



INSTITUTE
FOR WORK & HEALTH
INSTITUT DE RECHERCHE
SUR LE TRAVAIL ET
LA SANTÉ

The Effectiveness of Occupational Health and Safety Management Systems: A Systematic Review

Full Report

About this report:

Authors:

Lynda Robson¹, Judy Clarke¹, Kimberley Cullen¹, Amber Bielecky¹,
Colette Severin¹, Philip Bigelow¹, Emma Irvin¹, Anthony Culyer^{1,2},
Quenby Mahood¹

Methodological Consultants:

Jill Hayden¹, Selahadin Ibrahim¹, Victoria Pennick¹

Librarian and Library Staff:

Emma Irvin¹, Quenby Mahood¹, Doreen Day¹, Dan Shannon¹

Knowledge Transfer & Exchange:

Jane Brenneman Gibson¹

Administrative Assistant:

Elizabeth Resendes¹, Shanti Raktoc¹

Affiliations:

¹ Institute for Work & Health, Toronto, ON Canada

² University of York, England, UK

If you have questions about this or any other of our reports, please contact us at:

Institute for Work & Health
481 University Avenue, Suite 800
Toronto, Ontario, M5G 2E9

Email: info@iwh.on.ca

Or you can visit our web site at www.iwh.on.ca

Please cite this report as: Robson L, Clarke J, Cullen K, Bielecky A, Severin C, Bigelow P, Irvin E, Culyer A, Mahood Q. The Effectiveness of Occupational Health and Safety Management Systems: A Systematic Review. Toronto: Institute for Work & Health, 2005.

For reprint permission contact the Institute for Work & Health

© Institute for Work & Health, 2005

Table of Contents

1. Overview.....	5
2. Introduction	11
2.1. Aims of the review	11
2.2. Background on systematic reviews	12
2.3. Challenges in conducting a systematic review on OHSMSs.....	12
2.4. Defining OHSMSs	13
2.5. Voluntary and mandatory OHSMS initiatives	16
2.5.1. Voluntary OHSMS initiatives.....	17
2.5.2. Mandatory OHSMS initiatives	19
2.6. Controversy surrounding OHSMSs.....	20
2.7. Existing literature reviews on OHSMSs	22
2.8. Framework for the review of OHSMS effectiveness	24
3. Methods	27
3.1. Literature search	27
3.2. Selection of studies.....	27
3.2.1. Inclusion and exclusion criteria	27
3.3. Quality appraisal (QA)	31
3.3.1. Quality appraisal of quantitative evidence.....	31
3.3.2. Quality appraisal of qualitative evidence.....	32
3.4. Data extraction (DE)	32
3.5. Evidence synthesis	32
4. Results.....	34
4.1. Publications identified through the literature search.....	34
4.2. Quality of the literature investigating OHSMSs	38
4.3. Effectiveness of voluntary OHSMSs	41
4.3.1. Studies of implementation	41
4.3.2. Studies of intermediate outcomes	44
4.3.3. Studies of economic outcomes.....	46
4.3.4. Summary of results in studies of voluntary OHSMSs	48
4.4. Effectiveness of mandatory OHSMSs.....	48
4.4.1. Studies of implementation	51
4.4.2. Studies of intermediate outcomes	54
4.4.3. Studies of final OHS outcomes.....	55
4.4.4. Studies of economic outcomes.....	59
4.4.5. Studies of results in studies of mandatory OHSMs	61

4.5	Evidence synthesis	62
4.5.1	Evidence for the effectiveness of voluntary OHSMS interventions.....	63
4.5.2	Evidence for the effectiveness of mandatory OHSMS interventions	65
4.5.3	Evidence for facilitators and barriers of OHSMS implementation and effectiveness once implemented	67
4.5.4	Evidence for cost-effectiveness of OHSMS interventions	67
5.	Discussion	68
5.1.	Identifying and addressing research gaps.....	68
5.2.	Strengths and limitations of the review	70
5.2.1	Strengths of the review.....	70
5.2.2	Limitations of the review	71
6.	Conclusions and Recommendations.....	73
7.	References.....	75
8.	Appendices.....	85
	List of Appendices.....	87
A.	Details of bibliographic databases	89
B.	Search strategy	92
C.	Quality appraisal forms	94
C.1	Primary quality assessment (QA) form	94
C.2	Supplementary QA form for qualitative evidence	101
D.	Data extraction (DE) form	102
E.	Guide to quality assessment and data extraction.....	105
F.	References for included publications	123
G.	Characteristics of studies considered in this review	125
H.	Detailed summary of evidence.....	127
I.	Summary of inclusion criteria	152

List of Tables

Table 4.2.1	Nature of methodological limitations in the 18 studies selected for quality appraisal.....	40
Table 4.3.1	Summary of Evidence in Voluntary OHSMS Studies.....	42
Table 4.4.1	Summary of Evidence in Mandatory OHSMS Studies.....	52
Table 4.5.1	Summary of effects from studies of voluntary OHSMSs	64
Table 4.5.2	Summary of effects from studies of mandatory OHSMSs	66
Table B.1	Search Terms	92
Table G.1	Study characteristics for studies proceeding to data extraction (n=9).....	125
Table G.2	Study characteristics for studies not-proceeding to data extraction (n=9).....	126
Tables H.1.1	Data Extraction summary tables	
- H.1.9	127
Table H.2	Presence of OHSMS Elements	151
Table I.1	Criteria for Inclusion of Studies.....	152

List of Figures

Figure 2.8.1	Review Framework.....	24
Figure 4.1.1	Flowchart of studies in literature review	34
Figure 4.1.2	Number of studies considering mandatory vs. voluntary OHSMS	35
Figure 4.1.3	Study jurisdiction	35
Figure 4.1.4	Types of outcomes	36
Figure 4.1.5	Study designs	36
Figure 4.1.6	Number of workplaces	37
Figure 4.1.7	Size of workplace.....	37
Figure 4.1.8	Industrial sector.....	38
Figure 5.2.1	Volume of literature retrieved in OHSMS search.....	70
Figure B.1	Search strategy	93
Figure E.1	Study design algorithm	122

1. Overview

1.1 Introduction

Occupational health and safety management systems (OHSMSs) have developed considerably over the last 20 years. There are now more OHSMSs in place and available than ever before. Yet little is known about the effectiveness of these systems on employee health and safety and on relevant economic outcomes.

This systematic review of the literature was undertaken to assess the research evidence on the effectiveness of OHSMSs. Many countries including Canada are in the process of developing management standards for occupational health and safety, so a better understanding of the impact of these systems is timely.

A systematic literature review uses strict, explicit methods to identify, select and critically appraise relevant studies on a certain topic. This review initially set out to investigate three key issues:

- What is the relative effectiveness of mandatory and voluntary OHSMSs on employee health and safety and associated economic outcomes?
- What facilitators and barriers are there to the adoption and the effectiveness of OHSMSs?
- What is the evidence on the cost-effectiveness of OHSMSs?

1.2 What are OHSMSs?

A specific challenge was the lack of consensus on the definition of an OHSMS and how it might be distinguished from other occupational health and safety programs. After reviewing the various definitions found in the literature, the review team devised the following definition:

“An OHSMS is the integrated set of organizational elements involved in the continuous cycle of planning, implementation, evaluation, and continual improvement, directed toward the abatement of occupational hazards in the workplace. Such elements include, but are not limited to, organizations’ OHS-relevant policies, goals and objectives, decision-making structures and practices, technical resources, accountability structures and practices, communication practices, hazard identification practices, training practices, hazard controls, quality assurance practices, evaluation practices, and organizational learning practices.”

OHSMSs are generally distinguished from traditional occupational health and safety programs by being more proactive, better internally integrated and for incorporating stronger elements of evaluation and continuous improvement.

The scope of the review included both mandatory and voluntary OHSMSs.

Voluntary OHSMSs arise through private enterprise, employer groups, government and its agencies, insurance carriers, professional organizations, standards associations and are not directly linked to regulatory requirements. Government-affiliated agencies or insurance carriers sometimes offer incentives to organizations that adopt particular voluntary OHSMSs.

Mandatory OHSMSs arise from government legislation and its enforcement through inspections, fines, etc. In general, mandatory OHSMSs are simpler in terms of what they require of organizations, since they are intended for all or most workplaces, including small workplaces.

1.3 What research was included?

The review team searched seven electronic databases covering a wide range of journals. These contained mainly abstracts of peer-reviewed articles from a variety of disciplines. Reviewers sought relevant studies on OHSMSs, including both implementation and effectiveness research. ***The initial search produced 4807 studies.***

1.3.1 Study relevance The studies were screened for relevance by testing their titles and abstracts against an explicit set of inclusion criteria. For inclusion, reported research had to address at least two of the 27 elements in a comprehensive OHS framework; one of these two had to be a management element. After this initial screening, potentially relevant publications were tested again against the inclusion criteria. ***At this point, 18 studies were considered eligible and were appraised for their methodological quality.***

1.3.2 Quality appraisal, data extraction and evidence synthesis The methodological quality of each study was rated independently by at least two reviewers using a set of explicit criteria. The reviewers then met to reach consensus.

For evidence to be included in the data extraction and evidence synthesis, reviewers had to agree that it met the standard for being of at least “moderate” quality. ***Nine studies reached this minimal quality requirement.***

Of the nine studies, four examined voluntary systems and five evaluated mandatory systems. None of these studies provided evidence of sufficient quality on facilitators and barriers but they did provide information on implementation and effectiveness of OHSMSs.

1.4 Findings

1.4.1 Voluntary OHSMSs All studies involving voluntary OHSMS interventions reported positive findings. While the outcomes measured varied among studies, the findings included increased implementation over time, better safety climates, increased hazard reporting by employees, more organizational action taken on occupational and health issues and decreased workers' compensation premiums. It is likely that the size of the observed declines in premium rates (23 and 52 percent) would be considered important by stakeholders.

However, all the studies in this group were of only "moderate" quality, largely because their study designs were so simple. This meant that reviewers had some doubts regarding the validity of the study findings. Three of the four studies involved single workplaces, making the applicability of the findings to other workplace settings uncertain. There was, moreover, reason to suspect that publication bias might account for the consistency of positive results in these studies.

1.4.2 Mandatory OHSMSs Of the five studies examining mandatory OHSMSs, three involved Norwegian worksites and two evaluated the impact of Canadian provincial regulations introduced in the 1970s. All five studies reported positive findings. Some documented increased OHSMS implementation over time. Others found that OHSMS implementation improved employee perceptions of the physical and of the psychosocial working environment and increased workers' participation in health and safety activities. Implementation also reduced rates of lost time injury in workplaces and led to increases in productivity. It is likely that the size of the observed changes in OHSMS implementation and injury rate would be considered important by stakeholders.

All the studies in this group had moderate methodological limitations. These arose again from the simplicity of the study designs, especially the cross-sectional studies where the direction of causality was not clear.

No studies were found that compared voluntary and mandatory OHSMS interventions directly. No studies of sufficient quality were found that examined facilitators of and barriers to OHSMS implementation or effectiveness, nor were any found that estimated the cost of OHSMS implementation.

1.5 Strengths and limitations of this systematic review

1.5.1 Strengths of the review The studies included in the review were drawn from the research literature of a broad range of disciplines. They confirmed that there has been no other systematic review of the effectiveness of OHSMSs. The questions for the review were framed in collaboration with key stakeholders through formal and informal consultations thereby increasing the relevance of the review results for these stakeholders.

1.5.2 Limitations of the review The large volume of literature from the seven databases meant that reviewers could only carry out a preliminary search and screen of other literature, i.e., that which is not peer-reviewed and published. A search of thesis dissertations and “grey literature” identified government reports and publications which did in fact meet the systematic review’s inclusion criteria. More research would be needed to determine whether these articles are of sufficient quality to contribute to the evidence base on OHSMS.

1.6 Identifying and addressing research gaps

The review identified a number of gaps in the research. The most important was the lack of research whose explicit purpose was to study the effectiveness of voluntary and mandatory OHSMSs on employee health and safety and economic outcomes. The research designs were not sufficiently rigorous to allow for a high degree of confidence in the findings. This paucity of high quality studies may reflect, at least in part, how difficult it is to carry out applied research in workplaces.

1.7 Conclusions and recommendations

The synthesis of the best evidence available showed consistently positive effects in workplaces for voluntary and mandatory OHSMSs. However the absolute number of studies producing these results was not large and their quality was not high. The current applicability of these results to Canada is also questionable.

There is insufficient evidence in the published, peer-reviewed literature on the effectiveness of OHSMSs to make recommendations either in favour of or against OHSMSs. This is not to judge these systems as ineffective or undesirable; it is merely to say that it would be incautious to judge either way in the present state of our research knowledge.

Given the current state of evidence regarding OHSMSs effectiveness:

The review team recommends that those who fund Canadian research should support studies examining the effectiveness or cost-effectiveness of OHSMSs. Support should also be given to research aimed at

identifying facilitators of and barriers to OHSMS implementation and effectiveness.

The generalizability and practical application of this research would be greatly enhanced if stronger research designs were used. This would include the use of comparison groups and longitudinal designs. It is also important for researchers to carefully identify and control for potential confounders, to use larger samples selected through random means, and to include more rigorous economic evaluations. Research using both qualitative and quantitative methodologies would be helpful.

The review team recommends that when the topic of OHSMSs is reviewed in future, researchers seek evidence from sources outside the peer-reviewed, published literature.

This review involved an extensive search for research literature on OHSMS interventions. It focused on the published, peer-reviewed literature in order to concentrate on high-quality studies. However, relatively few studies were found and they were only of “moderate” quality. Others who are interested in doing similar research should consider using additional sources of literature.

Because all mandatory and some voluntary OHSMSs are initiated by the government and its agencies, these organizations are a likely source for evaluative reports on the subject. Thesis dissertations might also provide a valuable pool of high-quality studies, since their production involves peer-review.

2. Introduction

The Workplace Safety and Insurance Board committed funds for a four-year systematic review initiative at the Institute for Work & Health, starting in 2004. The products of these reviews are best-evidence syntheses of research on injury prevention and the topics for review are developed through consultations with stakeholders. This review represents one of the two systematic reviews¹ on injury prevention produced by the Institute for Work & Health in 2004.

2.1 Aims of the review

Occupational health and safety management systems (OHSMSs) have developed considerably over the last 20 years. There are now more OHSMSs in place and available than ever before. Yet little is known about the effectiveness of these systems on employee health and safety and associated economic outcomes.

This systematic review of the literature was undertaken to assess the research evidence on the effectiveness of OHSMSs. Many countries including Canada are in the process of developing management standards for occupational health and safety, so a better understanding of the impact of these systems is timely.

A systematic literature review uses strict, explicit methods to identify, select and critically appraise relevant studies on a certain topic. This review initially set out to investigate three key issues:

- What is the relative effectiveness of mandatory and voluntary OHSMSs on employee health and safety and associated economic outcomes?
- What facilitators and barriers are there to the adoption and the effectiveness of OHSMSs?
- What is the evidence on the cost-effectiveness of OHSMSs?

These questions were developed through formal and informal consultations with representatives of employers, labour, and Ontario's public sector prevention system.

A secondary aim of the review was to characterize the content and methodology of existing research literature on OHSMSs. The purpose here was to identify gaps and weaknesses in the literature, which could help guide future research in this area.

¹ *The other systematic review is about the effectiveness of participatory ergonomics interventions (Cole et al. 2004).*

2.2 Background on systematic reviews

A systematic literature review uses specific explicit, thorough methods to identify, select and critically appraise relevant research studies on a well-defined topic. The basic elements of a systematic review are:

- an extensive, systematic search of research literature for relevant studies
- systematic evaluation of the quality of relevant studies
- systematic synthesis of the findings in the best quality studies

Systematic reviews aid decision-makers by sifting through an enormous literature to find the best quality studies and then synthesize them. Decision-makers usually do not have the time, resources and/or expertise to undertake these tasks themselves.

Systematic reviews are more highly regarded than traditional literature reviews because they employ such rigorous, transparent methods. This means they are less vulnerable to the biases of a single researcher.

As with previous systematic reviews on primary prevention at IWH (Cole et al., 2004; Tompa et al., 2004) this review is a particular type of systematic review known as “best-evidence synthesis” (Slavin, 1995). This approach advocates using explicit, thorough methods to select the best studies from the literature which address a particular research question.

The Slavin (1995) approach includes generating a detailed presentation of individual study characteristics, along with either a qualitative or quantitative synthesis of results². This allows readers of the systematic review to actually judge the quality of individual studies. Slavin also encourages reviewers to critique and interpret the body of extracted literature in the manner of a narrative review.

2.3 Challenges in conducting a systematic review on OHSMSs

There were several challenges in carrying out this review:

- There is no consensus on the definition of an OHSMS. Yet a working definition was needed to support decisions about which studies would be included in the review and which would be excluded.
- The literature in which evidence on OHSMS effectiveness is found is diverse. Diversity increases the time and resources required during

² *Systematic reviews can sometimes yield a quantitative summary of the effect of an intervention, based on the pooling of the results from the review's studies, as in a meta-analysis. This requires there to be several studies of a similar nature. If this condition is not met, then a qualitative summary of the studies is appropriate.*

the initial literature search step. The search strategy needs to be adapted to each database, and any duplication of abstracts between databases eliminated.

- There was an overall scarcity of high-quality evidence about the effectiveness of OHSMSs on employee health and safety and associated economic outcomes.

2.4 Defining OHSMSs

There is no consensus on what an OHSMS is and its scope is potentially wide. Some definitions are simply too vague to be helpful in determining which literature should be included in a systematic review: e.g., the definition used by the International Labour Organization (ILO, 2001):

“A set of interrelated or interacting elements to establish OSH policy and objectives, and to achieve those objectives.”

The definition used in the Australian-New Zealand OHSMS standard AS/NZS 4801:2001³ is more specific, but is still very broad:

“...that part of the overall management system which includes organizational structure, planning activities, responsibilities, practices, procedures and resources for developing, implementing, achieving, reviewing and maintaining the OHS policy and so managing the risks associated with the business of the organization.”

It is not clear from the second definition whether the management system includes only management components or technical/operational components as well. After all, the technical/operational components play a role in “implementing...OHS policy” and “managing the risks.” This problem has also been noted by Nielsen (2000):

“OHSM systems are not, of course, a well-defined set of management systems. Indeed there are not clear boundaries between OHS activities, OHS management, and OHSM systems.”

OHSMSs, as commonly understood, are distinguished from traditional OHS programs by being more proactive, better internally integrated and for incorporating elements of evaluation and continuous improvement. Some OHSMS documents (e.g., ILO, 2001; Chemical Industries Association, 1995; HSE, 1997) explicitly ascribe their basic source as the Plan-Do-Check-Act model of continuous quality improvement that W. Edwards Deming introduced to the management field (Tartorella, 1995). Indeed, many OHSMSs involve the following continuous cycle:

³ Quotation provided by Victoria Workcover Authority website (www.workcover.vic.gov.au)

- Plan (i.e., goals, objectives, standard-setting, accountability, etc.),
- Do (i.e., implementing organizational processes like training and joint-health-and-safety-committee meetings),
- Check (i.e., evaluate through injury statistics reviews, inspections, root-cause analyses, audits, etc.), and
- Act (i.e., based on the evaluation results, make changes to improve the OHSMS and its effectiveness).

In contrast, traditional OHS programs can be characterized as having relatively little in the Check and Act domains. Furthermore, action tends to be reactive in response to workplace accidents, legislation, or enforcement, rather than proactive.

Redinger and Levine (1998) gave detailed consideration to what constitutes an OHSMS. After reviewing 13 publicly available management system documents for occupational health and safety, environment, or quality, they selected four from which to construct a “universal OHSMS instrument.” The four management systems on which this instrument was based are:

- the U.S. Occupational Safety and Health Administration (OSHA)’s Voluntary Protection Program, which was the most comprehensive management system within OSHA
- ISO 14001, the International Organization for Standardization (ISO) environmental management system standard
- BS8800, a voluntary standard from the British Standards Institute, based on both the Health and Safety Executive’s HSG65 model (HSE, 1997) and ISO 14001
- the American Industrial Hygiene Association’s OHSMS (Dyjack, 1998), which was designed to align with ISO 9001, the quality management system standard.

The authors selected these models because each was comprehensive, and together they represented the essential management system elements present in all 13 management system documents.

With the help of an expert team consisting of people from labour, government, industry, academia and professional associations, the text of the four management system documents was deconstructed until each text fragment represented a distinct, simple element of the management system. The fragments were then used to construct the universal OHSMS model with its 27 elements, onto which each of the simpler elements could be mapped.

This work is valuable from at least two points of view. First, their instrument operationally defines the scope of OHSMSs, assuming one can be confident that the 13 management system documents used at the outset

collectively define the possible scope of an OHSMS. Their model can therefore be used to assess the completeness of other OHSMSs, as it was in a review of OHSMSs prepared for the International Labour Organization (Dalrymple et al., 1998). This latter group of authors, which included Redinger and Levine, used the framework to assess the completeness of 23 management system documents targeting occupational health and safety or the environment. The present review also used the same framework to assess the content of the OHSMSs under study (Appendix H2).

The completeness of Redinger and Levine's framework is attested to by the similarity between it and the later ILO Guidelines (2001). The ILO document had input from employer, labour, and government representatives, enabling serious omissions from the background document to the ILO (i.e., Dalrymple et al., 1998), to be rectified.

The second valuable contribution of Redinger and Levine's work is its analysis of the 27 elements from the viewpoint of general systems theory, drawing on von Bertalanffy (1950) and a policy analysis model by Brewer and deLeon (1993).

The 27 elements were grouped into five categories: OHS inputs, OHS process (formulation), OHS process (implementation/ operations), OHS feedback, and open system elements. Drawing on system concepts of self-regulation, integration among system elements, and integration of the OHS sub-system with the rest of the organization and the external environment, Redinger and Levine identified the five primary (and four secondary) OHS feedback and open system elements as those that are distinctly system-like. These elements are a concrete means by which one can distinguish an OHSMS from an OHS program. The five primary elements are⁴:

- Communication System
- Evaluation System
- Continual Improvement
- Integration
- Management Review

This is not, however, to say that an OHS arrangement must incorporate all of these elements in order to be classified as an OHSMS.

Drawing on the literature just described and other insights encountered during the course of the systematic review, the review team proposed the following definition of an OHSMS:

⁴ *The four secondary elements are: Document and Record Management, Auditing and Self-Inspection, Incident Investigation and Root Cause Analysis, Health/Medical Program and Surveillance*

“An OHSMS is the integrated set of organizational elements involved in the continuous cycle of planning, implementation, evaluation, and continual improvement, directed toward the abatement of occupational hazards in the workplace. Such elements include, but are not limited to, organizations’ OHS-relevant policies, goals and objectives, decision-making structures and practices, technical resources, accountability structures and practices, communication practices, hazard identification practices, training practices, hazard controls, quality assurance practices, evaluation practices, and organizational learning practices.”

2.5 Voluntary and mandatory OHSMS initiatives

Voluntary OHSMSs are developed and disseminated through a variety of mechanisms. Expert organizations develop and distribute information on standards and best-practice guidelines. Other organizations, both non-profit and for-profit, provide technical expertise or a certifying mechanism to organizations wanting to meet one of these standards. Government-affiliated agencies or insurance carriers sometimes offer incentives to organizations that adopt OHSMSs.

Mandatory OHSMSs arise from government legislation and its enforcement through inspections, fines, etc. Voluntary OHSMSs arise through private enterprise, employer groups, government and its agencies, insurance carriers, professional organizations, standards associations and are not directly linked to regulatory requirements.

The following section gives a brief overview of some of the better known voluntary and mandatory initiatives.

In general, the voluntary OHSMSs, especially those marketed through commercial industries, target large companies. They are characterized by being more thoroughly specified but are too complex for the majority of employers (Frick and Wren, 2000). Voluntary OHSMS schemes marketed through public agencies are targeted not only on large companies but also on smaller companies (Frick and Wren, 2000). These schemes accordingly either involve simpler OHSMSs or have a menu of options, including simple ones, for companies of different sizes or at different stages of OHSMS development.

In general, mandatory OHSMSs are simpler in terms of what they require of organizations, since they are intended for all or most workplaces, including small workplaces.

2.5.1 Voluntary OHSMS initiatives The Dupont Corporation has long been regarded as a leader in occupational health and safety, which led to the development of a Dupont OHSMS and consultancy (Wokutch and VanSandt, 2000). Other well-known proprietary systems include the International Safety Rating System (ISRS), the Five Star system, the chemical industry's Responsible Care system, the Oil Industry International Exploration and Production Forum (E & P Forum) system, and the American Industrial Hygiene Association's OHSMS.

The International Organisation for Standardization (ISO) considered developing an international management standard for OHS, similar to those already established for quality (ISO 9001) and the environment (ISO 14001). Support for this development among member organizations was insufficient, however, and the project was disbanded in 2001 (Bennett, 2002).

Companies have nevertheless sought certification to an ISO-compatible occupational health and safety standard. Indeed, the British Standards Institute developed OHSAS 18001 in response to this demand (Abad et al., 2002). This standard is internationally recognized and has been adopted by industry as a proxy for an ISO standard. In some industries, there are even pressures from trading partners to adopt the OHSAS 18001 standard, as there have been previously for ISO 9001 and ISO 14001. In addition, some companies see potential efficiencies by adopting an OHSMS that can be integrated with their existing ISO-based management systems for quality and the environment (Winder, 1997; Wright, 2000).

The international OHS standard project was referred by ISO to the International Labour Organization, which was thought to be a more appropriate forum for it. Following this referral, ILO developed Guidelines on Occupational Health and Safety Management Systems in 2001 (ILO, 2001) through a consensus process that included equal representation from government, labour and employers. The ILO guidelines were envisioned as models for national standards.

Dalrymple et al. (1998) found that national voluntary standards for OHSMSs in draft or final form existed in Australia/New Zealand (AS/NZS 4804), Ireland, Jamaica, the Netherlands, Norway, Spain, and the United Kingdom. The forerunner of the British Standards Institute's OHSMSs (BS 8800, OHSAS 18001) was developed by the public sector Health and Safety Executive (HSG65; HSE 1997). The American National Standards Association (ANSI) and the Canadian Standards Association (CSA) are both in the process of developing a national standard. The Canadian organization has included the ILO Guidelines, OHSAS 18001, and the draft ANSI standard as reference documents.

Some of the better known voluntary OHSMS programs in the United States, Australia and Canada are reviewed below.

OSHA's Voluntary Protection Plan in the United States The Voluntary Protection Plan (VPP) in the United States grew from an experimental program begun in California in 1979, and was formally announced by OSHA in 1982. Its aim is to promote effective worksite-based safety and health, with cooperative relationships between management, labour and OSHA. Workplaces can qualify for one of three programs: Star, Merit, or Star Demonstration (the latter recognizing worksites that address unique safety and health issues). To do so they must have implemented a comprehensive safety and health management system, and meet a set of performance-based criteria for a managed safety and health system after on-site evaluation by a team of OSHA safety and health experts. Both federal and state VPPs exist; the number of organizations holding VPP accreditation has grown from 11 in 1982 to 1189 in 2004. The size of those organizations ranges from fewer than 100 workers (n=246) to more than 4,000 (n=15).

The OSHA website (www.osha.gov/dcsp/vpp) states that the average VPP worksite has a lost workday incidence rate 52 per cent below average for its industry. On the other hand, a report by the United States General Accounting Office (2004) notes that "OSHA's voluntary compliance programs appear to have yielded many positive outcomes, but the agency does not yet have adequate data to assess their individual and relative effectiveness."

WorkCover's promotion of OHSMS in Australia The Australian government has been active in the past decade in creating what Saksvik and Quinlan (2003) call a "hybrid mixture of regulatory mandate and incentives to promote the 'voluntary' adoption of OHSMSs by employers." The WorkCover Authority (the workplace insurer) in the Australian Capital Territory describes their approach as a continuum from co-regulatory to enforcement strategies, ranging from recognition of best practice, through technical and systems development advice, provision of guidance materials, legislative advice, to the issuance of notices and prosecution of those who refuse to comply (www.workcover.act.gov.au/about/index.html).

The various state and federal governments have developed audit systems which have incorporated the essential elements of OHSMSs into their criteria; some of these being the WorkSafe Plan in Western Australia, TriSafe Management Systems Audit in Queensland, the Safety Achiever Business System in South Australia, SafetyMAP in the State of Victoria, (Gallagher et al., 2003). However, a recent inquiry to the Victorian WorkCover Authority about the number of companies undergoing SafetyMap audits, and their premium rates in comparison to non-audited firms brought the response that the relatively small number of companies

having SafetyMAP accreditation made release of comparisons inappropriate due to privacy laws (Graham, 2004).

Nova Scotia's Certificate of Recognition The Workers' Compensation Board of Nova Scotia encourages all workplace parties to implement workplace safety systems, and awards their Certificate of Recognition to those who do so, and who meet the established criteria in a safety audit carried out by one of the WCB's approved audit firms. The Government of Nova Scotia may require that potential suppliers hold the Certificate of Recognition in order to bid on contracts.

Alberta's Partnerships in Health and Safety Programs This plan was first established in 1989 in the Province of Alberta; its vision is "Fostering a culture where health and safety becomes an integral part of every workplace." (www3.gov.ab.ca/hre/whs/partners/index/asp). Similar to the aforementioned plans, it is based on the premise that when employers and workers voluntarily build effective health and safety programs, the human and financial costs of workplace injuries and illnesses are reduced, and that larger reductions in injuries are achieved than by regulatory compliance alone. A total of 67 organizations are listed as part of the Partners in Health and Safety network which promotes and certifies organizations. The website lists approximately 4900 organizations which have achieved the Alberta Human Resources and Employment's Certificate of Recognition by meeting the program's health and safety management system. The WCB-Alberta offers premium rate discounts for implementing effective health and safety and disability management programs.

2.5.2 Mandatory OHSMS initiatives It is difficult to pinpoint the start of OHSMS legislation. During the 1970s, legislative reform took place throughout the industrialized world (Frick and Wren, 2000). Some of this legislation could be viewed as constituting a primitive OHSMS intervention, as a review of Ontario's Bill 70 and Quebec's LSST in Appendix H2 suggests.

In the early 1980s in Norway, the offshore industry became regulated by Internal Control legislation, which had its roots in OHSMSs developed by the petroleum industry. Internal Control is likely the most comprehensive of existing mandatory OHSMSs. In 1992, coverage was extended to all of the Norwegian onshore industry.

At a similar point in time, the European Community's Framework Directive (89/391) was promulgated, with a requirement for EC countries to embody it in their national legislation within four years. In Britain, the transposed legislation became the Management of Health and Safety at Work Regulations (Grayham and Rosario, 1997). European researchers refer to the EC Directive as "systematic occupational health and safety

management,” in order to distinguish it from the more complex voluntary systems. Dalrymple et al. (1998) included mandatory OHSMSs from two non-European countries in their review -- India (the 1988 revision of The Factories Act) and Korea (1990 Industrial Safety and Health Act) -- but considered the Korean system to be weak. Brazil was also included in the review, but was judged to not truly be a management system.

The Workwell audit in Ontario is an example of a special type of mandatory OHSMS intervention that does not arise directly from legislation. Instead, the Workwell audits are administered through a legislated agency, the Workplace Safety and Insurance Board. They are required of workplaces with poor performance in terms of workers' compensation claims or legislative compliance (Bennett, 2002). Failing an audit results in relatively severe penalties and requires remedial action. Several hundred companies go through the Workwell process on an annual basis (Richard Burton, WSIB, personal communication).

2.6 Controversy surrounding OHSMSs

There are several areas of controversy regarding OHSMSs. For instance, there is concern that OHSMS strategies, which foster self-regulation of OHS by workplaces, will weaken the external regulatory approaches developed to date (Bennett, 2002). Indeed, there is evidence that authorities view OHSMS strategies as a means of saving on the costs of enforcement (Frick and Wren, 2000, p. 40).

In addition, Quinlan and Mayhew (2000) argue that the current labour market trends towards precarious employment, outsourcing and subcontracting suggest that mandatory OHSMS strategies will not affect substantial portions of the population. The corresponding growth of smaller organizations means that workplaces will be more difficult to reach through either voluntary or mandatory initiatives. The authors also expect that the ability and motivation for employers to undertake OHSMS innovations will be weakened, due to the complexity of modern organizations and the lowered degree of responsibility they have towards some workers. Finally, they predict that as union membership and leverage decreases, so will their influence over OHSMSs and OHS in general.

