of injury, disease and exposure. Instead of building different comparisons for each type of inspection or penalty we addressed heterogeneity of studies with different types of inspections in subgroups. The type of injury, disease and exposure is used to define the outcome level of the comparison instead of defining the subgroup.

We did not mention panel studies in the protocol but included this type of study in the review. We stated in the protocol that we would include CBA studies and our search found panel studies which have a similar design to a CBA study (further described in the Methods section).

NOTES

Disclaimer: the findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.