There is significant concern expressed by labour representatives about the tendency for some of the proprietary systems, in particular, the Dupont and the Five Star systems, to use worker behaviour-based safety techniques. These can foster an authoritarian, “blame the worker” approach to safety (Nichols and Tucker, 2000; Wokutch and VanSandt, 2000). Also of concern to labour is that such systems can undermine the formal means of worker participation in the workplace, by shifting the control of OHS issues more towards management and by failing to integrate with existing union or joint health-and-safety committee arrangements (Lund, 2004; Nichols and Tucker, 2000). These tendencies are considered to be consistent with the

organizational theories underlying OHSMSs, according to Nielsen's (2000) analysis of OHSMSs. He finds OHSMSs reflective of the "rational" models of organizations, which include Tayloristic models, classic bureaucratic models, and particular general systems models wherein the social sub-system component is circumscribed.

Reports on worker participation within Norwegian Internal Control have been more favourable (Lindøe and Hansen, 2000). An Internal Control Committee found that IC improved the cooperative climate and gave employees greater influence over procedures. Lindøe's doctoral thesis based on three case studies reported that introducing IC had a positive impact on the status and roles of workers council, on safety delegates, as well as on the safety service.

The analysis by Nielsen (2000) also posits that the effectiveness of OHSMSs might be enhanced by the application of concepts found in more modern, "non-rational" theories of organizations, involving concepts of human relations, chaos, and the political nature of organizations. He implies that OHSMSs might benefit from a more complete rendering of a general systems model, as in socio-technical approaches. These aim for the joint optimization of the technical and social sub-systems of an organization. However, the elements in the Redinger and Levine (1998) framework demonstrate a primary focus on the technical sub-system and a secondary one on the social sub-system. Yet, certain workplace disasters have been attributed to an inadequate safety culture and safety climate (Kennedy and Kirwan, 1998, citing ACSNI, 1993).

Bennett (2002) identified further concerns with voluntary OHSMSs, particularly those developed outside public agencies. First, legal compliance is sometimes not specified in some OHSMSs, and this could take the focus off an important safeguard for OHS. Second, some systems focus on risk instead of hazards, the latter focus being preferred by labour (lest some hazards be deemed tolerable risks). Third, some systems are based on performance standards, leaving a wide range of discretion to the user, in contrast to specification standards, which do not. Fourth, while all formal OHSMSs require audit, some are quite vague about what that should involve.

The importance attached to audits was explored by Gallagher (2003), who found that experts in the safety field were concerned that the mere presence of an OHSMS, and especially of audit results, could give a false sense of security, and the picture they provide is distorted. She presented some of the findings of the Royal Commission investigating a fatal explosion at an Esso gas plant, as reported by Hopkins (2000). The Commission concluded that Esso's Operational Integrity Management System (OIMS), which was regarded as an exemplary OHSMS, tended to:

“...take on a life of its own, divorced from the operations in the field. In some respects concentration upon the development and maintenance of the system diverted attention from what was actually happening in the practical functioning of the plant.”

The issue of performance measurement has also been taken up by others (Frick and Wren, 2000), with criticism being directed at the tendency of some OHSMSs to focus excessively on lost-time injury statistics or behavioural measures. Regarding the latter, the concern has been that the behaviour of workers is overemphasized. In the case of lost-time injuries, the concern has been that claims management rather than prevention becomes a workplace's dominant focus. There may even be pressures to suppress injury reporting. Furthermore, with so much attention being directed to lost-time injuries, physical safety tends to overshadow other health issues, though OHSMSs are often ostensibly aimed at both.

2.7 Existing literature reviews on OHSMS

Two recent books (Frick et al., 2000; Walters, 2002) contain chapters by international groups of authors, who describe the implementation of OHSMSs in their respective jurisdictions and some of the major forces which have contributed to/affected this movement. The edition by Walters (2002) focuses on a mandatory OHSMS, the EU Framework Directive 89/391, and describes its implementation in seven different EU settings. The various chapters provide a detailed description of trends and influences in the national contexts in which the Framework was implemented. This includes countries in which the Framework Directive caused little fundamental change, being already consistent with existing workplace culture and legislation (Sweden, Denmark, the Netherlands), those in which the Directive was initially implemented in a “minimalist” fashion, but which has gradually been extended or has increased in interest (France, the UK), and those in which there was a more fundamental changes in political interest and/or regulatory structure (Italy, Germany).

Frick et al. (2000) contains chapters on the implementation and effects of mandatory OHSMS legislation in the EU, but this is not its primary focus. The book explores the roots of the movement in a more global perspective (including the EU, the US and Ontario contexts, Japan, Brazil) and the content includes political and strategic movements for promoting OHSMS, their applicability in changing labour markets and business structures, the promotion of health and safety in small workplaces, worker participation, and the integration of OHSMS into general business and management development. The overall tenor of the book is critical of OHSMSs, particularly the voluntary type, although a few positive cases are presented.

Saksvik and Quinlan (2003) compare the experiences of Norway and Australia in implementing OHSMS. They highlight the inter-dependence of occupational health and safety and industrial relations regulatory regimes.

In Norway, the major influences cited are the move to Total Quality Management related to business and environmental control, as well as a democratic tradition which emphasizes participation and cooperation of all parties. In Australia, the parallel environmental legislation and systems framework of international standards is seen as influential; in this context, the introduction of OHSMS is described as a “hybrid mixture of regulatory mandate and incentives to promote the ‘voluntary’ adoption of OHSM systems by employers”.

Finally, a review by Gallagher et al. (2003) cites several of the studies which are included in this systematic review, as well as Gallagher’s own primary research leading to her doctoral dissertation (Gallagher, 2000), and an exploratory study by the same authors, which was done for Australia’s National Occupational Health and Safety Commission (2001). They note that expert consultations turned up doubts and uncertainties as to the effectiveness of OHSMSs, and that research on their effectiveness tends to be inconclusive. They cite international research endorsing the value of the individual components of OHSMS, however, and note that the findings which recur in these studies were the critical role of senior management, and the importance of worker participation.

The article cites Gallagher’s dissertation research, which developed a taxonomy of different kinds of system (based on OSH control strategy and on the type of management structure and style). The former were categorized as “safe person” (the focus being on employee behaviour), and “safe place” (focusing on hazard control); the latter as “traditional” (OSH marginal, poorly integrated into strategy and operations, and involving top-down decisions), and “innovative” (driven by senior management, integrated into business planning and production, and involving teamwork and employee participation). Of 20 companies which Gallagher had studied, the three categorized as having “innovative management” and “safe place” strategies achieved above average performance on the SafetyMAP “output” data. They also had declining trends in injury data and against industry benchmarks. In contrast, those companies with “traditional management” and “safe person strategies” did less well.

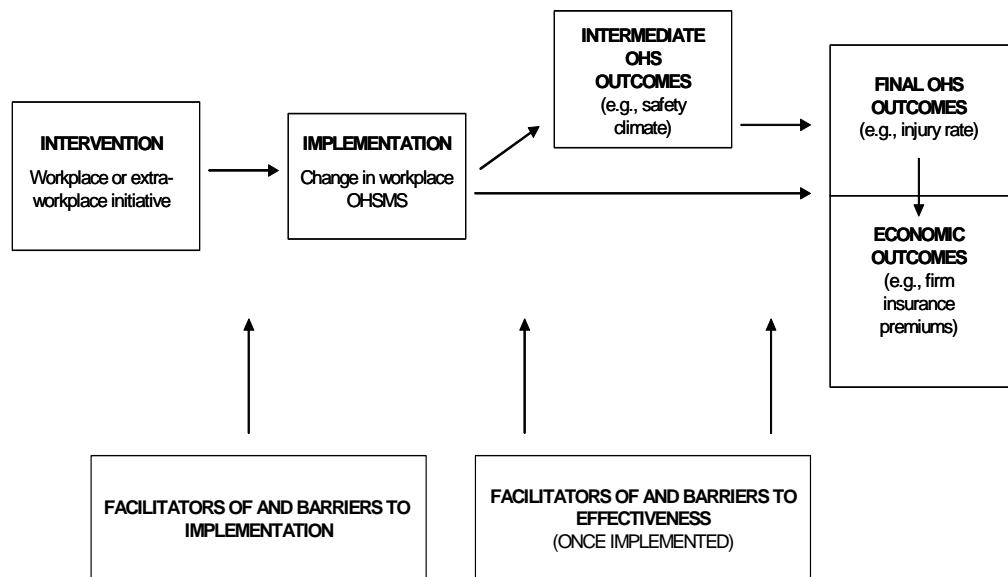
The “experts” who were consulted as part of Gallagher’s research emphasized some conditions they felt necessary for OHSMSs to succeed: OHSMSs customized to organizational needs; developed with stakeholder input; senior executives committed to OSH performance, willing to commit resources and make line managers accountable, and to lead by example; integration of OHSMS operations and other organizational functions; encouragement for employee participation and independent employee representation. The same experts cited barriers: failure to meet the above-mentioned conditions, especially those involving manager commitment and employee involvement; inappropriate use of audit tools; specific contextual

barriers, e.g. areas in which OHSMS implementation is especially difficult, such as small business, part-time or temporary employment, or contractors. The overall conclusion by Gallagher et al. (2003) is that although the evidence is “suggestive rather than conclusive”, it would seem that OHSMSs can work, if/when a set of very demanding conditions are met.

2.8 Framework for the review of OHSMS effectiveness

This section describes the major concepts applied to the studies involved in the current systematic review. For inclusion in the review, a study had to describe an OHSMS intervention and then do at least one of the following: i) provide change measures of OHSMS implementation; ii) estimate the effect of the intervention on intermediate/final OHS outcomes or economic outcomes; or iii) study facilitators of or barriers to OHSMS implementation or effectiveness. These concepts will be more fully explained below. The key relationships among these concepts are indicated in Figure 2.8.1.

Figure 2.8.1: Review framework



OHSMS intervention The scope of the review included interventions directed at developing the OHSMS in one or more workplaces. It therefore included studies of extra-workplace initiatives such as legislation, or voluntary programs arising through the government, its agencies, insurance carriers, groups of employers, etc. It also included studies of workplace-level initiatives, through which a workplace might attempt to improve its OHSMS, using either a scheme developed externally (e.g., British Standards Institute’s OHSAS 18001) or one developed internally. A minimalist definition of an OHSMS intervention was adopted. In order to count as

such, an intervention was required to address two or more of the 27 elements in the Redinger and Levine (1998) universal OHSMS framework, with at least one of these being a management element, as distinct from an operational/activity element.

OHSMS implementation While the primary focus of the review was on the effectiveness of OHSMSs, evidence was also sought about implementation at the workplace-level . There were two reasons for this. First, it was anticipated that for some mandatory initiatives there might be measures only of the OHSMS and not of its effects in workplaces. Second, implementation information allows one to distinguish between the following two possible explanations for an absence of effect: poor intervention content and poor implementation of the intervention. It is evidently possible for a well-conceived intervention to fail if it is poorly implemented.

An evaluation of extra-workplace OHSMS initiatives could measure implementation by measuring changes in structures and processes external to the workplace. In this review, however, implementation was considered only at the workplace-level, (as a change in the state of the workplaces' OHSMSs), in order to have consistency between the workplace-initiated and extra-workplace-initiated interventions included in the review. Examples of implementation changes include an increase in the number of OHSMS elements present in the workplace or improved quality of OHSMS elements.

Intermediate OHS outcomes Intermediate OHS outcomes were considered to be outcomes of secondary interest and potential proxies for final OHS outcomes. These would involve changes in constructs mediating between the OHSMS and final OHS outcomes. Examples would be: safety climate; employee knowledge, beliefs, values or perceptions; employee behaviours; OHS hazards; and risks of illness or injury.

Final OHS outcomes The review team identified the final outcomes of the interventions using a program evaluation approach (Rush and Ogborne, 1991). Outcomes were chosen so as to be consistent with the ultimate purpose of the intervention. For many stakeholders, this is improved employee health and safety. Thus, examples of final outcomes are changes in injury/illness statistics, musculoskeletal pain, and employee quality of life.

Economic outcomes Economic outcomes could also be considered to be final outcomes in the review framework, especially for stakeholders who had a primary interest in costs. A conventional notion of economic outcomes was adopted and so examples in this category are workplace workers' compensation premium rates and workplace productivity.

Facilitators of and barriers to OHSMS implementation and/or effectiveness It was expected that facilitators of and barriers to OHSMS implementation and effectiveness would fall into three categories: those internal to the OHSMS (e.g., management commitment to OHS, performance indicators, worker participation); those external to the OHSMS but in the workplace (e.g., company size, presence of other management systems, industrial relations, unionization); and those external to the workplace (e.g., trade pressures).

3. Methods

This section describes the methodological steps used in the review: 1) the search of the literature; 2) the selection of relevant studies through application of inclusion and exclusion criteria; 3) quality appraisal of relevant studies; 4) data extraction from higher quality studies; and 5) synthesis of the evidence.

3.1 Literature search

Seven electronic databases, abstracting primarily peer-reviewed research journal articles, were searched from their inception until July 2004: MEDLINE (from 1966), EMBASE (from 1980), PsycInfo (from 1887), Sociological Abstracts (from 1963), Safety Science and Risk Abstracts (SSRA, from 1981), EconLit (from 1969) and American Business Inform (ABI, from 1918).⁵ See Appendix A for details on the range of topics covered by these bibliographic databases. Since the search terms and language of the databases were found to differ significantly, the terms used in the search were customized for each database. A list of the broad terms used in the search can be found in Appendix B, Table B.1.

The search strategy combined two sets of keywords using an "AND" strategy (Appendix B, Figure B.1). The first set of keywords focused on OHSMS terms, while the keywords in the second set included evaluation and OHS effectiveness terms. The terms within each group were combined using an "OR" strategy. The titles, abstracts, case registry, and MeSH subject headings were searched for keywords. The search strategy was not limited in terms of language. In addition, the reference lists of all papers meeting the eligibility criteria and review papers on this topic were manually checked for relevant studies.

3.2 Selection of studies

Titles and abstracts of each article were screened by at least two reviewers. Full text articles were retrieved for those studies that appeared to meet the inclusion criteria (Appendix I), and for those in which the information presented in the title, abstract, and key words was insufficient to exclude them. A consensus method was used to resolve any disagreements between the two reviewers on study inclusion. A third reviewer was consulted for those studies in which agreement could not be reached.

3.2.1 Inclusion and exclusion criteria The review team needed some way to recognize an OHSMS intervention in the literature. This was a challenge at the outset of the study in the absence of a definition of an OHSMS.

⁵ Searches of CCIInfoWeb (Canadian Centre for Occupational Health and Safety (CCOHS), from 1900) and Dissertation Abstracts International (DAI, from 1861) were also conducted, but disbanded after the title and abstract screening step because of the large yield found in the seven databases.

One strategy would have been to review only the literature that referred to some variant of the term “occupational health and safety management system.” This approach seemed too narrow, since it would have potentially excluded relevant OHS interventions that happened to use different terminology, such as “Internal Control”, “systematic occupational health and safety management” or even “comprehensive safety program.”

Consideration was given to requiring an OHSMS intervention to address at least one element from each of the five conceptual categories of the Redinger and Levine (1998) framework (discussed in section 2.4), but this criterion was not adopted for fear that this too might exclude too many OHSMS interventions of interest. The review’s scope included mandatory OHSMS interventions, which are usually targeted to the majority of workplaces, and are thus less elaborate than some voluntary systems.

A minimalist definition of an OHSMS *intervention* was adopted. An intervention was required to address two or more of the 27 elements in the Redinger and Levine (1998) framework, with at least one of these being a management element, as opposed to an operational/activity element.⁶

The inclusion criteria also required that the studies involve an intervention, for which the OHSMS elements were described either explicitly or through reference to a known OHSMS. The criterion was intended to exclude studies that investigated the relationship between a researcher-defined measure of OHSMSs and outcomes out of the context of an intervention. In doing so, the review focussed on findings that had the most direct implications for decision-makers contemplating OHSMS interventions.

It was expected that facilitators of and barriers to OHSMS implementation and effectiveness would encompass factors internal to the OHSMS (e.g., management commitment to OHS, performance indicators, worker participation), factors external to the OHSMS but in the workplace (e.g., company size, presence of other management systems), and factors external to the workplace (e.g., trade pressures). Given this expectation, it would seem at first glance that almost any study looking at the effect of a single OHSMS element on OHS performance might have to be included. As a means of narrowing this unfeasible scope, publications looking at facilitators of or barriers to effectiveness were required to specify a level of OHSMS implementation. It was intended that this criterion would limit the retrieved studies to those where the reader would have an understanding of the implementation context, and thus, its applicability elsewhere.

⁶ *It should be noted that the requirement for at least two elements of the management system intentionally excluded the body of literature that has accumulated on single elements of OHSMSs (e.g., LaMontagne et al., 1996; Shannon et al., 1996; Habeck et al., 1998; Cohen, 1977; Simard and Marchand, 1994; Reilly et al., 1995), since the review’s focus was systems.*

The full set of inclusion and exclusion criteria used to determine the eligibility of studies for this review are described below and summarized in tabular form in Appendix I.

Publication type Studies published in peer-reviewed journals, book chapters, dissertations, reports, or Internet publications that contained a report on relevant research or a review of relevant research were included.⁷ Magazine articles, newspapers, newsletters, and conference proceedings were excluded.

Population of interest Workplaces could be located anywhere in the world.

Nature of intervention An OHSMS intervention, which was initiated at either the workplace or extra-workplace level (i.e., an initiative instituted by a government body or agency, or other extra-workplace organization, to encourage the development of OHSMSs in workplaces) was required in the study.

Extra-workplace or workplace initiatives designed to address isolated aspects of OHSMSs or particular risks (e.g., needle-stick injuries in a health-care facility) were not considered to be OHSMS interventions.

An OHSMS intervention could be identified by one of three means:

- i) directly, by a term synonymous with OHSMS or mention of specific types of OHSMS (e.g., ‘safety and health management system’, OHSAS 18001);
- ii) indirectly, by mention of OHSMS legislation or other extra-workplace OHSMS initiative (e.g., European Framework Directive 89/391, Internal Control); or
- iii) indirectly, by a term suggestive of OHSMS, and a description of its components that demonstrates that it is an OHSMS (e.g., comprehensive occupational health and safety program). In the case where terms were merely suggestive of an OHSMS, a description of the OHSMS must have been reported or referenced, and the description would qualify as that of an OHSMS if two or more system elements were specified, one of which was in the management domain (e.g., leadership, policy, planning, structure, evaluation), as opposed to both lying in the activity/technical domain (e.g., training, hazard control).

Multi-faceted management system interventions were included if they had an occupational health and/or safety component (e.g., a safety, health and environmental management system). However, the OHSMS intervention

⁷ Although a variety of publication types were permitted, the seven databases abstracted primarily peer-reviewed, published literature.

must have included primary prevention as a major component. Thus, management system interventions focusing on disability or health services were excluded.

Type(s) of evidence Studies which examined either i) OHSMS implementation, ii) effectiveness of OHSMS interventions, or iii) facilitators/barriers to OHSMS implementation or optimal outcomes following OHSMS implementation were included in this review.

Outcomes

- **Implementation studies** To meet the review's inclusion criteria, implementation studies were required to have a quantitative measure of change in the level or intensity of the OHSMS. Measures of OHSMS implementation typically assess the presence/absence/quality of OHSMS elements in workplaces (e.g., management audit). Implementation studies that did not report on a corresponding extra-workplace or workplace intervention were excluded.
- **Effectiveness studies** To meet the review's inclusion criteria, effectiveness studies were required to have a quantitative⁸ measure of one of the following outcomes:
 - i. *Intermediate OHS outcomes* such as, changes in knowledge, beliefs, values, perceptions, behaviours, hazards, or risks
 - ii. *Final OHS outcomes* such as, changes in injury/illness statistics or employee quality of life
 - iii. *Economic outcomes* such as, changes in the costs associated with employee illness/injury (at either the workplace or extra-workplace level).

Comparisons in outcome studies To be included in this review, implementation and effectiveness studies were required to make a *comparison* of outcomes with the presence and absence of an OHSMS intervention, or between OHSMS interventions of different intensities. These comparisons could be made within or across workplaces.

Facilitator/barrier studies To meet the review's inclusion criteria, quantitative or qualitative research methods were required to identify facilitators or barriers. Facilitators/barriers could relate to either OHSMS implementation or optimal outcomes following OHSMS implementation. The facilitators/barriers could relate to i) the OHSMS itself, ii) the workplace, or iii) the environment external to the workplace. Studies examining facilitators/barriers of optimal outcomes following OHSMS implementation that did not include a specification of the level of OHSMS implementation were excluded. Additionally, facilitator/barrier studies reflecting solely one expert's opinion were also excluded.

⁸ *For implementation and effectiveness evidence, we sought only quantitative findings, as the review's goal was to derive a quantitative estimate of effect.*

3.3 Quality appraisal (QA)

Studies which met the eligibility criteria were assessed for methodological quality using a process developed by the authors based generally on previous work (Franche et al., 2004; Cote, 2002; Drummond and Jefferson, 1996; Jadad, 1998; Oxman and Guyatt, 1991; Smith et al., 2000; Zaza et al., 2000; Kuhn et al., 1999; van Tulder et al., 2003; van Tulder et al., 1997; Abenhaim et al., 2000; Thomas et al., 2003; Tompa et al., 2004; Sculpher et al., 2000).

Although many of the reviewed quality assessment tools incorporated about a dozen questions, the tool developed for this review emphasized parsimony with an aim to streamlining the consensus procedure. This meant that issues which were often covered in more than one question, are covered in only one (for example all issues related to selecting and maintaining the sample were covered in one question). As well, the tool was designed to focus on internal validity.

The quality appraisal (QA) form and guide to reviewers can be found in Appendices C.1 (Primary QA form) and E (Guide to quality assessment and data extraction). The methodological quality of each study was rated independently by two reviewers, who then met for consensus. If consensus could not be reached, one or more other reviewers were consulted as needed.

The QA questions were structured around the type of evidence(s) examined by a study. As noted in section 3.2, there were three types of evidence included in this systematic review: (1) implementation, (2) effectiveness (on intermediate outcomes, final OHS outcomes, and/or economic outcomes), and (3) facilitators and/or barriers (using either quantitative or qualitative methods). It is important to note that the review's quality appraisal of the evidence refers to the quality of the research studies and reports, and not to the quality of the OHSMS interventions themselves. Furthermore, the quality was appraised from the point of view of the review's questions about OHSMSs. A study may in some cases have been assigned a higher quality if the research question had been different.

3.3.1 Quality appraisal of quantitative evidence For all types of quantitative evidence, reviewers were asked to rate the studies on i) selection and maintenance of the sample and ii) potential confounders. The ratings were based on a three-point scale (meets criteria - Yes, Partially, No). There were three additional quality criteria that were considered separately for each type of quantitative evidence: measurement methods, appropriateness of statistical analyses, and other issues (including contamination of comparison group with exposure to the intervention). The quality ratings for these criteria were assessed separately for each type of evidence.

An overall summary rating on quality for each type of quantitative evidence was given, using a four-point scale ranging from “Very Low” to “High”. This summary rating took all five quality dimensions into consideration.

Studies containing quantitative evidence rated as “Moderate” (moderate limitations) or “High” (no or minor limitations) in overall quality proceeded to data extraction. Studies containing evidence rated as “Low” (major limitations) or “Very Low” (serious limitations) were excluded from evidence synthesis after QA. In certain cases, where studies reported more than one type of evidence, some of the evidence may have proceeded to data extraction, while the remainder did not.

3.3.2 Quality appraisal of qualitative evidence Studies containing qualitative evidence on facilitators/barriers were assessed for methodological quality using an adaptation of a process developed by the National Centre for Social Research in the United Kingdom (Spencer et al., 2003; Appendix C.2). Two researchers independently conducted the quality assessment, and met for consensus.

To be consistent with accepted qualitative research methods, no fixed formula was used to determine inclusion of the qualitative studies. The major emphasis was, however, put on question 1, involving the overall credibility of the study. This judgment was carefully made on the basis of answers to the other 16 questions. Reviewers met to reach consensus on an overall rating of the study methodology as “High”, “Moderate”, “Low” or “Very Low”, with consideration of the “fit” of the study methods to the study’s purpose, whether the context of the study was considered where appropriate, the adequacy of the sample and/or analysis, and the explanatory value of the study. As with the quantitative evidence, those studies judged to be of Low or Very Low quality were not kept for final data extraction.

3.4 Data extraction (DE)

Pairs of reviewers independently extracted data from the included studies, and then met to reach consensus. A copy of the data extraction (DE) form and the guidance given to reviewers can be found in Appendices D (Data Extraction (DE) Form) and E (Guide to Quality Assessment and Data Extraction). Data were extracted on the type of OHSMS (mandatory versus voluntary), study design, research question, study population characteristics, sampling strategy, participation rates, outcomes of interest to the review, statistical analyses, statistical power, and results. They also recorded the presence or absence of individual OHSMS elements, according to the framework presented by Redinger and Levine (1998).

3.5 Evidence synthesis

Many systematic reviewers choose an explicit algorithm at the outset of the study for later translation of the findings into a summary statement about the level or strength of evidence they provide. (van Tulder et al., 2003; Franche

et al., 2004; Kuhnet al.,1999; Briss et al., 2000; Tompa et al., 2004). For example, application of the algorithm developed by Briss et al. (2000) would characterize the level of evidence provided by a group of studies as “Insufficient”, “Sufficient”, or “Strong”.

Criteria for these algorithms concern the study design, quality of research (as determined using the review’s quality assessment tool), the consistency of the results, and the quantity of research. The GRADE Working Group (2004) also includes consideration of the direct applicability of the studies to a new setting of interest, the strength of association, whether a dose-response gradient was seen, whether all plausible confounders would have reduced effect, and whether there was a reporting bias⁹.

In contrast, this review did not adopt an explicit algorithm at the outset. The reason was a lack of consensus in the OHS prevention field as to which synthesis algorithm was best. In addition, it was thought premature to base an algorithm upon so newly developed a QA tool. This review instead synthesized a summary statement in the style of a traditional narrative review, which is customary for systematic reviews in this field (Am J Prev Med (2000)) and permissible in best-evidence syntheses (Slavin, 1995).

⁹ *Reporting bias is also known as publication bias, which is the bias towards more positive results (i.e., in the direction intended by the intervention) within a body of literature, resulting from the reluctance of researchers to write and submit manuscripts with null finding and the reluctance of editors to publish such manuscripts.*

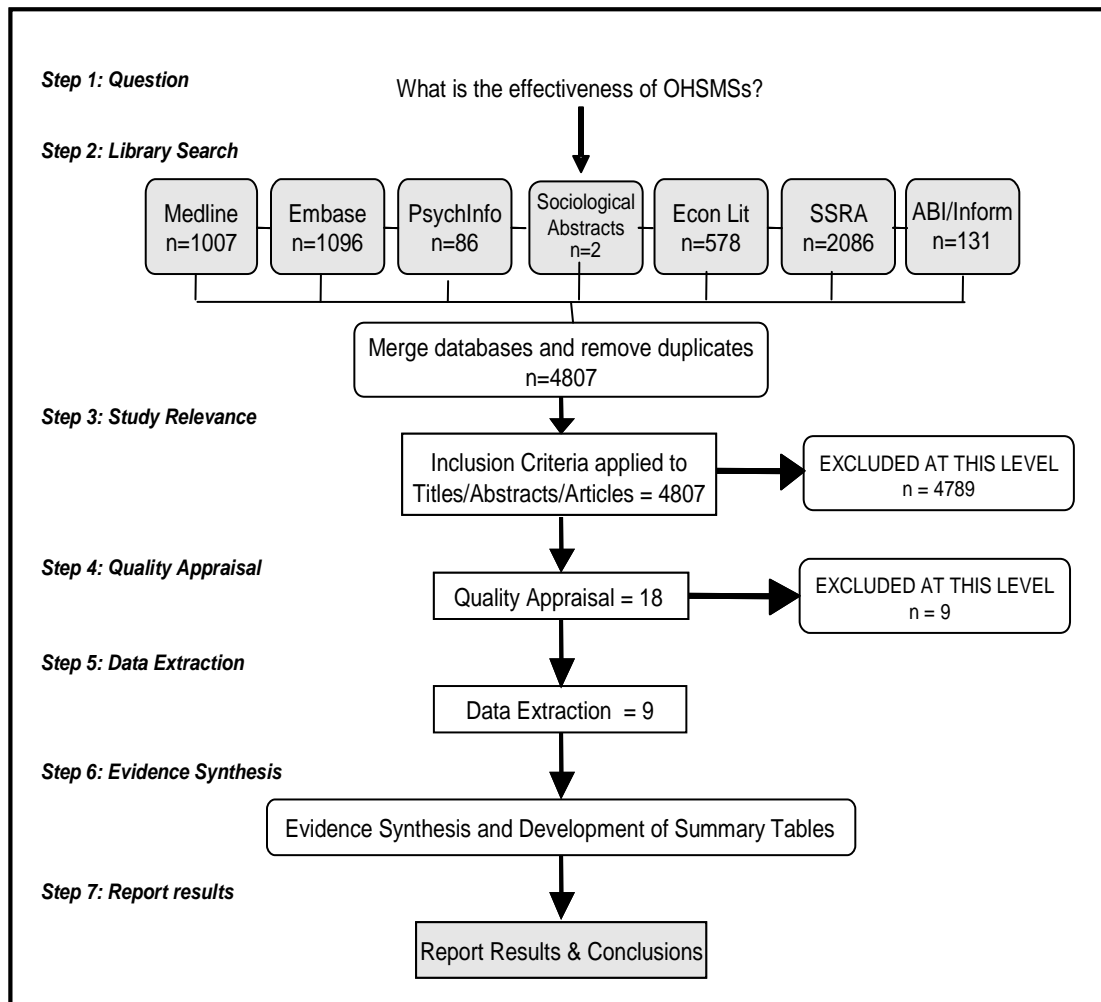
4. Results

4.1 Publications identified through the literature search

After merging the citations from the electronic search of the seven databases and removing duplicates, 4807 studies remained for possible inclusion in the systematic review. Following a review of titles and abstracts, and initial screening of full papers where necessary, 18 studies met the inclusion criteria and were appraised further for methodological quality.

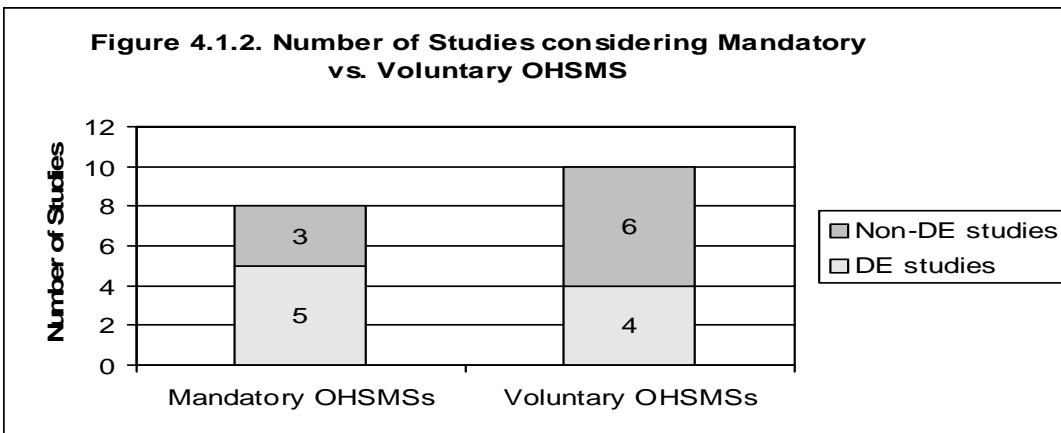
Of the 18 studies assessed for methodological quality, nine were of sufficient quality to proceed to the data extraction (DE) stage. A detailed breakdown of the flow of studies from the initial search strategy to evidence synthesis can be found in Figure 4.1.1.

Figure 4.1.1 Flowchart of studies in literature review

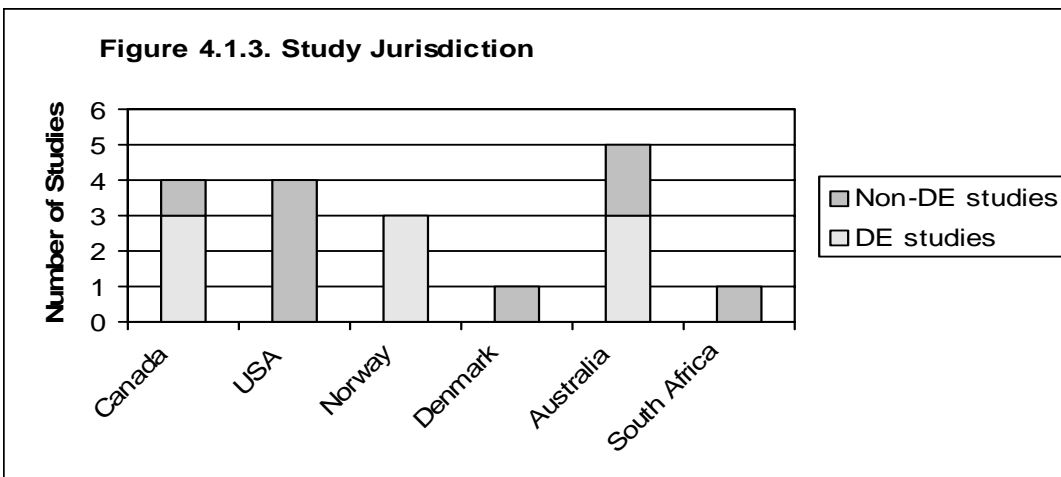


A list of all studies that were relevant to this review is in Appendix F. The list is divided into i) those studies which were relevant and of sufficient quality to proceed to data extraction (n=9), and ii) those studies which were relevant but not of sufficient quality to proceed to DE (n=9).

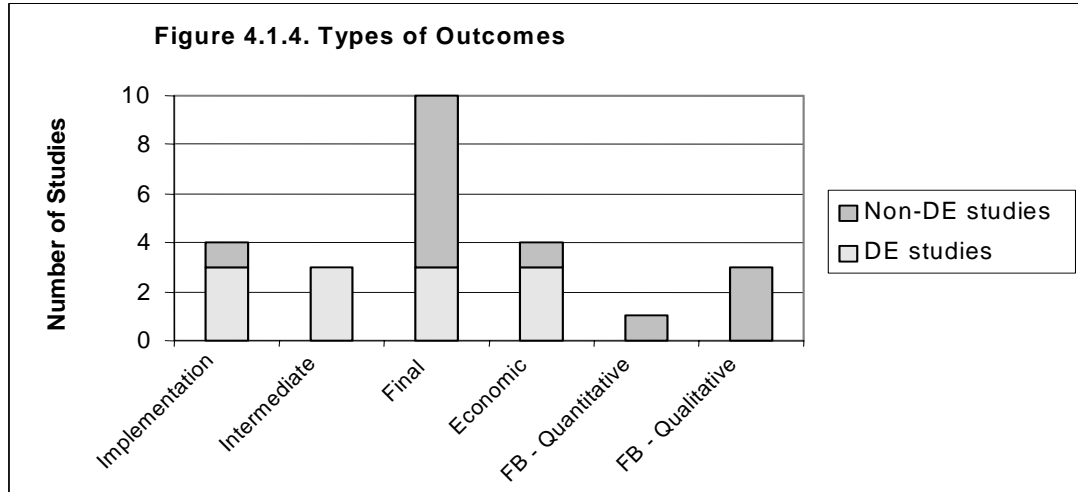
Study characteristics can be found in Appendix G, Table G.1. (studies proceeding to DE) and Table G.2 (studies not proceeding to DE). Overall, slightly more studies examining voluntary OHSMSs (n=10) entered QA, compared to those involving mandatory OHSMSs (n=8). However, fewer of the studies of voluntary OHSMSs met the minimum methodological quality to proceed to DE (4 voluntary vs. 5 mandatory) (see Figure 4.1.2).



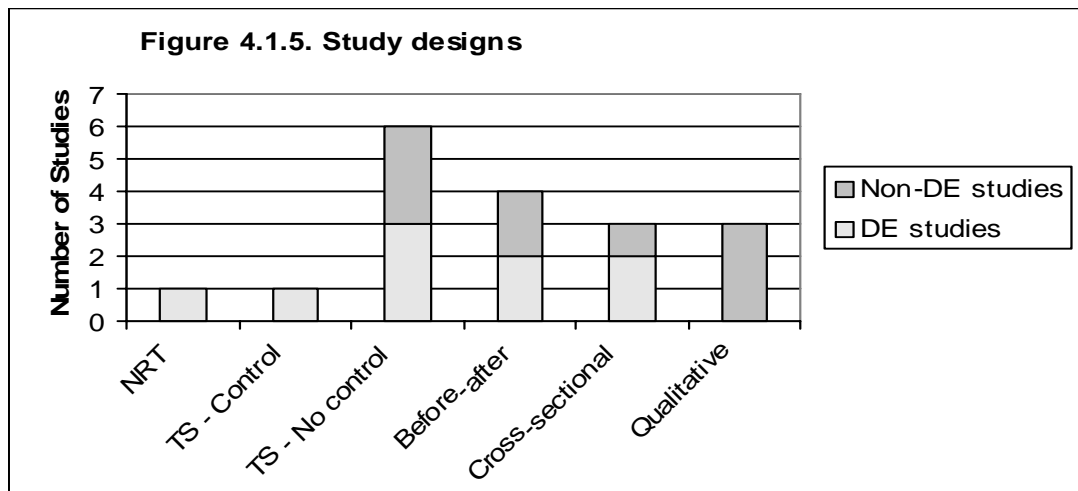
The majority of the studies were from North America (see Figure 4.1.3). Four studies were European: three from Norway and one from Denmark. Five studies were from Australia, and one was from South Africa.



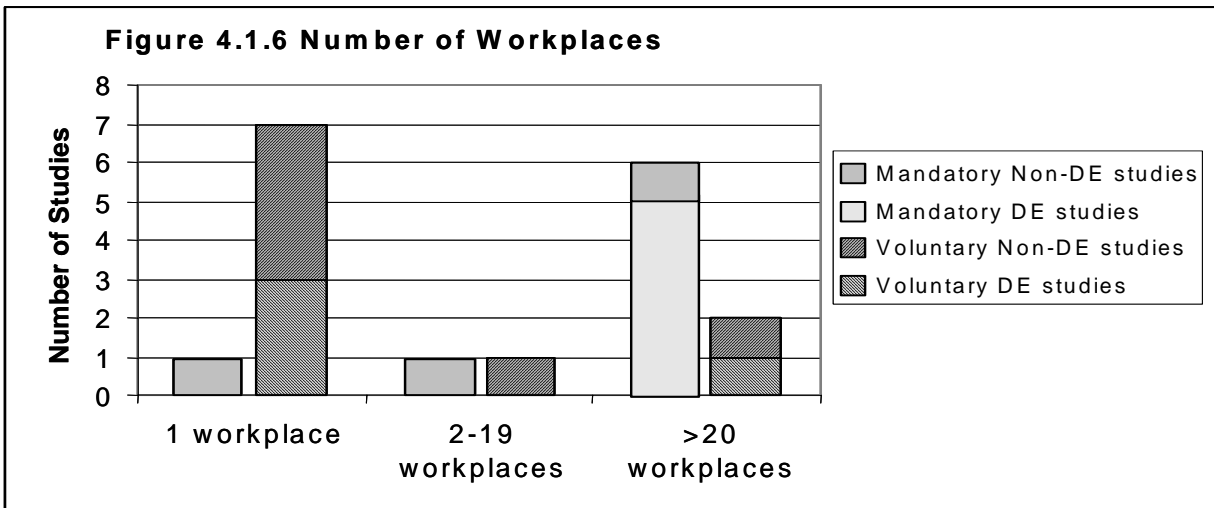
Four studies examined OHSMS implementation, 13 assessed OHSMS effectiveness (3 intermediate outcomes, 10 final OHS outcomes, 4 economic outcomes), and 4 addressed facilitators/barriers (1 with quantitative outcomes, 3 with qualitative). As can be seen in Figure 4.1.4, studies that proceeded to data extraction only considered implementation and effectiveness of OHSMSs. None of the studies examining facilitators or barriers were of sufficient quality to proceed to data extraction. This will be discussed further in Section 4.2 below.



Six types of study designs were identified: time series with a concurrent control group (TS – Control), time series without a control group (TS – No Control), before-after, cross-sectional, non-randomized trial (NRT), and studies involving qualitative data. Only two of these designs incorporate a control group (TS – Control and NRT). The majority of studies as shown in Figure 4.1.5, employed time series without control group or before-after designs. None of the studies with qualitative data were of sufficient quality to proceed to data extraction.



Most studies, as can be seen in Figure 4.1.6, considered either multiple (more than 20 worksites) or single workplaces. While more studies examining multiple workplaces proceeded to data extraction than those studying single workplace, there is a difference between studies exploring voluntary vs. mandatory programs. Of studies meeting the minimum quality to proceed to data extraction, those looking at voluntary OHSMSs generally involved only one workplace setting (1 workplace, n=3; 20+ workplaces, n=1); while those studies looking at mandatory OHSMSs all examined multiple workplace settings (i.e., >20 workplaces).



Studies examined workplaces of various sizes (see Figure 4.1.7). Those studies that proceeded to data extraction most frequently investigated workplaces of mixed sizes. However, small and medium enterprises (SME) and large workplaces were also well represented in the studies proceeding to data extraction. As might be expected, all of the studies that were unclear on workplace size were of insufficient quality to proceed to data extraction.

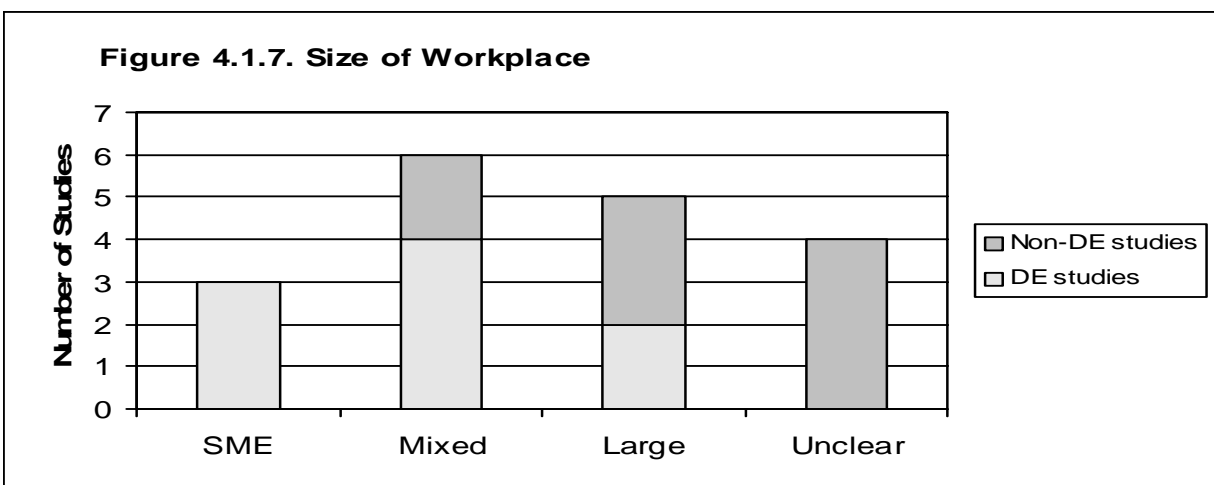
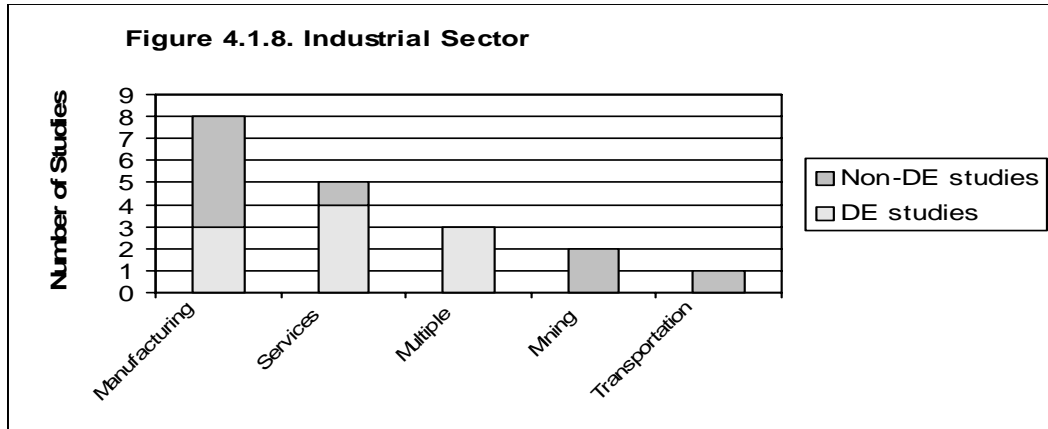


Figure 4.1.8 displays the industrial sector of studies proceeding and not proceeding to DE. There were five industrial sectors identified in this literature (manufacturing, services, multiple sectors, mining and transportation).



4.2 Quality of the literature investigating OHSMSs

A secondary aim of the review was to characterize the content and methodology of existing research literature on OHSMSs, thereby identifying research gaps and weaknesses, in order to provide guidance on future research in this area. Towards this end, this section characterizes the methodological limitations of studies of OHSMSs. The intention was that this would enable us to give some guidance for the design of future research in the field.

In this section, a summary of the methodological limitations of the 18 studies reviewed is presented, including both the nine studies retained for data extraction and the nine studies excluded after quality appraisal. The focus of the following discussion will be on the quality of the *different types of evidence* (implementation, intermediate OHS outcome, final OHS outcome, economic outcome, facilitators/barriers to OHSMS effectiveness). In addition, quality by mandatory/voluntary status of the OHSMS under investigation will be explored. Finally, special attention is given to controlling for confounding factors when investigating OHSMSs.

Summary of quality appraisals As shown in Figure 4.1.4, final OHS outcomes were the most frequently reported type of evidence in this body of literature (n=10; 40 per cent of all outcomes). Yet, with only three of the ten meeting the QA criteria, the quality of this type of evidence was, as a whole, low. The most common QA criterion not met with respect to the final OHS outcomes was “control for confounding factors” (see Table 4.2.1). The use of inappropriate statistical tests (including an inappropriate lack of a test) was the second most frequently encountered fault.

It is difficult to generalize about common limitations with respect to the quality of the evidence on implementation, intermediate outcomes, economic outcomes, and quantitative facilitators/barriers because of the small number of studies. However, if all effectiveness outcomes are considered together, an absence of any control for confounders and inappropriate statistical tests again appear to be the most frequent quality problems.

As noted in section 4.1, the search retrieved three articles containing facilitator/barrier evidence that had been based on qualitative data, i.e., data from interviews, focus groups, document review, or observations. In two articles, the interviews were part of an audit procedure. Although the means of selecting interviewees was sometimes reported, these audits were not described as containing an explicit inquiry into facilitating features or barriers to implementation, nor were the methods of analysis of the interview data described. In the third article, which involved case studies of two firms, data collection included taped interviews and observation, and the focus was not on OHSMSs but on the directness of employee involvement in management systems. In every case, an absence of any description of the studies' methodologies and/or the anecdotal nature of the results described led to a rating of the research methods as "Low" or "Very low".

The quality of studies also varied according to whether the OHSMS was voluntary or mandatory. More studies of a mandatory OHSMS met the QA criteria than did those of voluntary OHSMSs (see Figure 4.1.2). Specifically, the studies of mandatory OHSMSs were better at considering confounders, controlling for them, and discounting them than those of voluntary OHSMSs (see Table 4.2.1).

Table 4.2.1 Nature of methodological limitations in the 18 studies selected for quality appraisal (n=22*)

Type of Outcome	First Author (Year of Publication)	Study Design	Quality Assessment Criteria**					Overall Quality***
			Sample	Confounders	Measurement	Statistical tests	No additional bias	
Mandatory OHSMSs								
Implementation	Nytro (1998)	BA	+	+	-	+	+	Moderate
	Saksvik (1996)	BA	+	+	+	+	+	Moderate
Intermediate	Saksvik (1996)	CS	+	+	+	+	+	Moderate
	Torp (2000)	CS	+	+	+	+	+	Moderate
Final	Lanoie (1992)	TS	+	-	+	+	+	Low
	Lewchuk (1996)	TSWC	+	+	+	+	+	Moderate
	Saksvik (1996)	CS	+	+	+	+	+	Moderate
	Torp (2000)	CS	+	+	+	+	+	Moderate
Economic	Dufour (1998)	TS	+	+	+	+	+	Moderate
Facilitators/barriers (quantitative)	Nytro (1998)	BA	+	+	-	-	+	Low
Voluntary OHSMSs								
Implementation	Dotson (1996)	BA	-	-	-	+	-	Very low
	Pearse (2002)	BA	+	+	+	-	+	Moderate
Intermediate	Edkins (1998)	NRT	+	-	+	+	+	Moderate
Final	Alsop (1999)	TS	+	+	-	-	-	Low
	Anonymous (1993)	TS	+	-	+	-	+	Low
	Anonymous (1994)	BA	+	-	+	-	+	Very low
	Bolton (2001)	TS	+	-	+	+	+	Low
	Dotson (1996)	BA	-	-	+	+	-	Low
	Eisner (1988)	CS	-	-	+	+	+	Low
Economic	Alsop (1999)	TS	+	+	+	+	-	Moderate
	Dotson (1996)	BA	-	-	-	+	-	Very low
	Yassi (1998)	TS	+	+	+	+	+	Moderate
Facilitators/barriers (quantitative)	NONE	NA	NA	NA	NA	NA	NA	NA
<p>*n=22 because studies varied with respect to the number of different types of evidence examined: 13 studies examined one type of evidence, three studies examined two, and two studies examined three. (These totals consider quantitative study findings only).</p> <p>**Quality assessment criteria: "Sample" refers to sample selection and maintenance "Confounders" refers to the consideration and control of confounders "Measurement methods" refers to the methods used to measure the evidence "Statistical tests" refers to the tests performed on the evidence "No additional bias" refers to other sources of bias identified by the article reviewers</p> <p>***"Overall quality" refers to the overall quality rating provided by the article reviewers</p>					<p>BA = before-after; CS = cross-sectional; NRT = non-randomized trial; TS = time-series; TSWC = time-series with concurrent control</p> <p>"+" symbol: For "Sample", "Confounders", "Measurement methods", and "No additional bias" the "+" symbol means the study met that criterion (either fully or partially) For "Statistical tests" the "+" symbol means the study met the criterion, there were minor deficiencies which would have little or no impact on the conclusions, or no statistical tests were necessary</p> <p>"-" symbol: For all quality assessment criteria, a "-" symbol means the study did not meet the criterion (or it was unclear)</p>			

4.3 Effectiveness of voluntary OHSMSs

The term “voluntary OHSMS”, as used here, means those OHSMSs which have been developed and /or implemented by enterprises on a voluntary basis, rather than the result of legislation, regulations, and their enforcement. The impetus for undertaking voluntary OHSMSs may have been decisions taken within the workplace or sector, or there may have been encouragement by government bodies, their agencies or workers’ compensation insurers. In some cases, there may have been a research initiative. This was the case in Pearse (2002), where researchers recruited companies to investigate the suitability of OHSMSs for small and medium sized businesses.

The review’s search for relevant literature resulted in nine studies of voluntary OHSMSs, four of which (Edkins, 1998; Yassi, 1998; Alsop and LeCouture, 1999; Pearse, 2002) remained after quality assessment (summarized in Table 4.3.1). The outcomes used in these studies varied. One (Pearse, 2002) involved the degree of OHSMS implementation; the study by Edkins (1998) measured “intermediate outcomes” involving safety culture and employee hazard reporting; in the two remaining studies (Alsop and LeCouture, 1999; Yassi, 1998) the main focus was on economic outcomes..

Three studies (Pearse, 2002; Edkins, 1998; Alsop and LeCouture, 1999) were Australian; the other (Yassi, 1998) was Canadian, (Manitoba).

The voluntary OHSMSs in Australia were instituted in a context in which the various levels of government were strongly promoting a systematic approach to occupational health and safety, and agencies were supplying self-audit tools or system models (as described by Saksvik and Quinlan, 2003).

4.3.1 Studies of implementation Pearse (2002) described a project called Club Zero, funded by the WorkCover NSW Injury Prevention, Education and Research Grants Scheme. The purpose was to study the applicability of OHSMSs to small and medium-sized companies. The OHSMS intervention included guidelines tailored specifically for small to medium-sized metal manufacturers. In addition to the introduction of the OHSMS, the intervention involved the creation of a network of participating companies (20 fabricated metal products companies in southwest Sydney, Australia), and 10 network meetings in which participants shared resources and experiences and discussed particular aspects of OHSMS implementation.

At the start of the 27-month study period (before the distribution of guidelines and commencement of network meetings), participating companies completed a self-administered survey about their OHS

Table 4.3.1 Summary of Evidence in Voluntary OHSMS Studies

First Author (Year)	Sample	Intervention	Results	Conclusions/Interpretation
Alsop (1999)	1 large municipality in Australia	Integration of OHS management with quality and environmental management involving common corporate processes (training, procurement and supply, hazard identification and incident reporting, risk assessment and control), and risk factors specific to individual business units.	Economic outcomes: Reduction in Workers' Compensation premiums, 1995/6 to 1998/1999: from 4.00% to 1.90% Net premium rate as a percentage of remuneration in 1999-2000: lower premiums, compared to industry, in 19 of 20 business units	Authors: Significant improvements in key performance criteria for H&S management, including a major cultural change in intervention group. Reviewers: Implementation of Quality and Environmental Management Systems may also have contributed to results. Analysis of historical trend in premiums, and comparison of such with industry rates (or another municipal government) would be required to allow greater confidence in conclusions.
Edkins (1998)	2 medium-sized centres of a regional airline in Australia	Intervention group (I): Appointment of operational safety manager; focus groups to identify hazards; confidential safety hazard reporting system; safety meetings with management; safety information database; safety information to all staff. Comparison group (C): Confidential hazard reporting system only	Intermediate outcomes Airline Safety Culture Index (lower score = better safety culture): I: Pre 58; Post 45 C: Pre 69; Post 92 Hazardousness (perceptions of staff)(lower score = lower perceived hazardousness) I: Pre 140; Post 92 C: Pre 140; Post 120 Hazardousness likelihood (perceptions of staff) I: Pre 76; Post 44 C: Pre 76; Post 68 Number of hazard reports: I: 48; C: 9 Actions taken on identified safety hazards I:13; C:0	Authors: Program can have positive influence on airline safety performance, specifically: improving staff confidence in how safety is managed, increasing staff willingness to report hazards and incidents, improving organizational safety culture and reducing staff perceptions of the severity and likelihood of safety hazards occurring within the airline/ Reviewers: Uncertainty as to selection process for sites involved., and equivalency of the 2 sites. However, the study used multiple measures, some of which were objective, and all of which were consistent in their direction of change.
Pearse (2002)	20 small-medium sized fabricated metal product companies in Australia (4 withdrawals)	Guidelines tailored to participants' industry; networking meetings; audits with recommendations for change	Implementation: Change in % of OHSMS elements implemented between 1 st and 2 nd audit (mean): 9% Larger companies made more changes Of OHSMS elements, greatest change in Workplace Injury Management; least change in Performance Indicators (statistics)	Author: It was possible to improve the level of OHS management in small and medium companies although not all companies were motivated by the collaborative and voluntary approach. Reviewers: Non-random sample, and large refusal rate, so these results cannot be generalized to the larger population. The influence of the networking component was not analyzed in this study.
Yassi (1998)	1 large teaching hospital in Manitoba	Database to aid in record-keeping, data collection/analysis, hazard identification, risk assessment, performance measurement, continuous improvement, program and economic evaluation	Economic outcomes: % reduction in hospital's Workers' Compensation assessment rate, 1990-95: 23%. Estimated savings 1990-95, based on decrease in WC premium rates: \$2,866,000.	Author: Use of databases permitted targeting of areas requiring attention, substantial savings due to workers' compensation assessment reductions. Reviewers: Health care industry's premium rates' more moderate decline controlled for compensation system-related changes. Estimate of financial benefits would be improved by sensitivity analysis of premium rate assumed for the absence of intervention for the 1991-95 period. Available data support interpretation that financial benefit is attributable to prevention efforts, although analysis of historical trend in premium rates would be more conclusive.

management practices (32 yes-no questions drawn from SafetyMAP¹⁰), and were rank-ordered on this basis. Six months after the guidelines had been distributed and when networking meetings had begun, each company was audited. The audit included documentation review, observation of work processes and interviews with a range of individuals in the company. The audit tool included 77 closed, unambiguous questions. Recommendations and action plans were developed based on the audit. The audit was repeated six to eight months later, and scores from the two audits were compared.

Four of the 20 companies which had been recruited dropped out during the study. The final sample included six companies employing between 5-19 people, four with 20-99 employees, and five with 100-199 employees. One large company (with 350 employees) was also included owing to its position as an “industry leader”. It provided a partnering role in the intervention, demonstrating the fact that OHS management could be integrated into a business strategy. Audit scores of the 16 companies showed that implementation of the OHSMS had improved in all but one company in the period between the two audits. On average, the companies improved by nine percentage points (on a 100-point scale). The biggest improvement (15 per cent) occurred in a company that had ranked in the mid range (ninth out of 20) at baseline. The rank order of the companies remained roughly the same in the second audit as in the first. Although the large companies improved more than the smaller ones, it is reported that some smaller companies also made significant gains.

The companies’ average scores on individual OHSMS elements varied between 31 per cent (on performance indicators/statistics) and 87 per cent (on performance indicators/reporting) on the first audit; between 33 and 91 per cent (on the same two measures) on the second. The greatest change was in workplace injury management, which improved 20 percentage points, from 64 to 84 per cent in the same period; the next greatest change was on two measures (management commitment and policy; responsibility and accountability) on each of which a 12 per cent improvement was shown (from 61 to 73 per cent, and from 35 to 47 per cent respectively).

On the basis of audit results, the authors observed that the single factor most consistently associated with companies’ ability to make changes is whether they allocated specific resources (which could be as simple as allocating responsibility to an existing staff member) and developed a plan to carry out the changes recommended in the first audit.

¹⁰ *SafetyMAP (Safety Management Achievement Program), is an auditing tool designed by the Victorian WorkCover Authority to assist organisations improve their management of health and safety and implement a cycle of continuous improvement.*

The authors concluded that, although not all companies were motivated by this collaborative approach, for some small and medium sized companies, OHSMSs were applicable and improvements were gained by their introduction. Unfortunately, the role of the networking component (in which half the companies participated) was not analyzed. Reviewers noted that the results could not be generalized to a larger population of small and medium-sized companies because the sample was not randomly selected: in particular, the refusal rate among potential participants was substantial at the start of the project (67 per cent). No statistical analysis was provided, nor was there consideration of potential confounders which might have influenced the results.

4.3.2 Studies of intermediate outcomes The one study (Edkins, 1998) with intermediate outcomes investigated the effectiveness of a safety program called “Identifying Needed Defences In the Civil Aviation Transport Environment” (INDICATE) in improving the safety performance of small airlines. This OHSMS involved employee health and safety and also passenger safety and the prevention of catastrophic accidents. The study was done at a time when the aviation industry was being encouraged to be more active in identifying safety deficiencies following public inquiries into two airline crashes.

Two sites operated by a regional airline in Australia were involved in this non-randomized trial - the intervention site (with 81 employees) at which the INDICATE program was initiated, and the other (with 72 employees) acting as control.

The INDICATE program involved:

- appointment of an operational safety manager available to staff as a confidante for safety related issues;
- a regular series of staff focus groups to identify safety hazards within the organization;
- a confidential safety hazard reporting system;
- regular safety meetings with management;
- maintenance of a safety information database;
- regular distribution of safety information to all staff.

In the control group, only the confidential hazard reporting system was implemented.

The eight-month trial began in July 1995. A safety culture measure (the Airline Safety Culture Index (ASCI) was used. This had been developed as part of the project and was based on safety climate measures in the research literature. The culture measure and various risk perception measures were completed by control and experimental groups prior to the implementation of the program and again at the conclusion of the trial. Repeated measures ANOVAs were performed. The number of safety hazards reports submitted by each group was also tracked.

On the ASCI, the intervention group's scores indicated a better safety culture at the start of the trial and further gain over the course of the trial. The control group's scores got slightly worse over the same period. Repeated measures ANOVAs indicated that the inter-group difference in this outcome over the course of the trial was statistically significant, which the authors interpreted as an indication that the "safety culture improvement in the intervention group was a direct result of the safety program".

On the measure involving staff perceptions of risk (of common airline hazards), staff rated both the hazardousness of each factor and the likelihood of its occurrence. In each case, the intervention and control groups' scores were very similar at the start of the trial. At the conclusion of the trial, the intervention group's perception of hazardousness was lower than that of the control group and the inter-group difference was statistically significant. In the perceived likelihood of hazards occurring, the results were similar with the intervention group's perceptions changing more than those of the controls.

The authors also reported the number of safety hazard reports submitted in both groups, noting that 48 were submitted in the intervention group and nine in the control group. They list 13 action taken based on identified hazards which arose from the intervention group, and imply that none arose from the control group.

The author concluded that the results suggest that the INDICATE program can increase staff reporting of safety hazards and incidents, improve organizational safety culture, reduce staff perceptions of the severity and likelihood of safety hazards occurring within the airline, and improve staff confidence in how safety is managed. Reviewers had some concerns about the equivalency of the two sites, as there had been no investigation of staff characteristics at each site, and the initial differences in the ASCI scores furthered this concern. On the other hand, they noted that the study had used multiple measures, some of which were objective and all of which were consistent in their direction of change.

4.3.3 Studies of economic outcomes Two of the studies of voluntary OHSMSs (Alsop and LeCouture, 1999; Yassi, 1998) reported economic outcomes. Both described events in single organizations, and reported results as time series.

Alsop and LeCouture (1999) described the introduction of an OHSMS and its integration with existing risk management systems by the municipal council of Manningham, near Melbourne Australia in the mid-1990s. The municipality employed more than 500 people, although downsizing and restructuring was occurring during the time period covered by the study.

The OHSMS, which was based on quality (ISO 9001) and environmental (ISO 14001) management standards and the WorkSafe audit which the Victoria WorkCover had developed (Alsop, 2004) emphasized common corporate processes as well as risk factors that were specific to individual business units. The corporate processes involved training, procurement and supply, hazard identification and incident reporting, risk assessment and control. SafetyMAP audits led to certification of achievement during the subsequent four years.

The time series data indicated a downward trend in workers compensation premiums over the period 1991/92 to 1999/2000; rates were reduced by 52 per cent, from 4.00 per cent of remuneration in 1995/6 to 1.90 per cent in 1998/1999. Premium rates of the individual business units were also compared to the industry rates in the year 1999-2000, with Manningham having lower premium rates (as a percentage of remuneration) in 19 of the 20 units. Although results presented in graphs suggest that these differences were substantial, no statistical analysis was included.

Alsop and LeCouture concluded that, “The use of a systematic approach to OHS management has been a success at Manningham – it has helped to achieve sustained results that have been desirable and, in some cases, essential for financial viability.” Reviewers noted that one cannot disentangle the effects of the OHSMS from those of the quality management and environmental management systems, which were being implemented at roughly the same time. There would be greater confidence that the decrease in workers compensation premium had resulted from the OHSMS if Manningham’s premium rates over a longer historical period had been presented (to eliminate the possibility that this decrease was just a continuation of a long-term downward trend), and if they had been compared to the industry rate over this extended time period.

The paper by Yassi (1998) also involved the introduction of an OHSMS, and also reported financial outcomes. The focus of the paper was on the utilization of databases to guide occupational health programs in an acute and tertiary care teaching hospital in Winnipeg, Manitoba, with 6000 employees (a number which appears to have gradually increased over the years reported). In addition to revealing a downward trend in assessment rates, and comparing the hospital's workers compensation assessment rates for the years 1990-95 to the average rate paid in Manitoba, the authors reported estimated savings attributable to the reduced rate.

The intervention was a new risk assessment/risk management approach to different occupational hazards (biological/chemical, ergonomic, psychosocial) and was instituted in the early 1990's. It involved improved record-keeping; more systematic data collection and analysis using databases; hazard identification; risk assessment; planning programs to address risks; defining the programs' objectives and standards; assigning the responsibility for particular programs to particular individuals; program evaluation or surveillance' performance measurement; and continuous improvement; and economic evaluation. Some specific areas of intervention which resulted from this program were: promotion of staff vaccinations against influenza; pre-placement assessments for new employees; the introduction of new technology to prevent needle-stick injuries; a program of prevention/early intervention for back injuries.

The time series design covered yearly intervals from 1990-95. Accumulated savings were estimated by applying the hospital's 1990 premium rate to each of the payroll values for the five subsequent years, to estimate what the payments would have been in the absence of the program. The amounts actually paid for each of those years were subtracted from the estimated amounts and the differences presented as the estimated annual savings. The hospital's workers' compensation assessments decreased from \$1.51 (per \$100 payroll) in 1990 to \$1.17 in 1995 - a 23 per cent reduction. The average rates in Manitoba decreased 7 per cent in the same time period. Accumulated savings for the hospital from 1990-95 were estimated at \$2,866,000.

Yassi concluded that use of the databases permitted the analysis of departmental profiles of injury and health issues, systematic targeting of areas requiring attention, evaluation of the comprehensiveness, quality and efficiency of programs instituted, and appropriate allocation of resources. She described the use of databases as "invaluable" in effecting changes that resulted in reduced workers' compensation assessments and other significant savings over the five years following implementation of the program.

Reviewers noted that a sensitivity analysis of the premium rate that was assumed for the 1991-95 period in the absence of the intervention would have improved confidence in the estimates of financial savings but this was not performed. The reader is not given any information about the consistency of the workforce over the period, although payroll gradually increased by 13 per cent from 1990-92 and then subsequently levelled off suggesting that no major changes took place. As in the case of Alsop and LeCouteur's analysis, reporting of the longer-term trend in premium rates, and statistical analysis of the difference in rate changes between this hospital and others, would have allowed firmer conclusions.

4.3.4 Summary of results in studies of voluntary OHSMSs The studies involving voluntary OHSMSs were varied in nature. Two (Alsop and LeCouteur, 1999; Yassi, 1998) were descriptions of OHSMSs implemented in single workplaces (a hospital and a municipal government) and reported economic benefits. A third (Edkins, 1998) investigated the effect of an airline safety program (INDICATE) on a regional airline worksite compared to a control worksite and described the effect on safety culture, employees' perceptions of risk, employee reporting of hazards and actions taken. The fourth (Pearse, 2002) described a community intervention in which small and medium-sized metal fabrication companies were recruited into a program involving development and dissemination of OHSMS guidelines, group networking meetings and audits during and at the conclusion of the program, with the companies' progress in implementing the OHSMS being the major focus.

All four studies reported positive effects on their outcomes of interest. In spite of including some quantitative data, these studies were primarily descriptive in nature. Only Edkins reported any statistical analysis. The studies were all assigned a methodological quality rating of Moderate (moderate limitations), as none were entirely free of methodological shortcomings which could result in some bias in the results. None of the papers' authors attempted to situate their results (for example, their premium rates) within longer term trends, to give a more comprehensive picture of the nature of the change before and after the implementation of OHSMS. Nor were their study samples representative of a larger population of worksites/workers. As a consequence, the results of these studies cannot with confidence be generalized.

4.4 Effectiveness of mandatory OHSMSs

One approach to the implementation of OHSMSs is through regulatory mechanisms. Five studies of the evaluation of mandatory OHSMSs remained after quality assessment (summarized in Table 4.4.1). Three of the studies were based on the Norwegian regulations on Internal Control (IC) of health, safety, and the environment (Nytrø et al., 1998; Saksvik and Nytrø, 1996; Torp et al., 2000) and two focused on Canadian provincial regulations which included systematic occupational health and safety (OHS) approaches

(Dufour et al., 1998; Lewchuk et al., 1996). To facilitate the discussion of the studies, a brief overview of the regulatory approaches is provided first.

Norway Systematic approaches to OHS management were introduced into the Norwegian offshore oil industry in the 1980s with a major impetus being an oil rig disaster in March 1980 in which 123 workers were killed (Saksvik and Quinlan, 2003). The self regulation of the work environment through employee participation and monitoring was regarded as extremely successful in the offshore operations and laid the groundwork for a national regulatory system.

The Norwegian regulation requiring a systematic approach to OHS came into effect January 1, 1992. The rule, titled “Systematic Health, Environment, and Safety Work” and also referred to as the ‘Internal Control’ (IC) regulation, made it mandatory for every enterprise in Norway to establish an OHS system, regardless of trade or size (Saksvik and Nytro, 1996; Nytro et al., 1998). The regulation specified that systematic actions to comply with and document activities of health and safety control be performed in accordance with existing health, environment and safety (HES) regulations¹¹, the most important being The Work Environment Act of 1977.

In response to criticisms that the IC regulation was not understandable and the finding that only a small proportion of enterprises had established IC systems, the regulation was modified and the revised version was issued in 1997. The revised regulation placed more emphasis on activities and performance as opposed to documentation (Nytro et al., 1998).

Canada In Canada, the provinces have regulatory authority for health and safety for most workplaces within their jurisdiction (except federal employees and some exempt workplaces). One of the included studies was conducted in Quebec, using data from 1985-1988, which focused on the impacts of landmark legislation that created the Commission de la Securite du Travail (CSST) in 1980. The other study was conducted using data from the province of Ontario, over the period 1976 to 1989, and examined the effects of Bill 70 which took effect in late 1979 and which introduced the “Internal Responsibility System”.

In January 1977, Bill 139, “An Act Respecting Employees’ Health and Safety” came into effect in Ontario. This Act, which allowed workers the right to refuse any work they believed to be dangerous and empowered the Ministry of Labour to order the establishment of joint health and safety

¹¹ These included the Pollution Control Act, Product Control Act, Civil Defense Act, Act Relating to Electrical Installations and Electrical Equipment Changes, and legislation pertaining to fire and explosion hazards and fire prevention.

committees (JHSCs), moved Ontario toward a model under which emphasis was placed on having employees and employers regulate their own working conditions. The Internal Responsibility System was first suggested by the 1976 Ham Commission and it was anticipated that making both workers and employers responsible for OHS would be more effective and less costly (Lewchuk et al., 1996; Walters, 1983).

More extensive omnibus legislation in OHS was introduced to the Ontario legislature in October 1977 and led to the 1979 Occupational Health and Safety Act (Bill 70). Bill 70 formalized the Internal Responsibility System and the establishment of JHSCs became mandatory in firms employing more than 20 people or when ordered by the Minister.

A review of Bill 70 showed that a number of OHSMS components were mandated by the Act in workplaces with over 20 employees (see Lewchuk et al., 1996 in Table H.2).

In the province of Québec, landmark OHS legislation in 1979, *Loi sur la santé et la sécurité du travail (LSST)* (in English, ‘Act Respecting Occupational Health and Safety’), created *La Commission de la Santé et de la Sécurité du Travail (CSST)*. The CSST had a mandate to act as the workers’ compensation insurer and had the responsibility for direct regulatory control over OHS in the province. The CSST was mandated to administer the Act and six other acts pertaining to the health and safety of workers (CSST, 1988). The Act included the right to refuse hazardous work, the creation of JHSCs, the requirement of a prevention program, and the right of protective reassignment (Lanoie, 1992). Protective reassignment and right to refuse applied to all employees in all Quebec workplaces, whereas JHSCs and prevention programs applied only to firms with 21 or more workers in specific, higher risk, industry sectors. The CSST had the authority to serve violation notices to firms that failed to submit copies of their prevention program or if the submitted prevention program was inadequate (CSST, 1987).

A review of the requirements of the *Loi sur la santé et la sécurité du travail* indicated that many OHSMS elements were not specifically covered and the Act did not require integration of health and safety with other management activities (see Dufour et al., 1998 in Table H.2).

4.4.1 Studies of implementation Two of the five studies provided evidence on the implementation of the Norwegian Internal Control (IC) regulation (Nytro et al., 1998; Saksvik and Nytro, 1996). The first of these studies was a cross-sectional investigation by **Saksvik and Nytro (1996)** conducted in 1993, one year after the IC regulation came into effect (January 1992). The investigation had multiple objectives including the determination of the prevalence of implementation of IC systems, identification of IC induced organizational changes, and discovering whether absenteeism and accident rates were related to implementation. The survey, conducted amongst a randomized quota sample of 2092 public and private enterprises in Norway, was implemented using computer assisted telephone interviewing. The respondents, as described by the authors, "...for the most part were manager representatives" of each of the firms. The health and safety data that were obtained were derived primarily from archival records but the authors indicated some were based on recall or estimates provided by the respondent.

The 45 questions in the survey had been pre-tested in previous research involving 500 enterprises that had implemented IC prior to 1992. Of these questions, four provided information pertaining to IC implementation (this study also provided evidence for intermediate and final OHS outcomes). Study findings were reported based on a sample of 915.¹² Findings (percent firms) for reported changes due to the implementation of IC were as follows: more clear lines of responsibility (58 per cent), better documentation (58 per cent), more/better risk assessment (46 per cent), and new strategic plans (42 per cent). No statistical tests of significance or precision were conducted. An additional finding that did not fit into any of the review's evidence categories was that 53 per cent of respondents reported integration of IC with Total Quality Management as a result of IC implementation.

Findings for implementation outcomes from the Saksvik and Nytro (1996) study should be interpreted with caution as these were not the investigators' primary research questions. Additionally, the overall study design raises serious concerns about internal and external validity. Randomized quota

¹² It was not clear to the reviewers how this number of respondents was obtained as 66 per cent of respondents (n=1381) reported not having begun to implement the IC regulation (excluding these respondents from the data would leave 711 subjects). The authors contacted Dr. Saksvik and he suggested that the "extra" respondents to these questions were likely those who had developed systems similar to the IC system (e.g., TQM) yet answered "no" to actually have begun implementing IC.

Table 4.4.1. Summary of Evidence in Mandatory OHSMS Studies

First Author (Year)	Sample	Intervention	Results	Conclusions/Interpretation
Dufour (1998)	All Quebec workplaces in manufacturing sectors subject to LSST regulations 1985-88.	LSST (1979)	Economic outcomes: Prevention program (PREVENT) and INFRACTION variables had statistically significant and positive regression coefficients, so appeared to have a positive effect on productivity growth. Implied contributions of both variables on productivity growth 0.007 (at sample means).	Authors: Prevention programs and penalties have reduced the incidence of workplace injuries in manufacturing sector, leading to reduction of direct and indirect costs sufficient to have an enhancing effect on productivity growth. Reviewers: The question remains as to the nature of the OHSMS - IC regulations have the intention of moving firms to OHSMS and the PREVENT variable may be a good indicator of companies actually implementing them.
Lewchuk (1996)	637 Ontario workplaces (497 who responded to 1991 IAPA survey; 140 others)	Bill 70 (1979)	Final outcomes: Regression coefficient, $\beta = -0.015$ for LTI frequency regressed on legislation change. In other words, 18% decrease in LTI (10 yrs post-intervention vs. 4 yrs pre-intervention)	Authors: Bill 70 had a significant effect in the direction of lower frequencies for manufacturing but was not significant for retail sector. The reduction in lost-time accident frequencies was around 18% for manufacturing workplaces. Time variable was significant and positive in both regression analyses, indicating growing accident/illness rates in both sectors. Reviewers: A major confounder is the change in workers' compensation administration and shift to New Experimental Experience Rating (NEER). Weaknesses in the statistical analysis: potential for selection bias; confounders.
Nytro (1998)	Randomized quota sample, representative of Norwegian firms re: size, industry, location. 1993: n=2092 1996: n=1182	Internal Control (IC) regulation 1992	Implementation: Increase in % of companies that reported having fully implemented IC, 1993-96: 37%. 1993: 8% implemented; 25% in progress; 67% not yet started. 1996: 45% implemented; 36% in progress; 19% not yet started.	Authors: No specific conclusions on implementation (study's data on other outcomes not reported here due to exclusion at QA stage) Reviewers: Social desirability bias in reporting is possible, but is unlikely to explain all of the change seen here. There do not seem to be alternative explanations for the observed change.
Saksvik (1996)	Randomized quota sample, representative of Norwegian firms re: size, industry, location. N=2092	Internal Control (IC) regulation 1992	Implementation (% respondents report category of change) ¹³ : More clear lines of responsibility 25% More/better risk assessment 21% Better documentation 25% New strategic plans 18% Intermediate outcomes (% change) ¹³ : HES awareness 39% Final outcomes: <i>Regression results for absenteeism development:</i> T values: IC status 2.94 (p<0.01); <i>Regression results for accident development:</i> T values: IC status 1.66 (p >0.05);	Authors: The most frequently reported changes in HES practices after introduction of IC regulations were as follows: increased HES awareness, clearer lines of responsibility, improved risk assessment and documentation, and new strategic plans. IC status contributed significantly to explaining the variance in development of absenteeism 1990-1992, but not for accidents. Two regression models were able to explain only a small part of the total variance in absenteeism and accidents. The variables that contributed most to the models were level of absenteeism and accidents prior to implementation. Reviewers: Study examined a stratified random quota sample, but response rate not reported. Possible recall bias for final outcomes (absenteeism and accident rates). These were based for the most part on administrative archival data, but some workplaces presented recalled self-report data.
Torp (2000)	1567 managers and garage workers in 237 garages in 1996	Internal Control (IC) regulation 1992	Intermediate outcomes: IC status significantly correlated with 7 of 9 outcomes, indicating IC had positive effects. Final outcomes: Employees at garages with higher IC states reported significantly fewer musculoskeletal symptoms (p<0.01). No significant relationship between IC measures and sick leave. Standardized regression coefficients for (1) internal control index; and (2) internal control status rated by manager: Musculoskeletal symptoms: -0.026; -0.076**. Sick leave in last 30 days: -0.048; -0.013.	Authors: Systematic HES activities positively and significantly correlated with following outcomes, measured at individual level: satisfaction with HES activities, physical working environment, social support, HES-related management support, health-related support, and control and workers' participation in activities related to occupational health. Reviewers: Some sources of potential bias noted, e.g., selection bias, (low response rate); social desirability bias (exclusive use of self-report measures). Conclusions about causality not possible because of cross sectional design.

¹³ These values are derived from the values given in the report for the sub-sample that answered these questions (n=911) by accounting for the size of the sub-sample that responded divided by the size of the entire sample (the non-respondents were those who hadn't started implementing IC at all)

sampling may have introduced selection bias but no information on response rates was provided to assess this possibility.¹⁴ The interview questions assessed implementation in two different ways. In one, respondents were asked for their assessment of implementation. In the other, researchers deduced the degree of implementation from responses to questions about particular IC-related functions.

Nytro et al. (1998) conducted a cross-sectional survey of Norwegian firms to determine the change in prevalence of IC implementation, using data from the previously described study (Saksvik and Nytro, 1996) for comparison.¹⁵ Both studies used a randomized quota sample of private and public enterprises and computer assisted telephone interviewing. The Saksvik and Nytro (1996) study found that 66 per cent of firms had not begun implementation, 25 per cent reported they were in the process of implementation, and 8 per cent reported that IC had been implemented. Of the 66 per cent of firms that had not started implementation, 58 per cent had not heard of the regulation and of those most were small in size. The findings, 4 years after the regulation had come into effect, were that 19 per cent of firms had not yet begun implementation (47 per cent decline), 36 per cent were in the process of implementation (11 per cent increase), and 45 per cent had implemented IC (37 per cent increase) (Nytro et al., 1998).¹⁶

The study had a number of design characteristics affecting the internal and external validity of the implementation findings. Randomized quota sampling, based on firm size, industry, and geographic location, was used to obtain the final sample in both cross-sectional surveys (sample sizes were 2092 in 1993 and 1182 in 1996) in order to ensure that the final sample was representative of all Norwegian firms. The sizes of the firms in the 1996 survey were reported as: 43 per cent with less than 10 employees, 39 per cent less than 100, and 18 per cent had over 100. Refusal rates for the computer assisted telephone interviews were not provided and so the possibility of selection bias cannot be ruled out.

As the Nytro et al. (1998) investigation did not have the study of implementation as its primary objective, the published article provided little detail about the items in the interview that collected information on the

¹⁴ It should be noted that in a more recent investigation completed by the same research team using similar sampling methods, their response rate was 45 per cent (Saksvik et al., 2003).

¹⁵ The study design for the implementation evidence was classified as before after based on their use of data from the previously conducted survey.

¹⁶ A more recent study by the authors (Saksvik et al., 2003), which was not received in time for full review, provides a further update of the implementation rates in Norwegian firms. The survey was conducted in 1999 and findings were: 14 per cent of firms had not yet begun implementation (5% decline from 1996), 39 per cent were in the process of implementation (3% increase from 1996), and 47 per cent had implemented IC (2% increase).

degree of IC implementation. Referring to the methods used in the 1993 survey (Saksvik and Nytro, 1996), the level of implementation was based primarily on the response to a question about how far the implementation had progressed on a five point scale from “not started” to “finished”. Findings from Gaupset (2000) that measures of implementation constructed from responses to questions about specific activities were similar to measures from self-reports provide support for the validity of the measure used. However, the measure used was subjective and was vulnerable to response bias. It is unlikely that this would explain all of the effects reported.

4.4.2 Studies of intermediate outcomes Two studies provided evidence on intermediate outcomes that included changes in health, environment and safety (HES) awareness (Saksvik and Nytro, 1996) and changes in employee perceptions of the work environment (physical and psychosocial) and HES activities as a result of OHSMS implementation (Torp et al., 2000). Both investigations were conducted in Norway and focused on the level of implementation of the IC system in relation to the intermediate outcome variables.

The cross-sectional survey conducted by **Saksvik and Nytro (1996)** was described previously. A series of questions on the survey were used to construct an index of ‘status of IC implementation’. Respondents were asked about nine specific activities that the authors thought to be a part of, or a consequence of, IC implementation. Amongst those who reported changes due to IC implementation, 69 per cent reported increased health and safety awareness (representing 30 per cent of the total sample). In addition to the limitations described previously, for the subjective question about a change in health and safety awareness, there was the potential for response bias.

Torp et al. (2000) presented the findings of a cross-sectional study performed in the Norwegian motor vehicle repair industry in 1996. Motor vehicle repair garages (n=130) scheduled to participate in a health and safety management course along with 181 non-participating garages (approximately 130 matched to the participating garages by firm size and geographic location) were invited to participate in the study. A manager from each garage was sent a questionnaire on IC characteristics at the organizational level and also was asked for consent to distribute questionnaires to employees. The response rate of managers was 80 per cent and 2174 questionnaires were distributed to employees in 237 garages (warehouse and office workers excluded). The response rate among employees was 72 per cent. Managers’ responses were used to construct two indices of implementation of IC at the firm level and the study’s aims were to investigate the relationships between the level of IC implementation and the satisfaction of workers with HES activities, their working environment, participation in HES activities, as well as other outcome variables (discussed in the final outcomes section).

One of the IC indices, termed “IC status”, was based on the manager’s stated assessment of the level of implementation at their garage (from “don’t know what an IC is” to “implemented IC”, on a five-point scale). The other, called the “IC index”, was based on 16 questions considered relevant to establishing IC in firms (e.g., “does the garage have a health a safety deputy?”). Variables from the garage workers’ questionnaires were scored on a 7-point scale (poor=1, good=7). The results of the multiple regression analyses showed that the IC index based on the 16 items was significantly correlated with 6 of the 9 intermediate OHS variables.¹⁷ These significant correlations all indicated that the IC had positive effects. Higher levels of correlation were obtained in the regression analysis with IC status as rated by managers (7 of the 9 were statistically significant). The highest correlations were with “satisfaction with HES activities at the garage” and the IC index (0.16, $p<0.001$) and the same variable with IC status as rated by the manager (0.15, $p<0.001$).

These findings from Torp et al. (2000) do provide some evidence of correlation between measures of IC implementation and some intermediate outcomes. The cross-sectional design is discussed by the authors as limiting the conclusions drawn from the study because of ambiguity in the direction of causality.¹⁸ Additionally, the respective response rates for the companies participating in the training and those not scheduled for training were not provided and differences may present the potential for selection bias towards better or poorer performers in HES. The exclusive use of self-reported measures suggests that reporting bias may be problematic.

4.4.3 Studies of final OHS outcomes Three of the studies of mandatory OHSMSs reported final OHS outcomes (Saksvik and Nytro, 1996; Torp et al., 2000; and Lewchuk et al., 1995).

In the previously described Norwegian study by **Saksvik and Nytro (1996)**, the hypothesis was that the status of IC implementation was related to changes in the rates of absenteeism and accidents. Three levels of IC implementation were developed based on responses to 13 questions. The types of questions and IC implementation measures were as follows: one question (rated from “not started” to “finished” on a 5 point scale), “IC status after own opinion”; nine questions related to number of completed IC activities, “IC status quantitative measure”; and three questions based on the

¹⁷ *The nine intermediate OHS variables were: satisfaction with HES activities at garage, satisfaction with physical working environment, satisfaction with psychosocial working environment, psychological job demands, decision authority, social support, HES-related management support, health-related support and control, and workers’ participation in HES activities.*

¹⁸ *The reviewers agreed and thought that factors such as “commitment to HES”, that could cause both better IC implementation and better HES outcomes, could result in a correlation between implementation and intermediate outcomes.*

authors' view of the critical factors for the implementation process (better risk assessment, new action plans, and better documentation), "IC status qualitative measure".

The survey consisted of 45 items including questions on the size of the enterprise, sector (private or public), HES activity level prior to IC implementation (HES activities accomplished in last 3 years such as: assessed psychosocial work environment factors, risk assessment analyses, having worked out action plans to improve work environment, etc.), HES result level prior to IC introduction (absenteeism and lost time accident levels in 1990), HES competence (having a worker HES representative and an occupational health service), and time pressure (an ordinal variable with 4 points related to increased productivity and reduced staff). Although no information on the validity or reliability of these measures was provided, the interview had been tested in previous research with 500 enterprises that had started implementation of IC before 1992.¹⁹

In the Saksvik and Nytro (1996) study, information on accidents and absenteeism was obtained from respondents during interviews and was, in most cases, from archival data. Ordinal variables with three levels (higher – stable – lower) were developed for both the rate of absenteeism and rate of loss time accidents (one day of ill health after the accident) based on reported levels in 1990 compared to 1991-1992 (up to first half of 1992). Multiple regression models for absenteeism and accident development from 1990 to 1992 were developed. They found that all three measures of IC implementation (status) performed well in the model for absenteeism and were all statistically significant; the strongest IC status measure was the qualitative variable ($p < 0.01$, $\beta 0.09$). In the accident model, none of the IC status measures were statistically significant. R-squared values for the two final models were 0.05 and 0.12, respectively.

The primary findings from the Saksvik and Nytro (1996) study were that the reported decrease in absenteeism was, as hypothesized, related to IC status but that there was no statistically significant relationship between IC status and change in accident rates. The authors pointed out that the finding for absenteeism may have been an artifact as the IC regulations may favour enterprises with a tradition for systematic long-range planning in other parts of their business management (thus, this may lead to confounding or effect modification). Additionally, the authors point out that their analyses were conducted at the enterprise level and that part of the variance in the dependent variables can be attributed to individual factors or interactional factors (subject-enterprise). Although the direction of change in accident rates was from high to low, the relationship with IC status was not significant. It is possible that the period of observation was too short to

¹⁹ *Firms in the offshore oil industry and the onshore aluminum industry were required to establish IC program before 1990.*

observe the impacts of IC status on change in accident rates.²⁰ Additionally, measurement error in the dependent variables along with the categorization of responses may have limited the findings. The reviewers also commented that measurement bias may have been a factor because some of the safety and health data were based on recall (most from archival data but recall on remaining). An additional concern was the fact that the two dependent variables appear to have been ordinal, yet the results of an ordinary least squares regression were provided. The authors should have considered a modeling technique more appropriate for ordinal data.

Torp et al. (2000) (described in the intermediate outcomes section above) also provided evidence on musculoskeletal symptoms (including pain), since the questionnaire distributed to motor vehicle repair garage employees contained seven questions related to musculoskeletal symptoms experienced in the last 30 days. Six were derived from the Subjective Health Complaints Questionnaire and one question on knee pain was added. The severity of pain was rated on a four-point scale. One question on sick leave asked respondents whether they had been away from work in the last 30 days (dichotomous response).

There were two measures of IC implementation, as discussed previously. Findings were presented from a number of multiple regression analyses with each of the final outcome variables as the dependent variable and including each IC measure along with control variables for company size and unionization. For the model involving musculoskeletal symptoms, coefficients for managers' rated IC status were significant (0.077, $p < 0.05$). These results indicated a negative relationship between IC status and musculoskeletal symptoms (employees in garages in which the manager rated IC status as higher reported fewer symptoms). Neither coefficient for the two measures of IC implementation in their respective models for sick leave were statistically significant.

The evidence for final outcomes from Torp et al. (2000) is unreliable because of the possibility of selection bias, response bias, and most importantly, because of the limitations of their study design. The authors point out that conclusions about causality are not possible because of the cross sectional design. Although the investigators made efforts to select the most appropriate garages for the study, the selection was non-random as 130 had managers enrolled in an OHS course and 200 others were selected by other means.²¹ Response rates for the respective groups of garages were not provided, raising further concern over possible section bias.

²⁰ Additionally, the reviewers felt that the finding of a significant decrease in absenteeism may have been an artifact because it was found over such a short period of observation and up to only 6 months after the legislation.

²¹ 130 of the garages in the non training group were matched to the firms participating in the course and another 70 were randomly selected.

The aim of the investigation by **Lewchuk et al. (1996)** was to examine whether the legislation in Ontario, Bill 70 had any impact on workplace level OHS performance. The authors stated that with Bill 70 "...a number of the rules of the game were changed at that time and caused changes in the workplace – whether or not these were immediately translated into JHSCs (which became mandatory for most workplace after Bill 70)." Their research question was whether or not there was a reduction in work-related injury and illness frequency after the act came into effect.

To address this question, a multiple, pooled cross-sectional time series study with a comparison group was performed. The sample consisted of 637 workplaces (436 in the manufacturing sector and 201 in retail) who were participants in a previous survey in 1991, along with 140 additional workplaces who were non-participants in the 1991 survey (selection method not specified). To be eligible for the previous survey, manufacturing workplaces had to have 50 employees and retail workplaces 20 employees in 1988. For these workplaces, annual data from 1976 to 1989 for lost time claims records, short and long run compensation costs, and employment level by rate group were obtained from the Workers' Compensation Board. Regression analyses were conducted for each sector with the indicator variable, Bill 70, assigned a value of 0 for the years up to 1979 and a value of 1 for 1980 and after. The authors felt that the retail sector could be thought of as a control group since critical sections of the legislation (including formation of JHSCs) did not apply.

The results of the regression analyses, in which the dependent variable was injury/illness frequency and included employment, time, union status, and industry as control variables, indicated that Bill 70 had a significant effect in the direction of lower frequencies for manufacturing whereas it was not significant for the retail sector. The authors indicated that the size of the coefficient for Bill 70 was such that the reduction in lost-time accident frequencies was in the order of 18 per cent for manufacturing workplaces. The time variable was significant and positive in both regression analyses, indicating growing accident/illness rates in both sectors; the authors mentioned that they conducted other analyses using the time variable, including fitting time trend as a quadratic in the models, but that the final linear model was appropriate.

The important finding from Lewchuk et al. (1996) is that there was a significant reduction of work-related injury and illness rates in manufacturing after Bill 70 compared to before the regulation whereas no significant reduction was observed in the retail sector. They also found that the effects were progressively larger for years 1980, 1981, and 1982 in manufacturing, indicating that the impact of the legislation grew over the study period (this trend was not observed in the retail group). The study nonetheless has a number of limitations. The authors suggested that a number of other variables may have had an impact on OHS performance and

may have biased the findings; one major confounder would have been a change in workers' compensation administration and they mention that over the study period there was a shift to the New Experimental Experience Rating (NEER) system in Ontario but that only one rate group (plastics) entered the program. The reviewers also identified some weaknesses in the statistical analysis and reporting, potential for selection bias, as well as the possibility of the effect of confounding variables.

4.4.4 Studies of economic outcomes One of the studies of mandatory OHSMSs reported economic outcomes (Dufour et al., 1998).

The **Dufour et al. (1998)** study was conducted to determine what impacts, if any, did environmental and OHS regulations have on the rate of total factor productivity (TFP) growth in the Quebec manufacturing sector during the 1985 to 1988 period. As discussed previously, OHS legislation in 1979, *Loi sur la santé et la sécurité du travail (LSST)*, resulted in the creation of CSST in 1980. The CSST had been innovative in the adoption of safety policies that included compulsory prevention programs, the right to refuse dangerous work, and protective reassignment. In addition to the study's broad aim to determine the overall effects of Quebec OHS regulations on TFG growth, the study also addressed the possible impacts of these specific OHS policies as well as the level of regulatory compliance activity.

Dufour et al. (1998) presented a theoretical discussion of the expected impact of OHS and environmental regulations on TFP growth. They pointed out that previous analyses of this type were more limited in terms of the variables included and did not allow for the potentially positive impact of regulation on productivity growth. Their theoretically derived equation allowed for positive impacts of regulations, included variables that provided a more extensive assessment of the impact of OHS regulation, and incorporated control of two important confounders: economies of scale and business cycle fluctuations.

In the empirical study, the variables that were important to the review's research questions were as follows: the dependent variable, TFP growth – difference between real output growth and real input growth with real output growth based on the annual value of shipments and real input growth determined from the cost of materials and supplies, cost of energy, wages (production, administration and other non-manufacturing employees), and capital costs; five variables for intensity of OHS regulation, Inspection – rate of inspections per 1000 FTE employees, Refusal – rate of interventions from CSST for work refusals, Protective – rate of protective reassignments, Infraction – rate of penalties imposed, and Prevent – percent of firms that had adopted prevention programs.

Annual industry level data for all Quebec workplaces in 19 manufacturing sectors were obtained from CSST annual reports, Statistics Canada, the

Bank of Canada, other federal sources as well as unpublished reports. Pooled time series regression analyses, based on the three years of annual data across the 19 manufacturing sectors, were used to develop six different specifications of the authors' theoretical equation explaining variation in TFP growth. All of the specifications had satisfactory explanatory power (R-squared values from 0.54 to 0.77) and were relatively consistent.²²

The specification with all the variables included was used to explain the impact of each on the TFP growth. Three of the five OHS variables were significant in the final model, with Protective (rate of protective reassignments) having a negative coefficient thus a negative impact on TFP growth. Infraction (rate of penalties imposed) and Prevent (percent of firms that have adopted prevention programs) both were significant and positive and the implied contribution to productivity growth was estimated at 0.007 (the Prevent variable itself was calculated by the reviewers to make an estimated TFP growth contribution of 0.006).²³ The Protective variable had a large impact on TFP growth (-0.019) and taken together the overall impact of the OHS variables was -0.012 which was larger than the impact found for OHS regulation in the United States (-0.003).

The findings of Dufour et al. (1998) provide insights into the impacts of important aspects of the Quebec regulation and CSST activities on the rate of productivity growth over the 3 year study period. The authors suggested that prevention programs and penalties may have reduced workplace injury incidence, leading to reductions in both direct and indirect costs related to accidents thus enhancing productivity growth. The finding of a statistically significant positive impact of prevention programs (the variable was significant at the 5 per cent level in all the specifications) on productivity growth has application to this review's research questions. The model was well developed theoretically and the statistical analyses were appropriate and comprehensive.

Although the findings for the prevention programs were consistent and the model robust, the nature of the study design raises concerns about the strength of the evidence. The time series design, conducted with data at the industry level and with no control group, may not account for the possible effects of other factors affecting the study outcome. Although numerous potential confounding variables were included in the model there remained the possibility of a common underlying factor (e.g., management competency) that could be associated with both a) the more rapid development and report to CSST of a prevention program over time in response to a legislative change and b) productivity growth over time. Additionally, the measure of the prevalence of prevention programs in the

²² *The reviewers questioned the accuracy of R-squared values of some of the reported models (model 3, with three significant variables removed, had a higher R-squared value as compared to model 5 with them included).*

²³ *Annual rate of growth was 0.0032 over the period of the study.*

study (from CSST annual reports), for this review's purposes, is a surrogate for OHSMS implementation and may be limited in terms of validity and accuracy.

4.4.5 Summary of results in studies of mandatory OHSMSs The five studies involving mandatory OHSMSs were conducted in Norway, Quebec, and Ontario. The Lewchuk et al. (1996) and Dufour et al. (1998) studies were based on data before and after the Quebec and Ontario regulations were enacted in 1979. The health and safety regulations that were introduced in Quebec and Ontario were progressive for their time, but in terms of contemporary approaches to integrated health and safety management their requirements were limited. The 1992 Norwegian Internal Control regulation, on the other hand, incorporated systematic management systems which were becoming more common as best practice models in business at that time. Therefore, the two studies of the Quebec and Ontario legislations were based on regulatory efforts that were more limited in the application of OHSMS concepts; whereas, the Norwegian-based studies focused on the IC regulation which was, and is still, highly focused on OHSMSs.

Three of the five studies were conducted on the Norwegian IC regulation and they provided evidence for positive effects across a variety of outcomes. Four years after the IC regulation, 45 per cent of surveyed firms had fully implemented IC (Nytro et al., 1998). At the same point in time, Torp et al. (2000) found significant relationships between measures of IC implementation (four years post IC regulation) in their cross-sectional study and intermediate variables that included satisfaction with HES activities, satisfaction with the physical working environment, HES related management support, and workers' participation in HES activities. Final OHS outcomes were reported in two of the studies. Saksvik and Nytro (1996) reported a significant decline in absenteeism in firms with higher levels of IC implementation, one year after implementation. A decline in accident rates with IC implementation was observed but was not statistically significant. Torp et al. (2000) found a significant negative relationship between IC implementation and musculoskeletal symptoms in their study of garage workers, indicating that firms with higher implementation had fewer workers reporting symptoms. A statistically insignificant negative relationship of IC implementation and sickness absence was also reported.

These three studies of the Norwegian IC regulation, which all showed positive effects across a range of outcomes, were limited in their study design. In two, subjects were selected by randomized quota sampling (Saksvik and Nytro, 1996; Nytro et al., 1998) which is vulnerable to selection bias, but of which the authors did not acknowledge. Selection bias also could not be ruled out in Torp et al. (2000), as their sample included a large number of managers who were scheduled to receive OHS training.

There were also issues in these studies arising from their limited handling of potential confounding factors and their use of self administered questionnaires and self reports from telephone interviews. However, the most crucial limitation for these studies relates to the cross-sectional design which two of them had, making temporal causality difficult to establish.

The two Canadian studies had more sophisticated study designs and used pooled times series data analyses. However, these studies were limited for reasons primarily related to their units of analyses and an inability to account for confounding factors. Dufour et al. (1998) incorporated industry-level data from a variety of publicly available sources and there was concern that the findings may have been affected by factors acting at lower levels. Thus, their finding of a significant positive impact of prevention programs on productivity growth based on aggregate data at the industry level may be subject to ecological fallacy.²⁴ Similarly, the findings of Lewchuk et al. (1996) that Bill 70 led to a significant reduction in work-related injury rates in manufacturing must be interpreted in light of possible co-intervention or confounding effects.

Although all five studies reported positive findings, the overall conclusions are more equivocal. The Norwegian investigations showed effects across a variety of outcomes and over time, indicating that the IC regulation was having some impact. However, the limitations related to study design and sampling leaves open the possibility that the studies had over-selected organizations that were complying with the regulation and managers and workers felt they should report positive effects. The studies of the Ontario and Quebec regulations are limited as the regulations themselves had only a limited focus on OHSMSs and likely used other regulatory actions to impact OHS. The inability of these studies to measure other regulatory activities, and other confounders impacting OHS performance made it difficult to feel confident that the findings reflected only the impact of OHSMSs.

4.5 Evidence synthesis

Sections 4.3 and 4.4 of the report described the individual studies of voluntary and mandatory OHSMSs, and summarized the results for each grouping. This section gives a higher level synthesis of each group of studies and discusses what this means in terms of the level of evidence for OHSMS effectiveness.

One of the original aims of this study was to examine the *relative* effectiveness of voluntary and mandatory OHSMS interventions. The ideal research design for studies addressing that question would involve *direct* comparisons of mandatory and voluntary OHSMS interventions. No studies incorporated such a design.

²⁴ A situation that can occur when a researcher makes an inference about outcomes at a lower level based on aggregate data obtained a higher level of observation.

When a body of literature does not allow a direct comparison, one needs recourse to *indirect* comparisons and this can, under certain conditions, be a valid procedure (Song et al., 2003). In this review, valid indirect comparisons would require the voluntary and mandatory OHSMS studies to have been conducted in similar contexts, with similar populations, outcome measures, and contrasts. Unfortunately, the studies found do not meet these conditions.

In the course of synthesizing evidence on a particular topic, systematic reviews pool the results from homogeneous groups of studies (i.e., homogeneous with respect to study population, intervention, outcome measures, and even research design and context). If studies are very homogeneous and the number of higher quality studies sufficient, a quantitative pooling (meta-analysis) of the measured effects of the interventions is appropriate (van Tulder et al., 2003). Where higher quality studies are sparse and heterogeneous, qualitative pooling should be done.

Among the nine studies included in this systematic review, there is little in common in terms of population, intervention or outcome measures (see Tables 4.3.1 and 4.4.1). Indeed, looking within each of the voluntary and mandatory groups of studies, there is only one instance in which more than one study provided data on the same outcome (workers' compensation premium rates in the Yassi (1998) and Alsop and LeCouteur (1999) studies). The findings of this review, therefore, require a qualitative synthesis.

The following discussion focuses separately on voluntary and mandatory OHSMSs, and describes the levels of available evidence for each.

In systematic reviews, the algorithms used for grading the level of evidence typically involve the following considerations: study design, the quality of the research, the amount of available evidence, and the consistency of the results (i.e., GRADE Working Group 2004; van Tulder et al., 2003; Briss et al., 2000; Franche et al., 2004; Kuhn et al., 1999; Tompa et al., 2004). Although this study did not adopt an explicit algorithm at its outset (for reasons described in section 3.5), similar considerations are used below to characterize the literature.

4.5.1 Evidence for the effectiveness of voluntary OHSMS interventions

The studies on voluntary OHSMS interventions are relatively few, but they all show positive effects – that is, when grouped together, the studies suggest that voluntary OHSMS interventions result in desirable outcomes. These include:

- increased OHSMS implementation over time (as assessed by a validation audit);

- intermediate effects (e.g., better safety climate, increased hazard reporting by employees, more organizational action taken on OHS issues); and
- decreases in workers' compensation premiums.

Although data on injury rates was not included in the review, the findings on workers' compensation premiums, and additional data provided within the studies, imply that injury rates declined too.

The declines in the premium rates would likely be considered of practical importance by stakeholders (declines of 23 and 52 per cent).

Table 4.5.1 Summary of effects from studies of voluntary OHSMSs

Type of Outcome	First Author, Year of Publication (Type of OHSMS)	Effect	Direction of Effect ²⁵	Statistical Significance ²⁶
Implementation (of OHSMS)	Pearse, 2002 (WorkCover)	Avg. increase of 9% in validation audit score among 16 small-medium enterprises (from 53% to 62% in 0-100% scale) in 6-8 mos.	+	n.d.
Intermediate OHS Outcomes (e.g., safety climate)	Edkins, 1998 (own)	Intervention (I): Decrease by 13 points on safety climate scale (0-100 scale) in 8 mos. Control (C): Increase by 6 points.	+	p < 0.001
		I: Decrease by 48 points on employee perception of hazardousness scale (0-176 scale). C: Decrease by 20 points.	+	p < 0.001
		I: Decrease by 32 points of employee perception of hazard likelihood scale (0-176 scale). C: Decrease by 8 points of scale.	+	p < 0.001
		I: 48 confidential hazard reports. C: 9 hazard reports.	+	n.d.
		I: 13 actions taken on identified safety hazards. C: 0	+	n.d.
Economic Outcomes (e.g., workers' compensation premium rates)	Alsop, 1999 (SafetyMap)	52% drop in premium rate (from 4.00% to 1.90% of payroll) in 3 yrs.	+	n.d.
	Yassi, 1998 (own)	23% drop in premium rate (from 1.51% to 1.13% of payroll) in 5 yrs.	+	n.d.

²⁵ Direction of effect is interpreted such that "+" is consistent with more implementation, better OHS outcomes, and better economic outcomes.

²⁶ n.d. = not determined.

However, only one or two studies involving each type of outcome could be found, and all studies in this group were considered to have moderate methodological limitations. These limitations arose largely from the simple study designs that were employed.

In addition to weaknesses of internal validity, three of the four studies involved single workplaces (one with two sites), which makes the direct applicability of the results to other workplaces uncertain. Similar concerns about generalizability arise in the fourth study; it had a 20-workplace sample, but this sample was recruited in the context of a research study in which the refusal rate was high.

The single workplace studies must also be regarded cautiously from the point of reporting bias (favouring 'positive' results). While researchers outside the workplace tend not to publish null findings, workplace representatives who champion interventions (and thus have a vested interest in their success) have an even greater tendency not to publish them. All three reports on single workplace interventions appeared to have been authored by workplace champions. The intervention failure rate seen in this small sample of studies (0 per cent) is markedly different than the rate of 67 to 93 per cent reported for quality management systems (Gardner, 2000). There is no reason to expect the failure rate of OHSMSs to be markedly different than that for quality, since many of the issues identified in the context of quality system implementation are relevant to OHS (e.g., management commitment, culture change).

In sum, there is insufficient evidence in the published, peer-reviewed literature on the effectiveness of voluntary OHSMSs to make recommendations either in favour of or against them.

4.5.2 Evidence for the effectiveness of mandatory OHSMS

interventions The studies on mandatory OHSMS interventions also indicated consistently positive effects. They suggest that mandatory interventions result in:

- increased OHSMS implementation over time;
- intermediate effects (e.g., increased HES awareness; improved employee perceptions of the physical working environment and the psychosocial environment; and increased workers' participation in HES activities);
- decreases in loss-time injury rates; and
- increases in workplace productivity.

The size of the observed changes in OHSMS development and the decline in injury rate observed are likely of practical importance to stakeholders.

Table 4.5.2 Summary of effects from studies of mandatory OHSMSs

Type of Outcome	First Author, Year of Publication (Type of OHSMS)	Effect	Direction of Effect ²⁷	Statistical Significance ²⁸
Implementation (of OHSMS)	Saksvik, 1996 (Internal Control (IC), Norway)	23% of companies perceived improvement in various aspects of OHSMS due to implementation of IC, 1 yr post-intervention (23% is median of results for 4 different aspects of the OHSMS) ²⁹	+	n.d.
	Nytro, 1998 (IC, Norway)	Increase from 8% (at 1 yr post-intervention) to 45% (at 4 yrs) of companies which have fully implemented IC.	+	n.d.
Intermediate OHS Outcomes (e.g., safety climate)	Saksvik, 1996 (IC, Norway)	30% of companies' report increased HES awareness due to implementation of IC, 1 yr post-intervention ²⁹	+	n.d.
	Torp, 2000 (IC, Norway)	Median std $\beta = 0.11$ for 9 intermediate outcomes (incl. psychosocial working environment, physical working environment) regressed on degree of IC implementation, 4 years post-intervention	+	p < 0.05 for 7 out of 9 intermediate outcomes
Final OHS Outcomes (e.g., injury rate)	Lewchuk, 1996 (Bill 70, Ontario)	Regression coefficient, $\beta = -0.015$ for LTI frequency regressed on legislation change. In other words, 18% decrease in LTI from pre- to post-intervention	+	p < 0.001
	Saksvik, 1996 (IC, Norway)	$\beta = 0.09$ for change in absenteeism rate regressed on IC status, 1 yr post-intervention; $\beta = 0.05$ for change in accident rate regressed on IC status	+	P < 0.01 n.s.
	Torp, 2000 (IC, Norway)	Std $\beta = -0.076$ for musculoskeletal symptoms regressed on IC status, 4 yrs post-intervention; Std $\beta = -0.013$ for sick leave regressed on IC status	+	P < 0.01 n.s.
Economic Outcomes (e.g., firm financial benefits)	Dufour, 1998 (LSST, Quebec)	$\beta = 0.154$ for productivity regressed on % of companies with prevention programs. In other words, adoption of prevention programs contributed positively to productivity growth, with the magnitude of the contribution (0.006) being two-fold greater than productivity growth observed for the period studied (0.003).	+	P < 0.05

²⁷ Direction of effect is interpreted such that "+" is consistent with more implementation, better safety, and more financial benefits

²⁸ n.d. = not determined.

²⁹ This figure derived from the study by weighting the results so that they account for the fact that only a sub-sample (n=911) from the entire sample (n=2092) responded to these questions (the non-respondents were those who hadn't started implementing IC at all)

On the other hand, all studies in this group had moderate methodological limitations. These limitations arose largely from the simple study designs employed, especially cross-sectional studies where the direction of causality was uncertain. In addition, there was difficulty eliminating the possibility that there had been confounding or co-intervention in some studies.

As well as these internal validity weaknesses, two of the five studies were concerned with Quebec and Ontario legislation from the late 1970s and are therefore not applicable to most or all Canadian provinces now. The remaining three studies were from Norway, so their applicability to a Canadian jurisdiction is presently uncertain.

In sum, there is insufficient evidence in the published, peer-reviewed literature on the effectiveness of mandatory OHSMSs to make recommendations either in favour of or against them.

4.5.3 Evidence for facilitators of and barriers to OHSMS implementation and effectiveness once implemented Within the scope of the review, **no evidence was found concerning facilitators and barriers of OHSMS implementation or effectiveness that was of sufficient methodological quality.** Little systematic research has been conducted in this area. One study containing quantitative data on this question (Nytro et al., 1998) and three containing qualitative (Chinander, 1998; Lund, 2004; Mitchell, 1998) met the inclusion criteria for the review, but their evidence on facilitators/barriers was judged to be of either Very Low or Low quality (from the perspective of the present review).

4.5.4 Evidence for cost-effectiveness of OHSMS interventions Within the scope of the review, **no studies of the cost-effectiveness of OHSMSs or of the cost of OHSMS implementation were found.** We are aware of only one study with data on costs (Kjellen, 1997) which was a combined quality, environment and OHS intervention. This study was not included in the review because it arrived after the date for inclusion in the study.

5. Discussion

5.1 Identifying and addressing research gaps

The review identified a number of gaps in the research. The most important was the lack of research whose explicit purpose was to study the effectiveness of voluntary and mandatory OHSMSs on employee health and safety and economic outcomes. There was an absence of research focused on the relative effectiveness of these initiatives. Moreover, the studies were seldom sufficiently rigorous methodologically to allow for great confidence in the reported findings. Their limitations also prevent certainty about their applicability to other workplaces.

The following were common limitations in the studies:

- simple research designs (e.g., lack of comparison group, use of cross-sectional designs)
- lack of consideration or control of confounding (through design or statistical adjustments)
- lack of information about the sample (e.g., lack of information about refusal rates or the manner in which sample was drawn from sampling frame)
- convenience samples
- small samples (i.e., several single workplace studies)

The scarcity of high-quality published research on the implementation/effectiveness of OHSMSs may relate to the difficulties in carrying out research under conditions ill-suited to scientific rigour. OHSMSs are – by their nature – complex, broad in scope, and continuously evolving. In addition, the conditions surrounding workplaces and within workplaces are difficult to control and difficult to measure.

Thus, while the randomized control trial design is the gold standard for many clinical effectiveness questions, it is often not feasible for OHS interventions with study units at the workplace- (Cole et al., 2003) or jurisdiction-level. Furthermore, to measure effectiveness, one must carry out resource-intensive new measurement or else resort to using limited, lower-quality administrative data like workers' compensation claims.

There are several reasons why a major intervention such as an OHSMS is extremely challenging to study and evaluate. First, it's often difficult to recruit workplaces. The refusal rate of 67 per cent experienced by Pearse (2002) in the Australian fabricated metal industry is not surprising. Also, those who agree to participate might need to invest a considerable amount of decision-making and resource allocation to support the intervention and evaluation. Finally, coordinating and tracking events in multiple workplaces is a formidable logistic challenge.

When conducting a controlled trial is not feasible, observational cohort studies are considered to be the next best option by those in the epidemiological field. To answer questions about OHSMSs, one would follow a large sample of workplaces (the cohort) over time, measuring the introduction or upgrading of OHSMSs and then measuring outcomes of interest at the workplace-level. Sampling workers within those workplaces would allow a more precise estimate of effects on employees. However, such research designs are very expensive and complex to implement, and as such, are used rarely. One example is the Statistics Canada Workplace and Employee Survey, but to date it has been used primarily to answer questions about organizational factors and productivity.

To answer a question about the relative effectiveness of mandatory and voluntary OHSMS initiatives, one would no doubt need to use research designs other than controlled trials, and would likely need to have jurisdiction as the unit of analysis. Such a study would ideally identify and measure other variables at the jurisdiction-level that could bear upon outcomes. Frick et al. (2000, p. 13) indicated what these might be when they listed the contextual factors that are important when establishing an OHSMS strategy (i.e., industrial relations institutions and traditions, labour market arrangements, policies of governance and general production structure). It would be difficult to sufficiently capture all the important variables at the level of country. One alternative might be analyses of jurisdictions that are smaller than countries, so that some of the social and cultural variables will have less variance across the units of analysis. An example of such a study is the one by Smitha et al. (2001), which compared various voluntary OHS initiatives in the United States, using state as the unit of analysis.

There was a dearth of literature on the topic of facilitators/barriers to OHSMSs. If the subject was mentioned at all, the discussion was brief and any related data was not systematically analyzed. Yet, the context in which these systems are implemented and the characteristics of each particular OHSMS may have an important bearing on success or failure. There is a need for more qualitative research that would explore people's perceptions and experiences of OHSMSs. Such research would enhance understanding of these systems and their contexts. It could identify factors determining the successful implementation and working of OHSMSs.

This review could not address the initial question regarding cost-effectiveness, since no studies contained information on the cost of an OHSMS intervention. This is an important gap. Clearly decisions about whether to adopt an OHSMS at the workplace- or societal-level would be more informed if research that considered cost-effectiveness was available. In two of the three studies that provided evidence on economic outcomes, the evaluation methods used were primitive.

Based on the research gaps that emerged from the literature review, here are several suggestions about how to improve research methodology which would in turn help answer key questions, both about the effectiveness and the cost-effectiveness of OHSMSs:

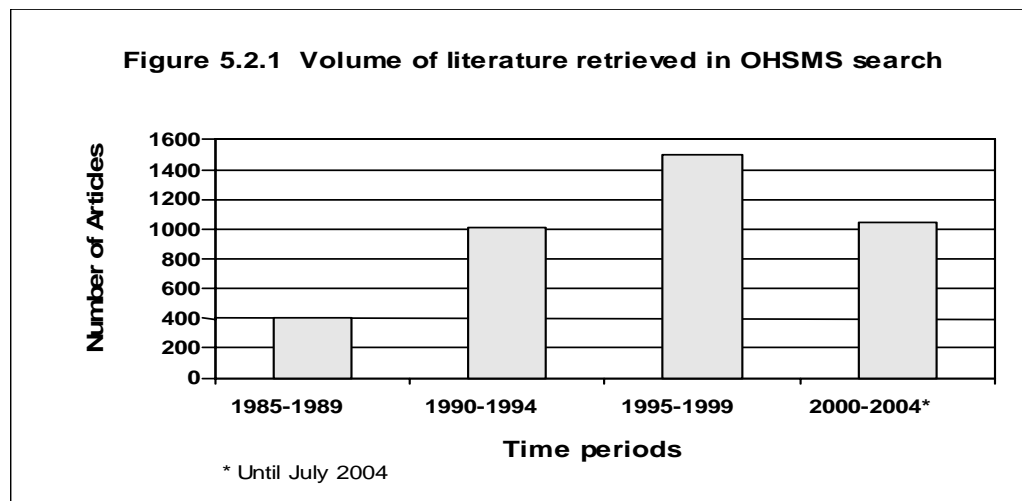
- Use stronger research designs (e.g., avoid cross-sectional designs, use of comparison groups).
- Consider and control confounding to a greater extent (i.e., increase use of control groups, increase measurement of confounders and statistical adjustment, endeavour to track and discount events coincident with the intervention).
- Use more rigorous sampling methods and reporting of those methods (e.g., avoid convenience sampling, report refusal rates).
- Use larger sample sizes, in particular, avoid single worksite studies for questions of effectiveness.
- Design more rigorous economic evaluations.

To address the gaps in research content, the review team recommends that more attention be paid in two areas:

- What are the facilitators of and barriers to OHSMS implementation and effectiveness? In answering this question, researchers should employ both quantitative and qualitative research methods.
- What are the costs of OHSMS interventions?

5.2 Strengths and limitations of the review

5.2.1 Strengths of the review The volume of studies published each year is more than most practitioners or researchers can easily keep track of or synthesize. This is confirmed for the literature on OHSMSs in Figure 5.2.1. The review has clearly eliminated the huge volume of work required to find those relatively few studies of interest and to summarize their findings. The use of explicit, systematic methods ensures that this summary is relatively



objective in its appraisal.

The review's extensive search of the current literature confirmed that no other systematic review has considered the effectiveness of OHSMSs. Until now there have only been high quality narrative reviews available (see section 2.7). The present review therefore makes a unique contribution to the research literature.

The research questions for the review, listed at the beginning of this report, were framed in collaboration with key stakeholders through formal and informal consultations. This increases the relevance of the review results for these stakeholders.

Although the review was necessarily restricted to the peer-reviewed published literature, it drew from a broad range of academic disciplines. The seven databases used represented the disciplines of occupational medicine, occupational safety, risk management, management, occupational psychology, and sociology.

Within the parameters set by the review questions and the included sources, the review team feels confident that the search has been both systematic and reasonably comprehensive and that it is unlikely that there are other items in the peer-reviewed, published literature that would dramatically alter the conclusions of the review.

5.2.2 Limitations of the review Time constraints limited consideration of the evidence to the published, peer-reviewed literature identified in seven academic databases. The usual expectation is that the literature of highest quality is in peer-reviewed journals. A large volume of articles was identified through these databases, despite multiple iterations of the search to enhance its specificity. This large number of articles meant that reviewers could only carry out a preliminary search and screen of other literature, i.e., that which is not peer-reviewed and published. Searches of CCInfoWeb and Dissertation Abstracts International³⁰ were initiated, but were discontinued after the title and abstract screening step, due to time constraints.

The preliminary search of these additional sources suggests that there may be valuable research of the quality required for inclusion in the review published in the form of thesis dissertations and government agency reports. Furthermore, some of these are concerned with OHSMSs not yet reviewed here (e.g., Voluntary Protection and Maine 200 Programs in the U.S., the Management of Health and Safety Regulations in Britain). More research would be needed to determine whether these articles are actually of

³⁰ *Details of these databases are found in Appendix A. In contrast to the academic databases, which primarily abstract journal articles, these sources together yield abstracts from books, book chapters, reports, and thesis dissertations.*

sufficient quality to contribute to the evidence base on OHSMS. Given the limited evidence base generated from the peer-reviewed, published literature, these other sources could have a sizeable impact.

There were a few specific limits placed on the scope of the review. First, only studies that examined a *clearly identifiable* voluntary or mandatory OHSMS *intervention* were included. This excluded an ambitious U.S. study that compared state voluntary initiatives, including prevention programs, while controlling for several other variables (Smitha et al., 2001) and another interesting paper on performance indicators (Simpson and Gardner, 2001). In both cases, the potential OHSMS variable was not clearly defined. Also excluded by this criterion were a few cross-sectional studies investigating the relationship between a researcher-defined measure of OHSMSs and injury outcomes (e.g., Mearns et al., 2003), because there had been no intervention.

The review required study to involve intervention on at least two OHSMS elements, since the review's focus was systems. This excluded a substantial number of cross-sectional studies of the effectiveness of single OHSMS elements such as joint-health-and-safety committees (e.g., LaMontagne et al., 1996; Shannon et al., 1996; Habeck et al., 1998; Cohen, 1977; Simard and Marchand, 1994; Reilly et al., 1995).

6. Conclusions and Recommendations

This systematic review found a relatively small quantity of published, peer-reviewed evidence involving OHSMSs, despite the fact that reviewers screened 4807 studies drawn from seven databases representing diverse disciplines. A qualitative synthesis of the available research was used, because of the small number of studies and the heterogeneity of the data they included.

The review's synthesis of the evidence showed consistently favourable results:

- Studies of voluntary OHSMS interventions reported desirable outcomes, including a more developed OHSMS as assessed by a validation audit, better safety climate, increased hazard reporting by employees, more organizational action taken on OHS issues, and decreases in workers' compensation premiums.
- Studies on mandatory OHSMS interventions also indicated consistently positive effects. These included: a more developed OHSMS; increased HES awareness; improved employee perceptions of the physical working environment and of the psychosocial environment; increased workers' participation in HES activities; decreases in loss-time injury rates; and increases in workplace productivity.

However, all of the studies included in the best-evidence synthesis had moderate limitations regarding their methodologies. The studies were seldom sufficiently rigorous to give great confidence in the reported findings. The most common limitations were: simple research designs (e.g., lack of comparison group, use of cross-sectional designs); lack of consideration or control of confounding; lack of information about the sample; use of small samples and convenience samples. Their limitations also prevent certainty about their applicability to other workplaces.

In conclusion, there is insufficient evidence in the published, peer-reviewed literature on the effectiveness of OHSMSs to make recommendations either in favour of or against OHSMSs. This is not to judge these systems as ineffective or undesirable; it is merely to say that it would be incautious to judge either way in the present state of our research knowledge.

Given the current state of evidence regarding OHSMSs effectiveness:

The review team recommends that those who fund Canadian research should support studies examining the effectiveness or cost-effectiveness of OHSMSs. Support should also be given to research aimed at identifying facilitators of and barriers to OHSMS implementation and effectiveness.

The generalizability and practical application of this research would be greatly enhanced if stronger research designs were used. This would include the use of comparison groups and longitudinal designs. It is also important for researchers to carefully identify and control for potential confounders, to use larger samples selected through random means, and to include more rigorous economic evaluations. Research using both qualitative and quantitative methodologies would be helpful.

The review team recommends that when the topic of OHSMSs is reviewed in future, researchers seek evidence from sources outside the peer-reviewed, published literature.

This review involved an extensive search for research literature on OHSMS interventions. It focused on the published, peer-reviewed literature in order to concentrate on high-quality studies. However, relatively few studies were found and they were only of “moderate” quality. Others who are interested in doing similar research should consider using additional sources of literature.

Because all mandatory and some voluntary OHSMSs are initiated by the government and its agencies, these organizations are a likely source for evaluative reports on the subject. Thesis dissertations might also provide a valuable pool of high-quality studies, since their production involves peer-review.

References

- (1) Systematic Reviews of Strategies to Prevent Occupational Injuries. *American Journal of Preventive Medicine* 2000;18(4S)
- (2) Abad J, Mondelo PR, Llimona J. Towards an international standard on occupational health and safety management. *International Journal of Occupational Safety & Ergonomics*. 2002; 8(3):309-319.
- (3) Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F et al. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. *Spine* 2000; 25(4 Suppl):1S-33S.
- (4) Advisory Committee on the Safety of Nuclear Installations. *Organising for safety: third report*. London: HMSO, 1993.
- (5) Alsop P. Personal communication. 2004.
- (6) Alsop P, LeCouteur M. Measurable success from implementing an integrated OHS management system at Manningham City Council. *Journal of Occupational Health & Safety - Australia & New Zealand*. 1999; 15(6):565-572.
- (7) Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S et al. Grading quality of evidence and strength of recommendations. *British Medical Journal*. 2004; 328(7454):1490.
- (8) Bennett D. Health and safety management systems: Liability or asset? *Journal of Public Health Policy* 2002; 23(2):153-171.
- (9) Brewer GD, deLeon P. *Foundations of policy analysis*. Chicago: Dorsey Press, 1993.
- (10) Briss PA, Zaza S, Pappaioanou M, Fielding J, Wright-De Aguerro L, Truman BI et al. Developing an evidence-based Guide to Community Preventive Services--methods. The Task Force on Community Preventive Services. *Am J Prev Med* 2000; 18(1 Suppl):35-43.
- (11) Burton R. Workplace Safety & Insurance Board. Personal communication. 2004.
- (12) Chemical Industries Association. *Responsible care management systems for health, safety and environment*. 2nd ed. London: Chemical Industries Association, 1995.

- (13) Cohen A. Factors in successful occupational safety programs. *Journal of Safety Research* 1977; 9(4):168-178.
- (14) Cole DC, Rivilis I, Van Eerd D, Cullen K, Irvin E, Kramer D. Effectiveness of Participatory Ergonomic Interventions, a Systematic Review. Toronto: Institute for Work & Health, 2004.
- (15) Cole DC, Wells RP, Frazer MB, Kerr MS, Neumann WP, Laing AC et al. Methodological issues in evaluating workplace interventions to reduce work-related musculoskeletal disorders through mechanical exposure reduction. *Scandinavian Journal of Work, Environment & Health* 2003; 29(5):396-405.
- (16) Commission de la Santé et de la Sécurité du Travail. Annual Report. Quebec: Commission de la Santé et de la Sécurité du Travail, 1987.
- (17) Commission de la Santé et de la Sécurité du Travail. Annual Report. Quebec: Commission de la Santé et de la Sécurité du Travail, 1988.
- (18) Côté P, Cassidy JD, Carroll L, Frank JW, Bombardier C. A systematic review of the prognosis of acute whiplash and a new conceptual framework to synthesize the literature. *Spine* 2001; 26(19):E445-E458.
- (19) Côté P. The association between patterns of visits to general practitioners, chiropractors and specialists and time-to-recovery from acute whiplash injuries: a population-based cohort study. University of Toronto, 2002.
- (20) Dalrymple H, Redinger C, Dyjack D, Levine S, Mansdorf Z. Occupational health and safety management systems : review and analysis of international, national, and regional systems and proposals for a new international document. Geneva, International Labour Organization, 1998.
- (21) Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the *British Medical Journal*. *British Medical Journal*. 1996; 313(7052): 275-283.
- (22) Dufour C, Lanoie P, Patry M. Regulation and productivity. *Journal of Productivity Analysis* 1998; 9(3):233-247.
- (23) Dyjack DT, Levine SP, Holtshouser JL, Schork MA. Comparison of AIHA ISO 9001-based occupational health and safety management system guidance document with a manufacturer's occupational health and safety assessment instrument. *American Industrial Hygiene Association Journal*. 1998; 59(6):419-429.

- (24) Edkins GD. The INDICATE safety program: Evaluation of a method to proactively improve airline safety performance. *Safety Science*. 1998; 30(3);275-295.
- (25) Franche RL, Cullen K, Clarke J, MacEachen E, Frank J, Sinclair S et al. Workplace-based return-to-work interventions: A systematic review of the quantitative and qualitative literature. Toronto: Institute for Work and Health, 2004.
- (26) Frick K, Wren J. Reviewing occupational health and safety management: multiple roots, diverse perspectives and ambiguous outcomes. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000: 17-42.
- (27) Gallagher C. *Occupational health & safety management systems: system types and effectiveness*. Melbourne: Deakin University, 2000.
- (28) Gallagher C, Underhill E, Rimmer M. *Occupational health and safety management systems: a review of their effectiveness in securing healthy and safe workplaces*. Sydney, National Occupational Health and Safety Commission, 2001.
- (29) Gallagher C, Rimmer M. Occupational safety and health management systems in Australia: barriers to success. *Policy and Practice in Health and Safety* 2003; 1(2):67-81.
- (30) Gallagher C, Underhill E, Rimmer M. Review of the effectiveness of OHS management systems in securing health and safe workplaces. Sydney, National Occupational Health and Safety Commission, 2001.
- (31) Gardner D. Barriers to the implementation of management systems: lessons from the past. *Quality Assurance* 2000; 8(1):3-10.
- (32) Gaupset S. The Norwegian internal control reform: an unrealized potential? In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000: 329-350.
- (33) Graham B. Victorian WorkCover Authority. Personal communication. 2004.

- (34) Grayham DA, del Rosario VO. The management of health and safety at work regulations 1992. *Journal of the Royal Society of Health* 1997; 117(1):47-51.
- (35) GRADE Working Group. Grading quality of evidence and strength of recommendations. *British Medical Journal*. 2004; 328:1490-1497.
- (36) Guastello SJ. Some further evaluations of the International Safety Rating System. *Safety Science*. 1991;14(3-4):253-259.
- (37) Habeck RV, Hunt H, van Tol B. Workplace factors associated with preventing and managing work disability. *Rehabilitation Counselling Bulletin* 1998; 42(2):98-143.
- (38) Health and Safety Executive. Successful health and safety management: HSG65. Sudbury, ON: Health and Safety Executive, 1997.
- (39) Hopkins A. Lessons from Longford: the Esso Gas Plant Explosion. Sydney: CCH Australia, 2000.
- (40) International Labour Organization. Meeting of experts on ILO guidelines on occupational safety and health management systems: final report. Geneva: International Labour Organization, 2001. p1-26.
- (41) Kennedy R, Kirwan B. Development of a Hazard and Operability-based method for identifying safety management vulnerabilities in high risk systems. *Safety Science*. 1998; 30(3):249-274.
- (42) Kjellen U, Boe K, Hagen L. Economic effects of implementing internal control of health, safety and environment: a retrospective case study of an aluminium plant. *Safety Science*. 1997; 27(2-3):99-114.
- (43) Kuhn M, Doucet C, Edwards N, Effective Public Health Practice Project. Effectiveness of coalitions in heart health promotion, tobacco use reduction, and injury prevention: a systematic review of the literature 1990-1998. Dundas (ON): Effective Public Health Practice Project. Public Health Branch. Ontario Ministry of Health, 1999.
- (44) LaMontagne AD, Rudd RE, Mangione TW, Kelsey KT. Determinants of the provision of ethylene oxide medical surveillance in Massachusetts hospitals. *Occupational and Environmental Medicine*. 1996;38(2):155-68.

- (45) Lanoie P. The impact of occupational safety and health regulation on the risk of workplace accidents: Quebec, 1983-87. *Journal of Human Resources* 1992; 27(4):643-660.
- (46) Lewchuk W, Robb AL, Walters V. The effectiveness of Bill 70 and joint health and safety committees in reducing injuries in the workplace: the case of Ontario. *Canadian Public Policy* 1996; 22(3):225-243.
- (47) Lindoe PH, Hansen K. Integrating internal control into management systems: a discussion based on norwedgian case studies. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000: 437-456.
- (48) Lund HL. Strategies for sustainable business and the handling of workers' interests: integrated management systems and worker participation. *Economic and Industrial Democracy* 2004; 25(1):41-74.
- (49) Mearns K, Whitaker SM, Flin R. Safety climate, safety management practice and safety performance in offshore environments. *Safety Science*. 2003; 41 (8):641-680.
- (50) Mitchell CS. Evaluating occupational health and safety programs in the public sector. *American Journal of Industrial Medicine* 1998; 34(6):600-606.
- (51) Nichols T, Tucker E. OHS management systems in the UK and Ontario, Canada: a political economy perspective. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000. p. 285-310.
- (52) Nielsen K. Organization theories implicit in various approaches to OHS management. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000. p. 99-124.
- (53) Nytro K, Saksvik P, Torvatn H. Organizational prerequisites for the implementation of systematic health, environment and safety work in enterprises. *Safety Science*. 1998; 30(3):297-307.
- (54) Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology* 1991; 44(11):1271-1278.

- (55) Pearse W. Club zero: implementing OHS management systems in small to medium fabricated metal product companies. In: Pearse W, Gallagher C, Bluff L. Occupational health and safety management systems: proceedings of the first national conference; 2000 Jul; Australia; 2001. p.83-99.
- (56) Pearse W. Club zero: Implementing OHSMS in small to medium fabricated metal product companies. *Journal of Occupational Health & Safety - Australia & New Zealand*. 2002;18(4):347-356.
- (57) Quinlan M, Mayhew C. Precarious employment, work-re-organisation and the fracturing of OHS management. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000. p. 175-198.
- (58) Redinger CF, Levine SP. Development and evaluation of the Michigan Occupational Health and Safety Management System assessment instrument: A universal OHSMS performance measurement tool. *American Industrial Hygiene Association Journal* Vol 59 1998;(8):572-581.
- (59) Reilly B, Paci P, Holl P. Unions, safety committees and workplace injuries. *British Journal of Industrial Relations* 1995; 33(2):275-288.
- (60) Rush B, Ogborne A. Program logic models: expanding their role and structure for program planning and evaluation. *The Canadian Journal of Program Evaluation* 1991; 6(2):95-106.
- (61) Saksvik PO, Nytro K. Implementation of internal control (IC) of health, environment and safety (HES) in Norwegian enterprises. *Safety Science* 1996; 23(1):53-61.
- (62) Saksvik PO, Quinlan M. Regulating systematic occupational health and safety management: comparing the Norwegian and Australian experience. *Relations Industrielles/Industrial Relations* 2003; 58(1):33-59.
- (63) Sculpher M, Fenwick E and Claxton K. Assessing quality in decision analytic cost-effectiveness models: a suggested framework and an example of application. *Pharmacoeconomics*. 2000;17(5):461-77.
- (64) Shannon HS, Walters V, Lewchuck W, Richardson J, Moran LA, Haines T et al. Workplace organizational correlates of lost-time accident rates in manufacturing. *American Journal of Industrial Medicine*. 1996; 29(3):258-268.

- (65) Simard M, Marchand A. The behaviour of first-line supervisors in accident prevention and effectiveness in occupational safety. *Safety Science*. 1994; 17(3):169-185.
- (66) Simpson I, Gardner D. Using OHS positive performance indicators to monitor corporate OHS strategies. *Journal of Occupational Health & Safety - Australia & New Zealand*. 2001;17(2):125-134.
- (67) Slavin RE. Achievement effects of ability grouping in secondary schools: a best-evidence synthesis. *Review of Educational Research* 1990; 60(3):471-499.
- (68) Slavin RE. Best evidence synthesis: an intelligent alternative to meta-analysis. *Journal of Clinical Epidemiology*. 1995; 48(1):9-18.
- (69) Smith B, Chronic pain expert advisory panel. Report of the Chronic Pain Expert Advisory Panel. 2000. Toronto, Workplace Safety & Insurance Board.
- (70) 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):1001-1010.
- (71) Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *British Medical Journal*. 2003; 326(7387):472.
- (72) Spencer L, Ritchie J, Lewis J, Dillon L. Quality in qualitative evaluation: a framework for assessing research evidence. Occasional paper No. 2, prepared by National Centre for Social Researcher's Office. London: Government Chief Social Researcher's Office, 2003. p i-98.
- (73) Thomas J, Sutcliffe K, Harden A, Oakley A, Oliver S, Rees R et al. Children and healthy eating: a systematic review of barriers and facilitators. London, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, 2003.
- (74) Tompa E, Trevithick S, McLeod C. A systematic review of the prevention incentives of insurance and regulatory mechanisms for occupational health and safety. Working Paper #213, Toronto: Institute for Work and Health, 2004.

- (75) Torp S, Riise T, Moen BE. Systematic health, environment and safety activities: do they influence occupational environment, behaviour and health? *Occupational Medicine (Oxford)* 2000; 50(5):326-333.
- (76) Tortorella MJ. The three careers of W. Edwards Deming. *Siam News* 1995 Jul 16. Available from: http://www.deming.org/theman/articles/articles_threecareers01.html. Accessed February 15, 2005.
- (77) United States General Accounting Office. Workplace safety and health: OSHA's voluntary compliance strategies show promising results, but should be fully evaluated before they are expanded. Report to the Chairman, Subcommittee on Workforce Protections, Committee on Education and the Workforce, House of Representatives. United States General Accounting Office: Washington, 2004.
- (78) van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003; 28(12):1290-1299.
- (79) van Tulder MW, Assendelft WJ, Koes BW, Bouter LM. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997; 22(20):2323-2330.
- (80) Victoria Workcover Authority. Available at <http://www.workcover.vic.gov.au> . Accessed December 16, 2004.
- (81) von Bertalanffy L. The theory of open systems in physics and biology. *Science* 1950; 111(2872):23-29.
- (82) Walters D. Regulating health and safety management in the European Union: a study of the dynamics of change. Bruxelles: P.I.E.-Peter Lang, 2002.
- (83) Winder C. Integrating quality, safety, and environment management systems. *Quality Assurance* 1997; 5(1):27-48.
- (84) Wokutch RE, VanSandt CV. OHS management in the United States and Japan: the DuPont and the Toyota models. In: Frick K, Jensen PL, Quinlan M, Wilthagen T, editors. *Systematic occupational health and safety management: perspectives on an international development*. Amsterdam: Pergamon, 2000: 367-390.

- (85) Wright T. IMS - Three into one will go!: The advantages of a single integrated quality, health and safety, and environmental management system. *Quality Assurance Journal*. 2000; 4(3):137-142.
- (86) Yassi A. Utilizing data systems to develop and monitor occupational health programs in a large Canadian hospital. *Methods of Information in Medicine* 1998; 37(2):125-129.
- (87) Zaza S, Wright-De Agüero LK, Briss PA, Truman BI, Hopkins DP, Hennessy MH et al. Data collection instrument and procedure for systematic reviews in the Guide to Community Preventive Services. Task Force on Community Preventive Services. *American Journal Preventive Medicine* 2000; 18(1 Suppl):44-74.

Appendices

List of Appendices

- A.** Details of bibliographic databases
- B.** Search strategy
- C.** Quality appraisal (QA) form
 - C.1** Primary quality assessment (QA) form
 - C.2** Supplementary QA form for qualitative evidence
- D.** Data extraction (DE) form
- E.** Guide to quality assessment and data extraction
- F.** References for included publications
- G.** Characteristics of studies considered in this review.
 - Table **G.1** Study characteristics for studies proceeding to DE
 - Table **G.2** Study characteristics for studies excluded after QA
- H.** Detailed summary of evidence
 - H.1.1 – H.1.9** Data extraction summary tables
 - H.2** Presence of OHSMS elements
- I.** Summary of inclusion criteria

Appendix A - Details of Bibliographic Databases

Medline

MEDLINE (MEDlars onLINE) is the National Library of Medicine's (NLM) premier bibliographic database covering the fields of medicine, nursing, dentistry, veterinary medicine, the health care system, and the preclinical sciences. The MEDLINE file contains bibliographic citations and author abstracts from approximately 3,900 current biomedical journals published in the United States and 70 foreign countries. The file contains approximately 9 million records dating back to 1966. Coverage is worldwide, but most records are from English-language sources or have English abstracts.

http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=2590&T=l

Embase

EMBASE, the Excerpta Medica database, produced by Elsevier Science, is a major biomedical and pharmaceutical database indexing over 3,500 international journals in the following fields: drug research, pharmacology, pharmaceuticals, toxicology, clinical and experimental human medicine, health policy and management, public health, occupational health, environmental health, drug dependence and abuse, psychiatry, forensic medicine, and biomedical engineering/instrumentation. There is selective coverage for nursing, dentistry, veterinary medicine, psychology, and alternative medicine. EMBASE is one of the most widely used biomedical and pharmaceutical databases because of its currency and in-depth indexing. Frequent updates allow access to the latest medical and pharmacological trends. Approximately 375,000 records are added yearly.

http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=14397&T=l

PsycInfo

PsycINFO is an abstract (not full-text) database of psychological literature from the 1800s to the present. An essential tool for researchers, PsycINFO combines a wealth of content with precise indexing so you can get just what you need easily.

http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=12424&T=l

Sociological Abstracts

Sociological Abstracts provides access to the latest worldwide findings in theoretical and applied sociology, social science, and policy science. Produced by Cambridge Scientific Abstracts, Sociological Abstracts features journal citations and abstracts; book, chapter, and association paper

abstracts; and book, film, and software review citations. The database also contains major and minor descriptors. Entries cover sociological aspects of twenty-nine broad topics, including anthropology, business, collective behaviour, community development, disaster studies, education, environmental studies, gender studies, gerontology, law and penology, marriage and family studies, medicine and health, racial interactions, social psychology, social work, sociological theory, stratification, substance abuse, urban studies, and violence. Sociological Abstracts is fundamental for interdisciplinary research in social sciences issues and for practitioners seeking the sociological perspective on various disciplines.

<http://www.ovid.com/site/catalog/DataBase/151.jsp?top=2&mid=3&bottom=7&subsection=10>

CCInfoWeb

CCInfoWeb is an electronic index from the Canadian Centre for Occupational Health and Safety (CCOHS) and was accessed through the University of Toronto Libraries e-resources. Although this index is comprised of several databases, only the OSHLINE, HSELINE and NIOSHTIC-2 databases were searched for this review.

http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=4469&T=I

Dissertation Abstracts International (DAI)

With more than two million entries, the Dissertation Abstracts database is the single, authoritative source for information about doctoral dissertations and master's theses. The database represents the work of authors from over 1,000 graduate schools and universities. We add some 47,000 new dissertations and 12,000 new theses to the database each year.

The database includes bibliographic citations for materials ranging from the first U.S. dissertation, accepted in 1861, to those accepted as recently as last semester. Citations for dissertations published from 1980 forward also include 350-word abstracts written by the author. Citations for master's theses from 1988 forward include 150-word abstracts. The full text of more than 1.7 million of these titles is available in paper and microform formats. Institutional subscribers to ProQuest Digital Dissertations receive on-line access to the complete file of dissertations in digital format starting with titles published from 1997 forward.

<http://www.il.proquest.com/umi/dissertations/individuals.shtml>

Safety Science & Risk Abstracts

Published in association with the University of Southern California and the University of Waterloo, this database provides comprehensive, timely information across the fields of public health, safety, and industrial hygiene. Cited studies are geared to help researchers identify, evaluate, and eliminate or control risks and hazards from environmental and occupational situations.

http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=2579&T=l

EconLit: Economic Literature Index

EconLit, the fundamental research tool in economics, provides bibliographic citations, with selected abstracts, to the international literature on economics since 1969.

EconLit covers a broad range of document types published world-wide, including journal articles, books, and dissertations, as well as articles in collective works, such as conference proceedings and collected essay volumes. The database also includes Abstracts of Working Papers in Economics from the Cambridge University Press database, Index of Economic Articles in Journals & Collective Volumes and the full text of the Journal of Economic Literature book reviews.

EconLit topics include economic development, forecasting, and history; fiscal theory; monetary theory and financial institutions; business finance; public finance; and international, labour, health care, managerial, demographic, regional, agricultural, and urban economics; country studies, and government regulations.

<http://www.ovid.com/site/catalog/DataBase/52.jsp?top=2&mid=3&bottom=7&subsection=10>

ABI/INFORM Global

One of the world's first electronic databases, ABI/INFORM has been a premier source of business information for more than 30 years. The database contains content from thousands of journals that help researchers track business conditions, trends, management techniques, corporate strategies, and industry-specific topics worldwide.

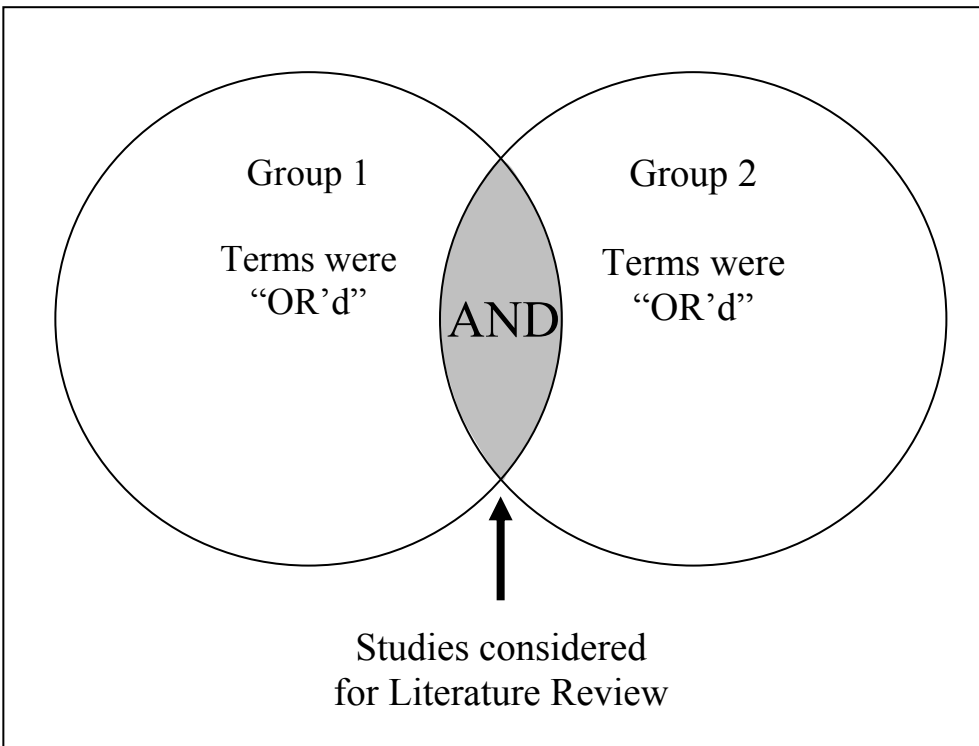
http://link.library.utoronto.ca/eir/EIRdetail.cfm?Resources_ID=4958&T=l

Appendix B - Search Strategy

Table B.1 Search Terms

Group 1	Group 2:
<ul style="list-style-type: none"> ▪ Health and safety management system(s) ▪ Safety and health management system(s) ▪ Systematic occupational health and safety management ▪ Occupational health management ▪ Safety management system(s) ▪ OHS program(s) ▪ OHS system(s) ▪ OHS management ▪ OHS legislation ▪ OHSAS 18001 ▪ OHSMS ▪ BS8800 ▪ Safety map ▪ 5 star/Five star ▪ International Safety Rating System(s) ▪ AS/NZS 4804 ▪ NR9 ▪ Respositble Care Management System(s) ▪ OSHA Voluntary Protection Program(s) ▪ OSHA VPP ▪ Framework directive ▪ Internal Control ▪ Working Environment Act ▪ Safety program(s) ▪ Safety system(s) ▪ Occupational health standard(s) ▪ Occupational safety standard(s) ▪ Occupational health guideline(s) ▪ Occupational safety guideline(s) ▪ Safety and health legislation 	<ul style="list-style-type: none"> ▪ Evaluat(ing/ion/ions/e/es/ed) ▪ Program evaluat(ing/ion/ions/e/es/ed) ▪ Implementat(ing/ion/ed/es) ▪ Management ▪ Self inspection(s) ▪ Operational audit(s) ▪ Climate ▪ Culture ▪ Cost-benefit analysis ▪ Effect(s) ▪ Impact(s) ▪ Knowledge ▪ Belief/believ(e/es) ▪ Value(s) ▪ Perception(s) ▪ Behavio(u)r ▪ Injury experience ▪ Product quality ▪ Property damage ▪ Quality of life ▪ Workman Compensation ▪ Work capacity ▪ Work limitation ▪ Workplace injur(y/ies) ▪ Injury ▪ Occupational health ▪ Occupational exposure ▪ Death ▪ Accident prevention ▪ Occupational accident ▪ Compensation cost(s) ▪ Compensation claims cost(s) ▪ Time loss/lost ▪ Lost workday(s) ▪ Wage replacement

Figure B.1 Search Strategy



Appendix C - Quality Appraisal (QA) Forms

C.1 Primary Quality Assessment (QA) Form

Level 4: Quality Assessment and Data Extraction

1. **What type of OHSMS intervention is being studied?**
 - Mandatory OHSMS (e.g., legislatively dictated)
 - Voluntary OHSMS (OHSAS 18001, ILO, BS8800, sector standards, proprietary OHSMSs)
 - Unclear/unknown from information provided

2. **How many work sites comprise the study sample?**
 - One work site
 - Between 2 and 19 work sites
 - 20+ work sites
 - Unclear/unknown from information provided

3. **Of what size of enterprise was the study sample representative?**
 - Small and/or medium enterprises (i.e., < about 100 to 200 employees)
 - Large enterprises
 - Mixed sizes
 - Unclear/unknown from information provided

4. **From what NAICS industrial sector was the study sample drawn?**

*Check all that apply up to four, then default to multiple sector option.
See*

<http://www.statcan.ca/english/Subjects/Standard/naics/2002/naics02-menu.htm>

 - Agriculture, Forestry, Fishing and Hunting (11)
 - Mining and Oil and Gas Extraction (21)
 - Utilities (22)
 - Construction (23)
 - Manufacturing (31-33)
 - Wholesale Trade (41)
 - Retail Trade (44-45)
 - Transportation and Warehousing (48-49)
 - Professional, Scientific and Technical Services (54)
 - Waste Management and Remediation Services (562)
 - Educational Services (61)
 - Health Care and Social Assistance (62)
 - Accommodation and Food Services (72)
 - Other Services (except Public Administration) (81)
 - Public Administration (91)
 - Other NAICS categories not mentioned above (51, 52, 53, 55, 561, 71)

- Multiple sectors (representative of workplace population)
 - Multiple sectors (five or more of the sectors mentioned above)
 - Unclear/unknown from information provided.
- 5. Does the publication give information on implementation/effectiveness or facilitators/barriers?**
- Implementation and/or effectiveness (I/E)
 - Facilitators and/or barriers (F/B)
 - Both implementation/effectiveness and facilitators/barriers
- 6. Does this publication give evidence relevant for our core search questions on the following? (check all that apply)**
- Implementation - quantitative
 - Intermediate OHS outcomes (e.g., climate, knowledge, attitudes, safety hazards, employee safety behaviours) - quantitative
 - Final OHS outcomes (e.g., injuries, claim rates) - quantitative
 - Cost of the intervention - quantitative
 - Financial benefits of the intervention (e.g., claim costs) – quantitative
 - Facilitators/barriers – quantitative
 - Facilitators/barriers – qualitative
 - Facilitators/barriers – non-systematic
- 7. What type of experimental design was used in the I/E and F/B portions of the study? (see algorithm³¹)**
- Non-comparative quantitative study
 - Cross-sectional
 - Before-after
 - Time series
 - Case control
 - Cohort study
 - Other designs with concurrent comparison groups
 - Randomized trial
 - Group randomized trial
 - Non-randomized trial
 - Qualitative
 - Unclear/unknown from information provided
- 8. Are you confident that the means of selecting and maintaining the sample minimized bias in the I/E portion of the study?**
- Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)
 - Unclear/unknown from information provided

³¹ Adapted from Zaza et al. (2000)

9. **Are you confident that potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias in the I/E portion of the study?**
- Yes (*Explain*)
 - Partially (some control evident) (*Explain*)
 - No (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
 - Reviewer not qualified to answer
10. **Was the degree of implementation of the OHSMS demonstrated or described clearly?**
- Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)
11. **Are you confident that the measurement methods did not introduce bias to the corresponding implementation findings?**
- Not applicable, no implementation evidence
 - Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)
12. **Were appropriate statistical tests conducted on the implementation data?**
- Not applicable, no implementation evidence
 - Yes (*Explain*)
 - Minor deficiencies, which will have little or no effect on the conclusions (*Explain*)
 - Inappropriate statistical tests, which might modify the conclusions (*Explain*)
 - Inappropriate statistical tests, which could substantially effect the conclusions (*Explain*)
 - No statistical tests conducted (and not needed). (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
 - Reviewer not qualified to answer
13. **Are you confident that the measurement methods did not introduce bias to the corresponding intermediate outcome(s) findings?**
- Not applicable, no intermediate outcomes evidence
 - Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)

14. Were appropriate statistical tests conducted on the intermediate outcome data?

- Not applicable, no intermediate outcomes evidence
- Yes (*Explain*)
- Minor deficiencies, which will have little or no effect on the conclusions (*Explain*)
- Inappropriate statistical tests, which might modify the conclusions (*Explain*)
- Inappropriate statistical tests, which could substantially affect conclusions (*Explain*)
- No statistical tests conducted (and not needed). (*Explain*)
- Unclear/unknown from information provided (*Explain*)
- Reviewer not qualified to answer

15. Are you confident that the measurement methods did not introduce bias to the corresponding final outcome(s) findings?

- Not applicable, no final outcomes evidence
- Yes (*Explain*)
- Partially (*Explain*)
- No (*Explain*)

16. Were appropriate statistical tests conducted on the final outcome data?

- Not applicable, no final outcomes evidence
- Yes (*Explain*)
- Minor deficiencies, which will have little or no effect on the conclusions (*Explain*)
- Inappropriate statistical tests, which might modify the conclusions (*Explain*)
- Inappropriate statistical tests, which could substantially affect conclusions (*Explain*)
- No statistical tests conducted (and not needed). (*Explain*)
- Unclear/unknown from information provided (*Explain*)
- Reviewer not qualified to answer

17. Are you confident that the measurement methods did not introduce bias to the corresponding financial benefits findings?

- Not applicable, no financial benefits evidence
- Yes, (*Explain*)
- Partially (*Explain*)
- No (*Explain*)

- 18. Were appropriate statistical tests conducted on the financial benefit data?**
- Not applicable, no financial benefits evidence
 - Yes (*Explain*)
 - Minor deficiencies, which will have little or no effect on the conclusions (*Explain*)
 - Inappropriate statistical tests, which might modify the conclusions (*Explain*)
 - Inappropriate statistical tests, which could substantially affect the conclusions (*Explain*)
 - No statistical tests conducted (and not needed). (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
 - Reviewer not qualified to answer
- 19. Are you confident that there are no additional potential sources of bias in the estimate of implementation/effectiveness not yet captured in the previous questions?**
- Yes, no additional potential sources of bias are likely
 - No, there are additional potential source(s) of bias of some concern (*Explain*)
 - No, there are additional potential source(s) of bias of major concern (*Explain*)
- 20. What is your overall appraisal of the quality of evidence about implementation provided by this study?**
- Not applicable, no implementation evidence
 - High (no or minor limitations)
 - Moderate (moderate limitations)
 - Low (major limitations)
 - Very low (serious limitations)
- 21. What is your overall appraisal of the quality of evidence about intermediate outcomes provided by this study?**
- Not applicable, no intermediate outcomes evidence
 - High (no or minor limitations)
 - Moderate (moderate limitations)
 - Low (major limitations)
 - Very low (serious limitations)
- 22. What is your overall appraisal of the quality of evidence about final outcome(s) provided by this study?**
- Not applicable, no final outcome(s) measured
 - High (no or minor limitations)
 - Moderate (moderate limitations)
 - Low (major limitations)
 - Very low (serious limitations)

- 23. What is your overall appraisal of the quality of evidence about financial benefit(s) provided by this study?**
- Not applicable, no financial benefit(s) evidence
 - High (no or minor limitations)
 - Moderate (moderate limitations)
 - Low (major limitations)
 - Very low (serious limitations)
- 24. If qualitative research methods were used to identify facilitators and barriers, how do you rate the quality of those methods?**
- Not applicable, qualitative research methods not used
 - High (no or minor limitations) (*Explain*)
 - Moderate (moderate limitations) (*Explain*)
 - Low (major limitations) (*Explain*)
 - Very low (serious limitations) (*Explain*)
 - Reviewer not qualified to answer
- 25. Are you confident that the means of selecting and maintaining the sample used in quantitative F/B study minimized bias?**
- Not applicable, quantitative research methods not used in F/B study.
 - Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
- 26. Are you confident that (in the quantitative study of F/Bs) potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?**
- Not applicable, quantitative research methods not used in F/B study.
 - Yes (*Explain*)
 - Partially (some control evident) (*Explain*)
 - No (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
 - Reviewer not qualified to answer
- 27. Are you confident the measurement methods did not introduce bias to the corresponding F/B findings?**
- Not applicable, quantitative research methods not used in F/B study.
 - Yes (*Explain*)
 - Partially (*Explain*)
 - No (*Explain*)

- 28. Were appropriate statistical tests conducted on the F/B data?**
- Not applicable, quantitative research methods not used in F/B study.
 - Yes (*Explain*)
 - Minor deficiencies only (*Explain*)
 - Inappropriate statistical tests, which might modify the conclusions (*Explain*)
 - Inappropriate statistical tests, which could substantially affect conclusions (*Explain*)
 - No statistical tests conducted (and not needed). (*Explain*)
 - Unclear/unknown from information provided (*Explain*)
 - Reviewer not qualified to answer
- 29. Are you confident that there are no additional potential sources of bias in the quantitative study of F/Bs not yet captured in the previous questions?**
- Not applicable, quantitative research methods not used in F/B study
 - Yes, no additional potential sources of bias are likely
 - No, there are additional potential source(s) of bias of some concern (*Explain*)
 - No, there are additional potential source(s) of bias of major concern (*Explain*)
- 30. What is your overall appraisal of the quality of evidence about facilitators and barriers provided by the quantitative study?**
- Not applicable, quantitative research methods not used in F/B study.
 - High (no or minor limitations)
 - Moderate (moderate limitations)
 - Low (major limitations)
 - Very low (serious limitations)
- 31. Should this publication go on to data extraction (Level 5)?**
- Yes
 - No
- 32. Is supplementary information required?**
- Yes
 - No
- 33. Is this the final version of the Level 4 assessment?**
- Yes
 - No

Appendix C.2 - Supplementary QA Form for Qualitative Evidence³²

1. How credible are the findings?
2. How has knowledge/understanding been extended by the research?
3. How well does the study address the original aims and purpose?
4. How well is the scope for drawing wider inference explained?
5. How defensible is the research design?
6. How well defended is the sample design/target selection of cases?
7. Sample composition/case inclusion - how well is coverage described?
8. How well was the data collection carried out?
9. How well was the approach to/formulation of the analysis conveyed?
10. How well are the contexts of data sources retained/portrayed?
11. How well has diversity of perspective and content been explored?
12. How well has detail, depth and richness of data been conveyed?
13. How clear are the links between data, interpretation and conclusions?
14. How clear and coherent is the reporting?
15. How clear are the assumptions/theoretical perspectives/values that shaped the form and output of the study?
16. What evidence is there of attention to ethical issues?
17. How adequately has the research process been documented?

³² The framework is provided in more detail in Spencer et al. (2003)

Appendix D - Data Extraction (DE) Form

The following questions are included for generating the evidence table. Keep this in mind when deciding how much information to include. If the information is different between the implementation/effectiveness and facilitators/barriers portions of the study, please provide both sets of information and indicate which portion they apply.

- 1. What was the research question?**
- 2. Describe the study sample and sampling frame, including the sample size, type of industry, size of workplace(s), types of occupations (if included), number of workplaces, and the method of sampling.**
- 3. Give any information pertaining to maintenance of the sample (i.e., withdrawals, etc.).** *(Put N/A if not applicable)*
- 4. Describe the intervention in detail.**
- 5. Describe the experimental design in detail.**
- 6. Describe relevant data collection, data transformation, and measurement properties under the applicable evidence category.**
(Not applicable, Implementation, Intermediate outcome, Final outcomes, Financial benefits, Facilitators/Barriers, Other)
- 7. Describe statistical tests used.**
- 8. Describe all findings regarding implementation.**
- 9. Describe all findings regarding intermediate OHS outcomes.**
- 10. Describe all findings regarding final OHS outcomes.**
- 11. Describe all findings regarding cost of the intervention.** *(Put N/A if not applicable)*
- 12. Describe all findings regarding the financial benefit of the intervention.**
- 13. Did the design lack statistical power?**
(Not applicable, Yes-explain, No-explain, Reviewer not qualified to answer)
- 14. List findings regarding facilitators and barriers.**

15. Were any harms of the intervention identified?

16. Provide any additional comments that you think should be included in the evidence table about this study? (Put N/A if not applicable)

Please indicate the presence of the following OHSMS elements, using the Dalrymple et al. (1998) scheme. Use the following code to indicate the presence/absence of an element: “X”, present, “/”, partially present, “-“, not present, “u”, unknown.

17. Management commitment and resources

18. Regulatory compliance and OHSMS conformance

19. Accountability, responsibility, and authority

20. Employee participation

21. Occupational health and safety policy

22. Goals and objectives

23. Performance measures

24. System planning and development

25. Baseline evaluation and hazard/risk assessment

26. OHSMS manual and procedures

27. Training system

28. Technical expertise and personnel qualifications

29. Hazard control system

30. Process design

31. Emergency response

32. Hazardous agent management

33. Preventive and corrective actions

34. Procurement and contractor selection

- 35. Communication system**
- 36. Document and record management system**
- 37. Evaluation system**
- 38. Auditing and self-inspection**
- 39. Incident investigation and root cause analysis**
- 40. Medical program and surveillance**
- 41. Continual improvement**
- 42. Integration**
- 43. Management review**
- 44. Is supplementary information required?**
- 45. Is this the final version?**

Appendix E - Guide to Quality Assessment and Data Extraction

Introduction

Quality Assessment

These forms aim to i) extract data with multiple choice questions to allow simple characterization of the group of studies ii) extract data with open-ended questions to allow the generation of detailed evidence table(s) and iii) appraise the quality of the data in terms of its internal validity.³³

1. What type of OHSMS intervention is being studied?

- Mandatory OHSMS interventions are those where the OHSMS is required by law. Voluntary OHSMS interventions include those available from commercial sources and government sources, providing their adoption is voluntary.

2. How many work sites comprise the study sample?

- For this question, report on the number of workplaces observed at the start of the study. (A larger sample might have been drawn, and a smaller sample decided to participate. For this question, indicate the latter number).
- If the facilitators and barriers (F/B) portion of a study is different than the implementation/effectiveness (I/E) portion of the study regarding the sample size, then answer the multiple choice question with the I/E study in mind and note any differences for the F/B study in the text box. If this isn't noted, the assumption will be that the samples are the same.

3. Of what size of enterprise was the study sample representative?

- The cutoff between medium and large varies among jurisdictions and sectors; go by the authors' definition
- If the facilitators and barriers (F/B) portion of a study is different than the implementation/effectiveness (I/E) portion of the study regarding the sample size, then answer the multiple choice question with the I/E study in mind and note any differences for the F/B study

³³ *Internal validity: "The index and comparison groups are selected and compared in such a manner that the observed differences between them on the dependent variables under study may, apart from sampling error, be attributed only to the hypothesized effect under investigation." This definition and others used later are taken from Last JM (ed.) A Dictionary of Epidemiology, 4th ed., 2001.*

in the text box. If this isn't noted, the assumption will be that the samples are the same.

4. From what NAICS industrial sector was the study sample drawn?

- Check all that apply up to four, then default to multiple sector option. See <http://www.statcan.ca/english/Subjects/Standard/naics/2002/naics02-menu.htm> for more information that will assist you in determining the sector.
- If the facilitators and barriers (F/B) portion of a study is different than the implementation/effectiveness (I/E) portion of the study regarding the sample size, then answer the multiple choice question with the I/E study in mind and note any differences for the F/B study in the text box. If this isn't noted, the assumption will be that the samples are the same.

5. Does the publication give information on implementation/effectiveness or facilitators/barriers?

- This question determines the sequence of questions the reviewer will see in the form after question 6. Answer I/E if the publication includes I/E information that meets our inclusion criteria. Answer F/B if the publication includes F/B information collected through systematic means OR is an I/E study with F/B information collected by any means. Subsequent questions will determine the nature of the F/B information and then focus on studies with F/B information derived through research methods.

6. Does this publication give evidence relevant for our core research question on the following?

- Check all responses that apply
- Check an answer only if complete information is provided, such that the corresponding type of evidence meets the inclusion criteria. For example, a publication might give information on implementation, but you would select "implementation" as an answer only if information provided was complete enough that the inclusion criteria for evidence on implementation was met.

Questions 7-23 are concerned with the implementation/effectiveness (I/E) aspects of the study.

7. What type of experimental design was used in the I/E and F/B portions of the study?

- See algorithm adapted from Zaza et al. (2000; Figure E.1)
- Choose the one design that best fits the I/E portion of the study.
- If your study has both I/E and F/B portions, then select the appropriate design(s) and indicate to which portion of the study the design is applicable with the use of “I/E” and “F/B” in the accompanying text box. Include both I/E and F/B in the same text box, if both portions of the study use the same design.

8. Are you confident that the means of selecting and maintaining the sample minimized bias in the I/E portion of the study?

- In the comment box following your response, note the rationale for your selection.
- This question addresses sampling bias, selection bias, bias due to withdrawals, selection threat to internal validity,³⁴ or any other bias arising from the selection and maintenance of the sample
 - Sampling bias: systematic error due to study of a non-random sample of a population
 - Selection bias: error due to systematic differences in characteristics between those who take part in a study and those who do not
 - Bias due to withdrawals: A difference between the true value and that actually observed in a study due to the characteristics of those subjects who choose to withdraw.
 - Selection threat to internal validity: A threat to internal validity arising from differing characteristics of subjects in the intervention and comparison groups

Intervention-only studies

- Before-after and time-series-without-comparison-group designs are unlikely to receive a response of “yes”, because the workplace selected for the study is typically not randomly selected
- Studies with samples comprised of volunteers are unlikely to receive a response of “yes” because volunteers are typically different than the target population.

³⁴ This term is from the key reference on quasi-experimental designs: Cook TD, Campbell DT (1979) *Quasi-experimentation design & analysis issues for field settings*. Boston: Houghton Mifflin. It also appears in the updated chapter: Cook TD et al. (1990) *Quasi experimentation In: Dunnette MD, Hough LM (eds.) Handbook of Industrial and Organizational Psychology, 2nd ed.* Palo Alto, Ca.: Consulting Psychologists Press.

- Studies with a large refusal rate (low participation rate) are unlikely to receive a “yes”.

Studies with comparison groups

- Consider whether blinding was used in the allocation of workplaces to intervention/control groups
- Differential workplace refusal rates and differential workplace dropout rates in designs with comparison/control groups are indicators of bias arising from sample selection and maintenance.
- Whenever groups have been created through non-random means, there is a threat to internal validity. If groups were selected through a matching procedure, the threat would be less.

9. Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias in the I/E portion of the study?

- In the comment box following your response, note the rationale for your selection.
- This question addresses confounding bias: distortion of the estimated effect of an exposure on an outcome, caused by the presence of an extraneous factor associated both with the exposure (i.e., intervention) and the outcome, but is not a mediator between exposure and outcome.
- Before-after and time-series designs are unlikely to receive a “yes” rating, because they are always susceptible to “history bias” (another term from the quasi-experimentation literature), which occurs when something else takes place in the workplace or extra-workplace environment that could bring about or alter the effect observed. A “partially” rating can be achieved, if efforts are taken to assure the reader that the observed effect can not be explained by another event taking place in the organization simultaneously.
- Demonstration of similar distribution of known confounders among comparison groups is required for a “yes” when groups are created through non-random means. If created through random means, this demonstration is important when there are small numbers in the comparison groups.
- We are concerned in this question about all potential sources of confounding, whether they arise through the sample, coincident events, differential measurement, etc.
- Control of confounders can be either through design or statistical analysis.
- Important confounders at the workplace level: management commitment to OHS, size of enterprise, other organizational change initiatives that could affect implementation/effectiveness, changes in the process technology

- Important confounders at the extra-workplace level: other extra-workplace OHS initiatives.

10. Was the degree of implementation of the OHSMS demonstrated or described clearly?

- The reason for including this question is to aid in our interpretation of studies looking at outcomes. We are looking for a description or demonstration of the degree of implementation that will give us confidence that any observed lack of effect on outcomes does not result from a lack of intervention implementation.
- Note that this question is concerned with what was actually done (i.e., what was implemented), as opposed to what was planned.

11. Are you confident that the measurement methods did not introduce bias to the corresponding implementation findings?

- In the comment box following your response, note the rationale for your selection.
- This question addresses measurement bias; i.e., systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s).
- Consider the validity of the measurement method; consider reliability to the extent that could bias findings. Some examples.
 - Consider whether blinding was used in the assessment of implementation. For example, was the person who carried out the management audit unaware of whether a workplace was in an intervention? Did the person have a vested interest in a certain set of results?
 - Consider inter-rater reliability of and how audits were distributed among raters. (i.e. could observed effect result from an audit with poor IRR and uneven distribution of auditors between intervention and control groups/)
- Consider the measurement methods independent of the sample issues. For example, in the Nytro et al. study, there was concern that the methods used for implementation might have created a bias in the direction of better compliance because company managers would have an incentive to appear as though they are complying with the law. If we had been told that the self-report method had been verified in a sub-sample through an external audit, we could have given a high rating for the measurement methods, even though the means of selecting the sample (quota sampling; no indication of refusals) could nevertheless have biased the results towards more compliance.

- Consider all measurement methods pertinent to the corresponding findings (not just the implementation variables). For example, in the case of the intermediate and final outcomes findings, the method of measuring OHSMS variables might also be relevant.

12. Were appropriate statistical tests conducted on the implementation data?

- In the comment box following your response, note the rationale for your selection.
- The issue of whether adjustment for confounders took place is covered in Q9. In this question, one is only concerned about whether such adjustments were done correctly, if they were done
- Do not fear choosing the option of “reviewer not qualified to answer.” We are simply using this response to flag where we need expertise from elsewhere in the group or even external to the group. If two reviewers have selected this option, then we seek expertise from elsewhere in the group, and failing this, use our external consult. If only one reviewer has selected this option, then a consensus answer can be formulated, if the remaining reviewer is highly confident of his/her answer. If not, then expertise should be sought elsewhere in the group or externally if required.
- Was the test appropriate for the sample size and type of data?
- If no statistical tests were used, and you think they were not needed, select the answer corresponding to this. If no statistical tests were used and you think they should have been, then use one of the following three options: minor deficiencies; inappropriate statistical test, which might modify the conclusions; inappropriate statistical tests, which could substantially affect conclusions.
- Consider issues of statistical power in Q43 not here.

13. Are you confident that the measurement methods did not introduce bias to the corresponding intermediate outcome(s) findings?

- In the comment box following your response, note the rationale for your selection.
- This question addresses measurement bias; i.e., systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s).
- Consider the validity of the measurement method; consider reliability to the extent that could bias findings. Examples:
 - The introduction of an OHSMS could alter reporting practices, so this potential source of bias should be eliminated by some means
 - Consider whether blinding was used in the assessment of intermediate outcome(s); e.g., was the person who observed

employee behaviour blind to the knowledge of intervention status

- Consider the measurement methods independent of the sample issues. For example, in the Nytro et al. study, there was concern that the methods used for implementation might have created a bias in the direction of better compliance because company managers would have an incentive to appear as though they are complying with the law. If we had been told that the self-report method had been verified in a sub-sample through an external audit, we could have given a high rating for the measurement methods, even though the means of selecting the sample (quota sampling; no indication of refusals) could nevertheless have biased the results towards more compliance.
- Consider all measurement methods pertinent to the corresponding findings (not just the intermediate outcome(s) variables). For example, in the case of the intermediate and final outcomes findings, the method of measuring OHSMS variables might also be relevant.

14. Were appropriate statistical tests conducted on the intermediate outcome data?

- In the comment box following your response, note the rationale for your selection.
- The issue of whether adjustment for confounders took place is covered in Q9. In this question, one is only concerned about whether such adjustments were done correctly, if they were done
- Do not fear choosing the option of “reviewer not qualified to answer.” We are simply using this response to flag where we need expertise from elsewhere in the group or even external to the group.
- Was the test appropriate for the sample size and type of data?
- If no statistical tests were used, and you think they were not needed, select the answer corresponding to this. If no statistical tests were used and you think they should have been, then use one of the following three options: minor deficiencies; inappropriate statistical test, which might modify the conclusions; inappropriate statistical tests, which could substantially affect conclusions.
- Consider issues of statistical power in Q43 not here.

15. Are you confident that the measurement methods did not introduce bias to the corresponding final outcome(s) findings?

- In the comment box following your response, note the rationale for your selection.
- This question addresses measurement bias; i.e., systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s).

- Consider the validity of the measurement method; consider reliability to the extent that could bias findings. Examples:
 - Consider whether blinding was used in the assessment of final outcome(s)
 - The introduction of an OHSMS could alter reporting practices, so this potential source of bias should be eliminated by some means (e.g., confirm that there is a constant ratio of major to minor injuries over time or assess change through more serious injuries (since changes in reporting are more likely to affect less serious injuries); verification of constant reporting practices by worker survey)
 - Injury measures should either i) adjust for exposure (e.g., have a denominator of FTEs or hours worked) or ii) confirm that exposure is constant if not adjusted.
- Consider the measurement methods independent of the sample issues. For example, in the Nytro et al. study, there was concern that the methods used for implementation might have created a bias in the direction of better compliance because company managers would have an incentive to appear as though they are complying with the law. If we had been told that the self-report method had been verified in a sub-sample through an external audit, we could have given a high rating for the measurement methods, even though the means of selecting the sample (quota sampling; no indication of refusals) could nevertheless have biased the results towards more compliance.
- Consider all measurement methods pertinent to the corresponding findings (not just the final outcome(s) variables). For example, in the case of the intermediate and final outcomes findings, the method of measuring OHSMS variables might also be relevant.

16. Were appropriate statistical tests conducted on the final outcome data?

- In the comment box following your response, note the rationale for your selection.
- The issue of whether adjustment for confounders took place is covered in Q9. In this question, one is only concerned about whether such adjustments were done correctly, if they were done
- Do not fear choosing the option of “reviewer not qualified to answer.” We are simply using this response to flag where we need expertise from elsewhere in the group or even external to the group.
- Was the test appropriate for the sample size and type of data? Some guidance is reproduced from material drafted by Harry Shannon,³⁵

³⁵ Robson LS, Shannon HS, Goldenhar LM, Hale AR. *Guide to evaluating the effectiveness of strategies for preventing work injuries: How to show whether a safety intervention really works.* DHHS (NIOSH) Publication No. 2001-119, 2001.

with Table 8.1 giving guidance on choice of test (though these aren't the only possibilities) and Appendix B illustrating some of those tests. Some cautions appear in the text:

- When numbers of injuries are less than five, “exact” methods should be used (p. 103)
- When comparing the change in rates between an intervention and control group, statistical tests should not be carried out on each group separately. Instead, the test should be carried out on the difference in rate changes (pp. 105-107)
- If groups are being compared and initial injury rates are quite different, adjustment needs to be made, e.g. through multiple regression (p. 107)
- If no statistical tests were used, and you think they were not needed, select the answer corresponding to this. If no statistical tests were used and you think they should have been, then use one of the following three options: minor deficiencies; inappropriate statistical test, which might modify the conclusions; inappropriate statistical tests, which could substantially affect conclusions.
- Consider issues of statistical power in Q43 not here.

17. Are you confident that measurement methods did not introduce bias to the corresponding financial benefits findings?

- In the comment box following your response, note the rationale for your selection.
- This question addresses measurement bias; i.e., systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s).
- Consider the validity of the measurement method; consider reliability to the extent that could bias findings. Examples:
 - Consider whether blinding was used in the assessment of final outcome(s) and whether it was necessary
 - The introduction of an OHSMS could alter reporting practices, so this potential source of bias should be eliminated by some means
- Consider the measurement methods independent of the sample issues. For example, in the Nytro et al. study, there was concern that the methods used for implementation might have created a bias in the direction of better compliance because company managers would have an incentive to appear as though they are complying with the law. If we had been told that the self-report method had been verified in a sub-sample through an external audit, we could have given a high rating for the measurement methods, even though the means of selecting the sample (quota sampling; no indication of refusals) could nevertheless have biased the results towards more compliance.

- Consider all measurement methods pertinent to the corresponding findings (not just the financial benefits variables). For example, in the case of the intermediate and final outcomes findings, the method of measuring OHSMS variables might also be relevant.
- Consider issues of statistical power in Q43 not here.

18. Were appropriate statistical tests conducted on the financial benefits data?

- In the comment box following your response, note the rationale for your selection.
- The issue of whether adjustment for confounders took place is covered in Q9. In this question, one is only concerned about whether such adjustments were done correctly, if they were done
- Do not fear choosing the option of “reviewer not qualified to answer.” We are simply using this response to flag where we need expertise from elsewhere in the group or even external to the group. If two reviewers have selected this option, then we seek expertise from elsewhere in the group, and failing this, use our external consult. If only one reviewer has selected this option, then a consensus answer can be formulated, if the remaining reviewer is highly confident of his/her answer. If not, then expertise should be sought elsewhere in the group or externally if required.
- Was the test appropriate for the sample size and type of data?
- If no statistical tests were used, and you think they were not needed, select the answer corresponding to this. If no statistical tests were used and you think they should have been, then use one of the following three options: minor deficiencies; inappropriate statistical test, which might modify the conclusions; inappropriate statistical tests, which could substantially affect conclusions.

19. Are you confident that there are no additional potential sources of bias in the estimate of implementation/effectiveness not yet captured in the previous questions?

- This is the place to discuss any other sources of bias; e.g., contamination of the comparison intervention.

20. What is your overall appraisal of the quality of evidence about implementation provided by this study?

- Consider your answers to questions 7, 8, 9, 11, 12, 19 and give an overall assessment.

21. What is your overall appraisal of the quality of evidence about intermediate outcomes provided by this study?

- Consider your answers to questions 7, 8, 9, 13, 14, 19 and give an overall assessment.

22. What is your overall appraisal of the quality of evidence about final outcome(s) provided by this study?

- Consider your answers to questions 7, 8, 9, 15, 16, 19 and give an overall assessment.

23. What is your overall appraisal of the quality of evidence about financial benefits provided by this study?

- Consider your answers to questions 7, 8, 9, 17, 18, 19 and give an overall assessment.

Questions 24-30 are concerned with any facilitators and barriers aspects of the study.

24. If qualitative research methods were used to identify facilitators and barriers, how do you rate the quality of those methods?

- In the comment box following your response, note the rationale for your selection. Refer to excerpt from Spencer et al. (2003) Quality in qualitative evaluation: A framework for assessing research evidence. London: Strategy Unit, Cabinet Office.

25. Are you confident that the means of selecting and maintaining the sample used in quantitative F/B study minimized bias?

- See Q8 above.

26. Are you confident that (in the quantitative study of F/Bs) potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?

- See Q9 above.

27. Are you confident the measurement methods did not introduce bias to the corresponding F/B findings?

- In the comment box following your response, note the rationale for your selection.
- This question addresses measurement bias; i.e., systematic error arising from inaccurate measurement (or classification) of subjects on study variable(s).

- Consider the validity of the measurement method; consider reliability to the extent that could bias findings.
- Consider the measurement methods independent of the sample issues. For example, in the Nytro et al. study, there was concern that the methods used for implementation might have created a bias in the direction of better compliance because company managers would have an incentive to appear as though they are complying with the law. If we had been told that the self-report method had been verified in a sub-sample through an external audit, we could have given a high rating for the measurement methods, even though the means of selecting the sample (quota sampling; no indication of refusals) could nevertheless have biased the results towards more compliance.

28. Were appropriate statistical tests conducted on the F/B data?

- See Q12 above.
- If no statistical tests were used, and you think they were not needed, select the answer corresponding to this. If no statistical tests were used and you think they should have been, then use one of the following three options: minor deficiencies; inappropriate statistical test, which might modify the conclusions; inappropriate statistical tests, which could substantially affect conclusions.
- Consider issues of statistical power in Q43 not here.

29. Are you confident that there are no additional potential sources of bias in the quantitative study of F/Bs not yet captured in the previous questions?

30. What is your overall appraisal of the quality of evidence about facilitators and barriers provided by the quantitative study?

- Consider your answers to questions 25-29 and give an overall assessment.

31. Should this publication go on to data extraction (Level 5)?

32. Is supplementary information required?

33. Is this the final version of the Level 4 assessment?

Data Extraction

The following questions are included for generating the evidence table. Keep this in mind when deciding how much information to include.

General instructions for Q1-Q45:

- Use numbers instead of text for expressing quantitative information
- Use the following abbreviations: N/A, not applicable; NR, not reported.
- Include page numbers to show from where information has been drawn

1. What was the research question?

- Use author's own words if provided.

2. Describe the study sample and sampling frame, including the sample size, type of industry, size of workplace(s), types of occupations (if included), number of workplaces, and the method of sampling.

- Include any information about withdrawals here.

3. Give any information pertaining to maintenance of the sample (i.e., withdrawals, etc.) (Put N/A if not applicable)

4. Describe the intervention in detail.

- Include all information about:
 - elements of the OHSMS being intervened upon
 - levels (e.g., government, head office, etc.) at which the intervention was initiated and delivered
 - by whom, with whose involvement, how, and with what intensity the intervention was delivered
- Include the corresponding information about any comparison group(s).
- Include any information about contamination of the comparison group, or compliance of the intervention group (excluding implementation results) here too.

5. Describe the experimental design in detail.

- Elaborate on the design selected in Q7, identifying the number of groups and the timing of measurements.

6. Describe relevant data collection, data transformation, and measurement properties under the applicable evidence category.

- If there were any other categories of measures not covered by the above categories, flag them here.
- Describe the way the data were collected and transformed into the final measure of effect.
- Detail any reliability and validity information.

7. Describe statistical tests used.

- For each type of measure, detail statistical approach taken, including mention of software when relevant.
- List any confounders for which there was adjustment.

8. Describe all findings regarding implementation

- For point estimates, include units and confidence intervals where relevant. Identify the time of the measures in a way consistent with the description of the design in Q35.
- For effect estimates, include units and confidence intervals, where relevant.
- State statistical test, statistical hypothesis, test statistic, p-value, and effect size, as provided in the text.

9. Describe all findings regarding intermediate OHS outcomes.

- For point estimates, include units and confidence intervals where relevant. Identify the time of the measures in a way consistent with the description of the design in Q35.
- For effect estimates, include units and confidence intervals, where relevant.
- State statistical test, statistical hypothesis, test statistic, p-value, and effect size, as provided in the text.

10. Describe all findings regarding final OHS outcomes.

- For point estimates, include units and confidence intervals where relevant. Identify the time of the measures in a way consistent with the description of the design in Q35.
- For effect estimates, include units and confidence intervals, where relevant.
- State statistical test, statistical hypothesis, test statistic, p-value, and effect size, as provided in the text.

11. Describe all findings regarding cost of the intervention. (*Put N/A if not applicable*)

12. Describe all findings regarding the financial benefit of the intervention.

- For point estimates, include units and confidence intervals where relevant. Identify the time of the measures in a way consistent with the description of the design in Q35.
- For effect estimates, include units and confidence intervals, where relevant.
- State statistical test, statistical hypothesis, test statistic, p-value, and effect size, as provided in the text.

13. Did the design lack statistical power?

- State your response and then support with information from the publication where available.
- Do not fear choosing the option of “reviewer not qualified to answer.” We are simply using this response to flag where we need expertise from elsewhere in the group or even external to the group. If two reviewers have selected this option, then we seek expertise from elsewhere in the group, and failing this, use our external consult. If only one reviewer has selected this option, then a consensus answer can be formulated, if the remaining reviewer is highly confident of his/her answer. If not, then expertise should be sought elsewhere in the group or externally if required.

14. List findings regarding facilitators and barriers.

15. Were any harms of the intervention identified?

- Please describe.

16. Provide any additional comments that you think should be included in the evidence table about this study. (*Put N/A if not applicable*)

17-43. Please indicate the presence of the following OHSMS elements, using the Redinger (2004) or Dalrymple, Redinger, Dyjack, Levine & Mansdorf (1998; ILO report) scheme. Use the following code to indicate the presence/absence of an element: “X”, present; “/”, partially present; “-“, not present; “u”, unknown.

17. Management commitment and resources

18. Regulatory compliance and OHSMS conformance
19. Accountability, responsibility, and authority
20. Employee participation
21. Occupational health and safety policy
22. Goals and objectives
23. Performance measures
24. System planning and development
25. Baseline evaluation and hazard/risk assessment
26. OHSMS manual and procedures
27. Training system
28. Technical expertise and personnel qualifications
29. Hazard control system
30. Process design
31. Emergency response
32. Hazardous agent management
33. Preventive and corrective actions
34. Procurement and contractor selection
35. Communication system
36. Document and record management system
37. Evaluation system
38. Auditing and self-inspection
39. Incident investigation and root cause analysis
40. Medical program and surveillance
41. Continual improvement

42. Integration

43. Management review

44. Is supplementary information required?

- Detail here any references needed to provide more complete information for the Level 4 assessment, including any documents (e.g., description of legislation or guidelines) needed to complete Q47-73.

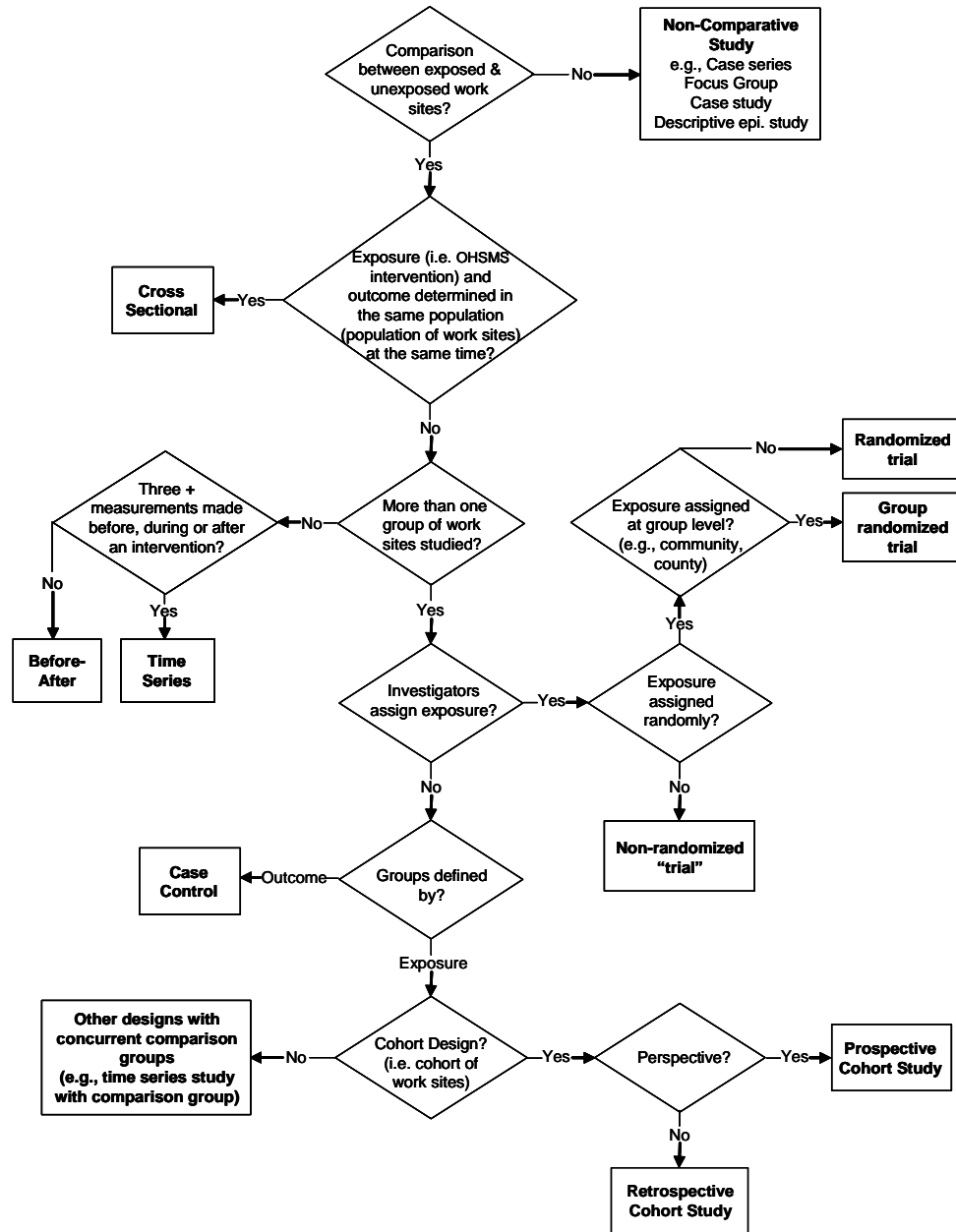
45. **Is this the final version?**

- Answer YES if this is the post-consensus version.

Attachments:

- Appendix B from Robson et al. (2001) “Guide to Evaluating the Effectiveness of Strategies for Preventing Work Injuries” (provided with previous draft)
- Excerpt from Spencer et al. (2003) “Quality on Qualitative Evaluation”

Figure E.1: Study Design Algorithm



Appendix F - References for Included Publications

Relevant Studies That Were Assessed for Quality and Proceeded to Data Extraction. (n=9)

Alsop P, LeCouteur M. Measurable success from implementing an integrated OHS management system at Manningham City Council. *Journal of Occupational Health & Safety - Australia & New Zealand*. 1999; 15(6):565-572.

Dufour C, Lanoie P, Patry M. Regulation and Productivity. *Journal of Productivity Analysis*. 1998; 9(3):233-247.

Edkins GD. The INDICATE safety program: Evaluation of a method to proactively improve airline safety performance. *Safety Science*. 1998; 30(3):275-295.

Lewchuk W, Robb AL, Walters V. The Effectiveness of Bill 70 and Joint Health and Safety Committees in Reducing Injuries in the Workplace: The Case of Ontario. *Canadian Public Policy* 1996; 22(3):225-243.

Nytro K, Saksvik P, Torvatn H. Organizational prerequisites for the implementation of systematic health, environment and safety work in enterprises. *Safety Science*. 1998; 30(3):297-307.

Pearse W. Club zero: Implementing OHSMS in small to medium fabricated metal product companies. *Journal of Occupational Health & Safety - Australia & New Zealand*. 2002; 18(4):347-356.

Saksvik PO, Nytro K. Implementation of internal control (IC) of health, environment and safety (HES) in Norwegian enterprises. *Safety Science* 1996; 23(1):53-61.

Torp S, Riise T, Moen BE. Systematic health, environment and safety activities: do they influence occupational environment, behaviour and health? *Occupational Medicine (Oxford)*. 2000; 50(5):326-333.

Yassi A. Utilizing data systems to develop and monitor occupational health programs in a large Canadian hospital. *Methods of Information in Medicine*. 1998; 37(2):125-129.

Relevant Studies That Were Assessed for Quality But Did Not Proceed to Data Extraction (n=9)

Anonymous. Best practice *case study*. Portland aluminium. Journal of Occupational Health & Safety - Australia & New Zealand. 1993; 9(4):375-378.

Anonymous. Best practice case study: Herbert Adams Bakeries. Journal of Occupational Health & Safety - Australia & New Zealand. 1994; 10(3):275-278.

Bolton FN, Kleinsteuber JF. A Perspective on the Effectiveness of Risk Assessment by First-Line Workers and Supervisors in a Safety Management System. Human and Ecological Risk Assessment. 2001; 7:1777-1786.

Chinander KR, Kleindorfer PR, Kunreuther HC. Compliance Strategies and Regulatory Effectiveness of Performance-Based Regulation of Chemical Accident Risks. Risk Analysis. 1998; 18(2):135-143.

Dotson K. An international safety and health measurement strategy: Corporate programs, systems and results. Journal of Occupational Health & Safety - Australia & New Zealand. 1996; 12(6): 669-678.

Eisner HS, Leger JP. The international safety rating system in South African mining. Journal of Occupational Accidents. 1988; 10(2):141-160.

Lanoie P. Safety Regulation and the Risk of Workplace Accidents in Quebec. Southern Economic Journal. 1992; 58(4):950-965.

Lund HL. Strategies for Sustainable Business and the Handling of Workers' Interests: Integrated Management Systems and Worker Participation. Economic and Industrial Democracy. 2004; 25(1):41-74.

Mitchell CS. Evaluating occupational health and safety programs in the public sector. American Journal of Industrial Medicine. 1998; 34(6):600-606.

Appendix G – Characteristics of Studies Considered in this Review

Table G.1 Study characteristics for studies proceeding to data extraction (n=9)

Author, Yr	Jurisdiction	Study Design*	OHSMS Type	Type of Evidence	Type of Outcomes	Number of Workplaces	Workplace Size	Industrial Sector**	Proceed to DE
Mandatory OHSMSs									
Dufour, 1998	Quebec, Canada	Time series	Mandatory	Effectiveness	Financial	Most manufacturing sectors	Mixed sizes	Manufacturing	Yes
Lewchuk, 1996	Ontario, Canada	Time series with comparison group	Mandatory	Effectiveness	Final	636 workplaces	Mixed sizes	Manufacturing Retail Trade	Yes
Nyto, 1998	Norway	Before-after	Mandatory	Implementation	Implementation	1184 workplaces	Mixed sizes	Multiple sectors	Yes
Saksvik, 1996	Norway	Cross-sectional	Mandatory	Implementation Effectiveness	Implementation Intermediate Final	2092 workplaces	Mixed sizes	Multiple Sectors	Yes
Torp, 2000	Norway	Cross-sectional	Mandatory	Effectiveness	Intermediate Final	311 workplaces	Small and/or medium enterprises	Other Services except Public Administration (auto repair)	Yes
Voluntary OHSMSs									
Alsop, 1999	Australia	Time series	Voluntary	Effectiveness	Financial	1 workplace - multiple worksites	Large enterprise	Public Administration	Yes
Edkins, 1998	Australia	Non-randomized trial	Voluntary	Effectiveness	Intermediate	1 workplace - 2 worksites	Small and/or medium enterprises	Transportation	Yes
Pearse, 2002	Australia	Before-after	Voluntary	Implementation Facilitator/Barrier	Implementation F/B – Non-systematic	20 workplaces	Small and/or medium enterprises	Manufacturing (fabricated metal product)	Yes
Yassi, 1998	Manitoba, Canada	Time series	Voluntary	Effectiveness	Financial	1 workplace	Large enterprise	Health Care and Social Assistance	Yes

* Study design identified through modified study design flowchart (Zaza et al., 2000)

** Industrial sector identified through NAICS classification. <http://www.statscan.ca/english/Subjects/Standard/naics/2002/naics02-menu.htm>.

Table G.2 Study Characteristics for Studies Not-proceeding to Data Extraction (n=9)

Author, Yr	Jurisdiction	Study Design*	OHSMS Type	Type of Evidence	Type of Outcomes	Number of Workplaces	Workplace Size	Industrial Sector**	Proceed to DE
Mandatory OHSMSs									
Chinander 1998	Multiple states, USA	Qualitative	Mandatory	Facilitator/Barrier	F/B – Qualitative	16 workplaces	Unclear	Manufacturing (chemical)	No
Lanoie, 1992	Quebec, Canada	Time series	Mandatory	Effectiveness	Final	Most industrial sectors	Mixed sizes	Multiple sectors	Yes
Mitchell, 1998	Mid-Atlantic state, USA	Qualitative	Mandatory	Facilitator/Barrier	F/B – Qualitative	1 workplace	Unclear	Public Administration	No
Voluntary OHSMSs									
Bolton, 2001	California, USA	Time series	Voluntary	Effectiveness	Final F/B – Non-systematic	1 workplace - multiple worksites	Unclear	Professional, Scientific and Technical Services	No
Dotson, 1996	Arizona, USA	Cross-sectional Before-after	Voluntary	Implementation Effectiveness Facilitator/Barrier	Implementation Final F/B – Non-systematic	1 workplace - multiple worksites	Large enterprise	- Mining and Oil and Gas Extraction - Manufacturing	No
Eisner, 1988	South Africa	Cross-sectional	Voluntary	Effectiveness	Final	33 workplaces	Unclear	Mining, Oil, Gas Extraction	No
Lund, 2004	Denmark	Qualitative	Voluntary	Facilitator/Barrier	F/B -	2 workplaces	Mixed sizes	Manufacturing (chemical)	No
Unknown, 1993	Australia	Time series	Voluntary	Effectiveness	Final	1 workplace	Large enterprise	Manufacturing	No
Unknown, 1994	Australia	Before-after	Voluntary	Effectiveness	Final	1 workplace	Large enterprise	Manufacturing	No

* Study design identified through modified study design flowchart (Zaza et al., 2000)

** Industrial sector identified through NAICS classification. <http://www.statcan.ca/english/Subjects/Standard/naics/2002/naics02-menu.htm>.

Appendix H - Detailed Summary of Evidence

Table H.1.1.

Study #1: Alsop, P. and LeCouteur, M. Measurable success from implementing an integrated OHS management system at Manningham City Council, *Journal of Occupational Health & Safety - Australia & New Zealand*. 1999; 15(6): 565-572.

Research Question:	
What were the achievements and key learning experiences that were obtained through the implementation of a health and safety management program integrated into other management systems at Manningham City Council?	
Experimental Design:	
TIME SERIES	
Time series design, with yearly measures of financial benefits (premium rates) from 1992/93 to 1999/2000	
Sample Characteristics:	
Sample consisted of one municipal government (Manningham City Council, a metropolitan local government near Melbourne, Australia) employing >500 people. The municipality was located 15 km east of Melbourne, Australia.	
Maintenance of Sample:	
Some employment changes over the period of observation are reported, with a 25% reduction in the size of the workforce. An amalgamation and restructuring of business units took place that the authors suggested caused slower than expected progress.	
Intervention:	
VOLUNTARY OHSMS	
In May 1995, work began on the implementation of an OHSMS based on "various Australian and international standards" (not further described in article). It involved the integration of OHS management with Quality Management (ISO 9001), and Environmental Management (ISO 14001) Systems. Emphasis was on common corporate processes (e.g. training, procurement and supply, hazard identification and incident reporting, risk assessment and control) and those factors/risk exposures which were business-unit specific (e.g., tree cutting, lifting of frail aged). SafetyMAP Achievement levels sought and achieved (Initial Level Achievement since Aug '98 for all operations and Advanced Level Achievement since June '99 for largest business unit.) Intervention initiated by municipal government.	
Data Collection, Data Transformation, and Measurement Properties:	
<input type="checkbox"/> Implementation	
<input type="checkbox"/> Intermediate Outcomes	
<input type="checkbox"/> Final OHS Outcomes	
<input checked="" type="checkbox"/> Economic Outcomes	Workers Compensation Premiums 1991/92 to 1999/2000. Results also given for net premium rates for individual business units, in comparison to industry rate. We don't, however, have the equivalent comparison from before OHSMS instituted.

<input type="checkbox"/> Facilitators/ Barriers	
Statistical Tests:	
Not applicable.	
Results:	
Implementation	Not applicable
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	WC premium rates declined each year over the study period and the authors report that rates were reduced from 4.7564% in 1993/94 to 1.323% in 1999/2000 (% of remuneration). Individual business units' "net premium rate as a percentage of remuneration" was also compared to Industry net premium rate, with Manningham having lower premiums in 19 of the 20 business units for the year 1999-2000.
Facilitators/ Barriers	Not applicable
Did the Design Lack Statistical Power?	
Not applicable: no statistical tests.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
The implementation of Quality Management Systems and Environmental Management systems at virtually the same time as the OHSMS may have contributed to the results seen here. Also, the number of claims was also decreasing at the time, suggesting that the change in premium rates was due to a change in injury experience. (However, we have not reported these latter results above, due to a judgment at the QA phase that there was potential bias in methods of measurement). There was already a trend of decreasing premium rates prior to the intervention. Thus, estimate of change due to the intervention might be overestimated.	

Table H.1.2.

Study #2: Dufour, C., Lanoie, P., and Patry, M. Regulation and Productivity, Journal of Productivity Analysis. 1998; 9(3): 233-247.

Research Question:	
What was the impact of occupational health and safety (OHS) and environmental regulations on the rate of productivity growth (total factor productivity) in the Quebec manufacturing sector during 1985-1988?	
Experimental Design:	
POOLED TIME SERIES	
Pooled time-series using three annual changes between 1985 and 1988, for each of the 19 manufacturing sectors.	
Sample Characteristics:	
The sample consisted of all Quebec workplaces in 19 manufacturing sectors that were subject to LSST legislation and CSST enforcement during the period 1985-88. The unit of analysis was the sector.	
Maintenance of Sample:	
Not reported.	
Intervention:	
MANDATORY OHSMS	
In Quebec, the 1979 Loi sur la santé et la sécurité du travail (LSST) legislation led to the creation of the Commission de la Santé et Sécurité du Travail (CSST). The study focused on specific aspects of the Quebec regulation: workplace inspections, investigation of refusal to work, penalties imposed, protective reassignment, and implementation of prevention OHS programs. Although the CSST policies and activities were wide ranging, four of the five selected aspects are cited in another publication (Lanoie, 1992) as being a major focus of their policies and enforcement activities. The level of implementation of the regulation at the firm level increased following the creation of the CSST in 1980, although a number of the requirements had phase-in periods (e.g., prevention programs were legislated in 1979 but employers in Group I and Group II had until July 3, 1983 and May 4, 1984, respectively for preparation of their programs). Additionally, the level of implementation varied across industry sectors as the requirements for OHS prevention programs and JHSCs pertained to firms employing 20 or more workers in 15 high hazard industries.	
Data Collection, Data Transformation, and Measurement Properties:	
<input type="checkbox"/> Implementation	
<input type="checkbox"/> Intermediate Outcomes	
<input type="checkbox"/> Final OHS Outcomes	
<input checked="" type="checkbox"/> Economic Outcomes	Dependent variable: total factor productivity growth (TFP) = difference between real output growth and real input growth. Real output growth was based on the annual value of shipments. Real input growth was determined from the following: cost of materials and supplies, cost of fuel and electricity, wages of production and related workers, wages of administrative, office and other non-manufacturing employees, and capital costs (stock and investment, adjusted for depreciation, taxation, etc.).

	<p>Independent variables: ENVIRONMENT = changes in the ratio of the value of investment in pollution-control equipment to total costs in industry; five variables for changes in the intensity of OHS regulations that served as proxies for firm expenditures for compliance, INSPECTION for inspection rate, INFRACTION for rate of penalties imposed, REFUSAL for rate of interventions for refusal to work, PREVENT for percent of firms having a prevention program, and PROTECT for rate of protective reassignment; control variables were CYCLE, change in capacity utilization index (to control for cyclic fluctuations in the presence of quasi-fixed costs), SCALE, change in level of output (to control for effect of changing economies of scale), and ENERSHARE, change in the cost share of energy (to control for differing rate of productivity growth across industries in the presence of oil price changes, arising from their different energy-intensities). The data were obtained from various published and unpublished reports from sources such as Statistics Canada, the Federal Government, Bank of Canada, and CSST. Dummy variables for manufacturing sectors and time were also included as independent variables.</p>
<input type="checkbox"/> Facilitators/ Barriers	
<p>Statistical Tests:</p>	
<p>Estimates were performed using generalized least squares procedures based on the cross-sectionally and time-wise autoregressive model presented by Kmenta (1986). Thus, the model accounted for possible serial correlation (annual industry data) and heteroskedasticity due to the cross sectional data from diverse sources. Additional statistical tests for exogeneity of variables were performed to see if productivity growth influenced the level of regulation (the exogeneity of the prevention program was rejected). Six models, based on 57 data points, were developed to investigate the effect of various groups of variables in the overall model (i.e., effect of removing ENVIRONMENT or ENERSHARE or SCALE). The final model for TFP contained nine variables.</p>	
<p>Results:</p>	
Implementation	Not applicable
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	<p>Effect estimates: The prevention program variable (PREVENT) had a statistically significant ($p < 0.05$) and a positive regression coefficient, meaning that it appears to have a positive effect on productivity growth. INFRACTION also had a significant and positive coefficient, and the implied contribution of both variables on productivity growth was 0.007 (at the sample mean). The authors suggest that these findings indicate that prevention programs and penalties have reduced the incidence of workplace injuries in the manufacturing sector, leading to a reduction of direct and indirect costs sufficient to have an enhancing effect on productivity growth. It should be noted, however, that the PROTECT variable was also significant and with its negative coefficient its implied contribution on productivity was -0.019. Thus the net effect of the OHS variables was -0.012 which is larger than the average impact found in a study of American OHS regulations (-0.003).</p>

	<p>Test statistics: t-statistics are presented for each of the independent variables in the 6 specifications of the TFP growth equation. The t-statistics for the final (selected) model for the OHS variables are as follows: INSPECTION 0.41 (NS), REFUSAL 1.14 (NS), PROTECTIVE -1.91 (P<0.01), INFRACTION 3.02 (P<0.05), PREVENTION 2.86 (P<0.05).</p>
Facilitators/ Barriers	Not applicable
Did the Design Lack Statistical Power?	
Using the rule of thumb of 10 observations per predictor variable, the study may be slightly underpowered (i.e., analysis is based on 57 observations and 9 independent variables are in the model).	
Were Any Harms of the Intervention Identified?	
None reported.	
IWH Reviewers' Comments:	
<p>This is a pooled time series study of the degree of implementation of LSST regulations for health and safety and effect on productivity growth. The equation for TFP was well developed theoretically and was an improvement upon those used in past analyses of OHS regulations in the US. The statistical analyses were quite thorough, however it was unusual that one of the specifications with 3 significant variables removed had a higher R-squared value than one with those variables included. The findings were unique and suggested that there was the potential for positive economic impacts from the implementation of prevention programs. The primary concerns pertained to the level of observation and analysis and the possibility for confounding. The model was based on 57 observations from aggregated industry-level data. Some of the variables were surrogates for constructs in the model so there was an issue of validity of the measures. It was impossible to rule out the effects of variables operating at the firm level or other possible co-intervening factors on the final outcome variable.</p>	

Table H.1.3.

Study #3: Edkins, G. D. The INDICATE safety program: Evaluation of a method to proactively improve airline safety performance, *Safety Science*. 1998; 30(3): 275-295.

Research Question:
To evaluate effectiveness of INDICATE safety program in improving airline safety performance.
Experimental Design:
NON-RANDOMIZED TRIAL Non-randomized trial with two groups, pre- and post-intervention (8 months) measures.
Sample Characteristics:
<u>INDUSTRY</u> : Airline passenger transportation <u>WORKPLACE (TYPE and NUMBER)</u> : 1 regional airline in Australia - 2 sites (regional centres): 1 control, 1 intervention <u>SIZE</u> : Intervention - 81 staff; Control - 72 staff. <u>OCCUPATIONS</u> : not explicitly described, but text suggested that it included pilots, cabin crew, ground crew, and maintenance. <u>SAMPLING METHOD</u> : not described, but likely convenience.
Maintenance of Sample:
The two worksites were maintained throughout the study, with the same staffing levels.
Intervention:
VOLUNTARY OHSMS 8 month trial (starting July 1995) <u>Intervention Group</u> <i>INDICATE Program:</i> <ol style="list-style-type: none">1. Operational safety manager that: coordinated meetings with mgmt; oriented staff to INDICATE program; coordinated safety training; encouraged staff to use confidential safety reporting system; served as confidential contact for staff on safety issues; served as technical resource to staff and management; communicated safety information to staff; participated in incident investigations; evaluated and improved safety program.2. Focus group discussions with "as many company staff and management as possible" to identify hazards local to their area, assess the associated level of risk, evaluate the current defences for dealing with those risks, and make recommendations for remedial safety action.3. Confidential hazard reporting system. Staff can report any concerns about safety including witnessing or experiencing unsafe work practices, or identifying deficient safety equipment. Reports flow to safety manager, who investigates and enters information into INDICATE software and gives feedback to reportee.4. Regular safety meetings. Purpose of the meetings is to manage, monitor, and address identified hazards in conjunction with managers.5. Safety information database. This is a tool for safety manager that facilitates i) tracking hazards for each area; ii) issuing recommendations and monitoring action on them; iii) outputting summaries.6. Communication of safety information with staff. Multiple methods were used including: safety newsletter; safety notice board; impromptu safety briefings; distribution of minutes; and a safety reading file with an account of action taken on hazards.

<u>Comparison Group</u>	
Only the confidential hazard reporting system described above (#3)	
Data Collection, Data Transformation, and Measurement Properties:	
<input type="checkbox"/> Implementation	
<input checked="" type="checkbox"/> Intermediate Outcomes	<p><u>Intermediate1:</u> Airline Safety Culture Index (ASCI; p. 284) scores. ASCI is a survey-based instrument consisting of 25 positively worded statements requiring a response on a 5-point Likert scale ranging from strongly agree to strongly disagree. Content covered the following concepts: management commitment, management action, employee commitment, level of perceived risk, beliefs about accident causation, emergency procedures, provision of safety training and safety communication. Seemed to be based on existing safety climate tools. Was pre-tested with 10 aviation safety professionals. High internal consistency measures (0.94 (T1) and 0.96 (T2)). Test-retest reliability established ($r = 0.67$). Author gives two citations supporting the link between climate and accidents and there are others that the author didn't cite; i.e. the predictive validity of climate scores has been shown elsewhere, but not shown for this particular tool. Apparently there was an 80% response rate at T1 and T2, though the numbers given for # of respondents at T1 (p. 286) is the same as that given for total employees (p. 283).</p> <p><u>Intermediate2a and b:</u> Staff hazardousness perception scores and staff hazard likelihood perception scores. p. 284. A list of 22 safety hazards developed from accident reports from 1990-1996. (i.e. content validity). Respondents were asked to rate each hazard as to its hazardousness (2a) and likelihood (2b) using a 9 point scale. The measurement method was based on some citations lacking peer review. No further information on validity and reliability of this particular tool. However, in the discussion (p. 292), it says that "research evidence suggests that there is little correlation between risk perception and actual accident or incident frequencies," suggesting a lack of predictive validity.</p> <p><u>Intermediate 3:</u> Staff willingness to report hazards. Number of hazards reports submitted to the confidential hazard reporting system. The period of collection was unspecified and is therefore assumed to be the 8-month intervention period.</p> <p><u>Intermediate 4:</u> Action taken on identified safety hazards. Number of hazards identified and addressed by each group.</p> <p><u>Intermediate 5:</u> Staff comments on safety management. Employees in both intervention and comparison groups were asked to comment on company safety management both before and after the intervention (method of data collection unspecified).</p>
<input type="checkbox"/> Final OHS Outcomes	
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	
Statistical Tests:	
Intermediate 1, 2a, 2b: Repeated measures ANOVA. Intermediate 1 was also tested for baseline differences using t-test. Intermediate 3 and 4: No statistical tests.	

Results:	
Implementation	Not applicable
Intermediate OHS Outcomes	<p>Effect estimates: <u>Intermediate 1:</u> Intervention (from figure): pre: 58; post: 45. Comparison (from figure): pre: 69; post: 75. <u>Intermediate 2a:</u> Intervention (from figure): pre: 140; post: 92. Comparison (from figure): pre: 140; post: 120. <u>Intermediate 2b:</u> Intervention (from figure): pre: 76; post: 44. Comparison (from figure): pre: 76; post: 68. <u>Intermediate 3:</u> Intervention: 48. Comparison: 9. <u>Intermediate 4:</u> Intervention: 13. Comparison (implied): 0.</p> <p>Test statistics: Intermediate 1: $F(1,148) = 109.05$, $t(1) = -4.63$. Intermediate 2a: $F(1,141) = 66.64$. Intermediate 2b: $F(1,139) = 41.34$.</p> <p>Qualitative findings: Intermediate 5: Comments from the intervention group were positive about the program and indicated belief in its effectiveness. Comments from the comparison group suggested a reluctance to report incidents because of fear of retribution.</p> <p>P-value: Intermediate 1, 2a, 2b: $p < 0.001$ for all F-statistics and t-statistics.</p>
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
No. Effect detected in all tests conducted.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
<p>Uncertainty existed as to reason that one site was selected for intervention and the other one not, as well as the equivalency of other events during the trial. The initial equivalency of the two sites was uncertain, because there was no investigation of staff characteristics at the two sites, and no other description that would reassure the reader of equivalency. The initial differences in ASCI scores between the two groups only furthered this concern. On the other hand, the study used multiple measures, some of which were objective, and all of which were consistent in their direction of change.</p>	

Table H.1.4.

Study #4: Lewchuk, W., Robb, A. L., and Walters, V. The Effectiveness of Bill 70 and Joint Health and Safety Committees in Reducing Injuries in the Workplace: The Case of Ontario, Canadian Public Policy. 1996; 22(3): 225–243.

Research Question:
Did Bill 70, which took effect at the end of 1979, change the work-related injury and illness frequency rate?
Experimental Design:
<p>POOLED TIMES SERIES WITH COMPARISON GROUP</p> <p>This pooled time-series cross-sectional design compares the period BEFORE implementation of the bill (1976-1979) to the period AFTER implementation of the bill (1980-1989). The comparison is made in BOTH manufacturing and retail sectors. The authors then contrast the before-after comparisons in the two sectors.</p>
Sample Characteristics:
<p><u>TYPE OF INDUSTRY:</u> 436 were manufacturing and 201 were retail.</p> <p><u>SAMPLE SIZE:</u> 637 (only 636 participated).</p> <p><u>SIZE OF WORKPLACES:</u> Manufacturing enterprises had at least 50 employees in 1988, Retail enterprises had at least 20 employees in 1988, There is no indication of the range in sizes of the workplaces, nor is there any indication of the size of the workplaces over time (data was presented for the period 1976 through to 1989, so the number of employees in 1988 may not accurately represent the number of employees in 1976 through to 1987 or in 1989), Retail enterprises had at least 20 employees in 1988, There is no indication of the range in sizes of the workplaces, nor is there any indication of the size of the workplaces over time (data was presented for the period 1976 through to 1989, so the number of employees in 1988 may not accurately represent the number of employees in 1976 through to 1987 or in 1989), Number of Workplaces: 637 (only 636 participated).</p> <p><u>SAMPLING METHOD:</u> Sample included all those workplaces that had provided at least some information in a 1991 IAPA survey of workplaces (n=497) plus some additional workplaces (n=140), The 140 additional workplaces were all non-respondents to the IAPA survey, Although the authors report that the additional workplaces were included to “maintain the confidentiality of the IAPA survey participants”, no specific criteria is provided with respect to how these workplaces were chosen from all non-responding workplaces, nor is there any indication of the proportion that were chosen from the retail sector versus the proportion chosen from the manufacturing sector, The response rate to the survey was low (539/1032 = 52%), Even with the addition of the 140 non-responding workplaces, the response rate would still be low (679/1032 = 65%), Therefore, sampling bias may affect this study whereby those workplaces which initially participated in the IAPA study and those workplaces that were chosen from the non-respondents do not accurately represent the OHS activities and outcomes of all workplaces, In fact, it may be that the sample represents OHS conscientious workplaces since those likely to respond to a survey about OHS activities and outcomes are likely those workplaces which are active in OHS activities and have positive outcomes, If overrepresentation of OHS conscientious workplaces occurred equally across the sectors (retail and manufacturing), the estimate of the effect of the legislation will be unbiased, however, the results will not be generalizable to workplaces outside of the study.</p>
Maintenance of Sample:
Not applicable. Cross-sectional time series.

Intervention:	
MANDATORY OHSMS	
<p>Bill 70 is a piece of legislation titled Occupational Health and Safety Act which took effect in October 1979. This legislation required organizations to take on an Internal Responsibility System. This system places greater emphasis on having employers and employees regulate their own working conditions. One key element of the legislation is the requirement for workplaces to form a <u>Joint Health and Safety Committee (JHSC)</u>. The legislation applies to workplaces with over 20 employees and exempts some non-manufacturing firms.</p>	
Data Collection, Data Transformation, and Measurement Properties:	
<input type="checkbox"/> Implementation	
<input type="checkbox"/> Intermediate Outcomes	
<input checked="" type="checkbox"/> Final OHS Outcomes	<p>Frequency of lost-time accidents: The authors note that the number of lost-time accidents is a good outcome because it is more likely to be reported than non-lost time accidents. However, lost-time accidents may be affected by socioeconomic circumstances of the injured person. For example, an injured worker may not take off the necessary time for recovery if he or she cannot afford such a loss in income.</p>
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	
Statistical Tests:	
<p><u>Multiple regression, t-test for the beta coefficients:</u> Regression analyses for the firms in the manufacturing and retail sectors are presented separately and then compared. There are two types of regression analyses conducted for each sector: (1) compare the beta coefficient for the Bill 70 variable in the manufacturing sector regression analysis TO the beta coefficient for the Bill 70 variable in the retail sector regression analysis (just two periods 76-79 VS 80-89). (2) compare the beta coefficients for three period variables (1980, 1981, & 1982+) in manufacturing sector regression analysis TO the beta coefficients for three periods in the retail sector regression analysis. The purpose of the second regression is to see if there are delayed effects of Bill 70. In both regression models, dummy variables for time trend, industry rate group, and union status were included. An analysis that included a quadratic trend for time was also conducted. The analyses were based on 5,400 workplace-year observations for the manufacturing firms and 2,494 for the retail firms.</p>	
Results:	
Implementation	Not applicable
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	<p>Effect estimates: (1st Regression – beta coefficient for Bill 70) Manufacturing = -0.015, Retail = -0.004; (2nd Regression – beta coefficients for 1980, 1981 and 1982+) Manufacturing = -0.011, -0.014, -0.024, Retail = -0.005, -0.3, -0.5</p> <hr/> <p>P-value: No p-values are stated. Only whether the coefficient was significant or not was reported. But because n is greater than 30, z scores apply. Therefore, (1st Regression – p value for Bill 70) Manufacturing less than 0.01, Retail greater than 0.05; (2nd Regression – p value for 1980, 1981 and 1982+) Manufacturing = less than 0.01, less than 0.01, less than 0.01, Retail = greater than 0.05, greater than 0.05, greater than 0.05</p>

	Statistical hypothesis: Beta coefficients will be equal to zero
	Test statistics: (1st Regression – t statistic for Bill 70) Manufacturing = -4.1, Retail = -0.6; (2nd Regression – t statistics for 1980, 1981 and 1982+) Manufacturing = -2.6, -3.0, -5.0, Retail = -0.6, -0.3, -0.5-
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
No. Since statistical significance was achieved in the manufacturing sector regression analyses, it is clear there was enough statistical power. As for the retail sector, if we use the rule of ten observations per predictor variable, we can see we have enough power. Specifically, five predictor variables are used in the first retail sector regression analysis and eight in the second analysis. Therefore, we would need 10*5 observations for the first analysis (n=50) and 10*8 for the second (n=80). Since there are 201 retail firms included in the analysis, the study was sufficiently powered to detect an association if one existed.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
The study does have a number of limitations that should be raised. The authors suggested that a number of other variables may have had an impact on OHS performance and may have biased the findings; one major confounder would be a change in workers' compensation administration and they mention that over the study period there was a shift to the New Experimental Experience Rating (NEER) system in Ontario but that only one rate group (plastics) entered the program. The reviewers also identified some weaknesses in the statistical analysis (a Poisson regression would have been a better approach - or the normality of the distribution of the injury/illness frequencies should have been provided to justify the use of multiple regression), potential for selection bias, as well as the possibility of the effect of confounding variables.	

Table H.1.5.

Study #5: Nytro, K., Saksvik, P., and Torvatn, H. Organizational prerequisites for the implementation of systematic health, environment and safety work in enterprises, *Safety Science*. 1998; 30(3): 297-307.

Research Question:	
To learn more about the necessary requirements (organizational factors) for achieving successful implementation of Internal Control (IC) of health, environment and safety (HES).	
Experimental Design:	
BEFORE-AFTER STUDY	
Cross-sectional surveys done in 1993 (shortly after regulation) and in 1996.	
Sample Characteristics:	
In each of 1993 and 1996, randomized quota samples of private and public enterprises were selected to represent the distribution of Norwegian firms according to size, industry and geographical location. In 1993, n = 2092; in 1996 n = 1182. Refusal rate not given. Size of workplaces in 1996: 43% had <10 employees; 39% had <100; 18% had >100.	
Maintenance of Sample:	
Not applicable.	
Intervention:	
MANDATORY OHSMS	
1992 regulation in Norway claiming all enterprises should implement <u>Internal Control (IC)</u> of health, environment and safety (HES) work. Must undertake systematic actions (at enterprise level) to ensure and document that the activities for the continuous betterment of HES are accomplished in accordance with requirements specified in laws and regulations. The systematic actions are to be described as administrative procedures.	
Data Collection, Data Transformation, and Measurement Properties:	
<input checked="" type="checkbox"/> Implementation	Implementation was measured by means of a cross-sectional survey (computer-assisted telephone interviews) of enterprises. Each enterprise had one respondent. The publication does not mention how respondents were selected, but they were most often managers (34% firm managers, 15% middle management, 12% managers responsible for health, environment and safety). The authors are unclear about the derivation of the implementation measure in Table 2, but with the aid of Saksvik and Nytro (1996) and Gaupset (2000), Saksvik et al (2003) we can deduce that it is based on a self-assessment of internal control implementation status ranging from “not started” to “finished” on a five point scale. The five categories are collapsed into three: implemented, implementation in progress, not started yet. The percentage of enterprises in each category is reported.
<input type="checkbox"/> Intermediate Outcomes	
<input type="checkbox"/> Final OHS Outcomes	
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	
Statistical Tests:	
Not applicable (none conducted).	

Results:	
Implementation	Effect estimates: Implemented; implementation in progress; not started yet (1993/1996): 8%/45%; 25%/36%; 67%/19%.
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
Not applicable.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
<p>The respondents would have a vested interest in overstating their implementation, but this is unlikely to explain all of the change seen between 1993 and 1996. Data from Gaupset (2000) show that a measure of implementation constructed from responses to questions about particular activities yields similar estimates as the self-assessment. There do not seem to be alternative explanations for the observed change.</p>	

Table H.1.6.

Study #6: Pearse, W. Club Zero: Implementing OHSMS in small to medium fabricated metal product companies, Journal of Occupational Health & Safety - Australia & New Zealand. 2002; 18(4): 347-356.

Research Question:
The Club Zero project was designed to provide information about how suitable OHSMSs were for smaller companies. There were 2 objectives: 1) to trial an approach for implementing OHSMS in a network of companies 2) to evaluate effectiveness of the OHSMS in improving the management of OHS in the companies involved in the study.
Experimental Design:
BEFORE-AFTER STUDY At baseline, each company completed a self-administered survey about their OHS management; companies were rank ordered on basis of this survey. Guidelines were created and given to them, and networking began. The companies had 6 months to implement and develop their OHSMS. After 6 months, the companies were audited. Recommendations were made for improvements. After another 6-8 months, the audit was repeated.
Sample Characteristics:
<u>INDUSTRY:</u> Fabricated metal product companies. <u>SAMPLE SIZE:</u> 20 companies in southwest Sydney, Australia recruited; but 16 left at follow-up. <u>WORKPLACE SIZE:</u> 6 had 5-9 employees; 4 had 20-99 employees; 6 had >100. Of the latter group, 1 company had 350 employees and had been included because it was an “industry leader”. <u>OCCUPATIONS:</u> not reported. <u>SAMPLING METHOD:</u> Convenience sampling - through mail-outs, word of mouth and cold-calling. About 60 companies were contacted, leading to 20 willing to participate. Authors have not analyzed how representative these companies are of other metal companies in the area, but in conclusions note that only companies which had prior commitment to change joined the project in the first place.
Maintenance of Sample:
80% follow-up over 27 months: 20 recruited however, 4 were lost to follow-up (1 bankrupt, 1 moved to Thailand, 1 takeover, 1 undisclosed)
Intervention:
VOLUNTARY OHSMS

27 month study of OHSMS implementation and effectiveness funded by the WorkCover Authority of New South Wales. The intervention involved:

1. A network of individuals participating in ongoing meetings to share resources and experience, which was established at the start of the study. Ten network meetings were held, lead by the author. These involved site visits of the host company, and an outline as to how it was managing its OHS. Discussion of issues raised, how to implement certain aspects of an OHSMS. Approximately ½ the companies actively participated in network meetings.
2. OHSMS guidelines were prepared specifically for the project, based on the principles outlines in the Australian and New Zealand Standard (AS/NZS 4804-1997)(REF), and tailored for small to medium metal manufacturing companies. They were developed around the following system elements: management commitment and policy; responsibility and accountability; OHS hazard management; purchasing and contractors; OHS training and education; emergency planning; performance indicators and records; workplace injury management.
3. Implement OHSMS - after receiving guidelines, companies had 6 months to begin to implement OHSMS before first audit.
4. 1st audit involved checking documentation, observing work processes and interviewing a range of people with different roles in the company.
5. Develop and implement action plans. Companies were provided feedback from 1st audit with recommendations for improvement and companies were asked to produce action plans designed to improve OHSMS.

Data Collection, Data Transformation, and Measurement Properties:	
<input checked="" type="checkbox"/> Implementation	At baseline the self-administered survey involved 32 yes-no questions about OHS management, drawn from SafetyMAP ³⁶ . The audit at 6 and 12 months used a 77-question audit tool based on the guidelines. Questions were closed, simple and unambiguous, with Yes, No or Partial. The audit also involved checking documentation, observing work processes and interviewing a range of people with different roles in the company. Audit scores, and % change from the first to the second audit were reported, along with the companies' rank order established at baseline. The audit tool was trialed by different external auditors and was found very reliable.
<input type="checkbox"/> Intermediate Outcomes	
<input type="checkbox"/> Final OHS Outcomes	
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	
Statistical Tests:	
Not applicable.	
Results:	

³⁶ SafetyMAP. *SafetyMap: Auditing health and safety management systems*. Melbourne: Victorian WorkCover Authority., 1997.

Implementation	Effect estimates: Companies' scores rank ordered A to Q (best – worst) on basis of self-administered survey at baseline (score out of 100 on first audit; second audit; % change). A: 72; 80; 8%. B: 75; 86; 11%. C: 65, 69, 4%. D:85, 94, 9%. E: 61, 75, 14%. F (data missing – withdrew). G: 60, 69, 9%. H: 44, 58, 14%. I: 71, 86, 15%. J: 41, 52, 11%. K: 36, 49, 13%. L: 31, 34, 3%. M: 69, 73, 4%. N: 43, 58, 15%. O: 42, 45, 3%. P: 44, 55, 11%. Q: 12, 12, 0%. Mean audit scores calculated by reviewers: 1 st audit 53.2%; 2 nd audit 62.2%; % change 9%. Management system elements rank ordered (% on 1st audit, 2nd audit, % change). Performance indicators (reporting): 87, 91, 4%. Consultation and participation: 79, 88, 9%. Workplace injury management: 64, 84, 20%. Risk management (hazard control): 68, 73, 5%. Management commitment and policy: 61, 73, 12%. OHS Training and Education: 52, 59, 7%. Emergency planning: 47, 58, 11%. Performance indicators (investigation): 47, 53, 6%. Risk management (inspections): 39, 49, 10%. Responsibility and accountability: 35, 47, 12%. Purchasing and contractors: 39, 47, 8%. Performance indicators (statistics): 31, 33, 2%. Audit changes by company size categories (1st audit mean score, 2nd audit mean score, % change): Small (5-19 employees) 35.5, 42.7, 7.2%; Medium (20-99 employees) 45, 66.5, 21.5%; Medium/Large (100+ employees) 55, 79, 24%.
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
Not applicable.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
Because the sample was non-random, and because there was quite a large refusal rate (40 of the 60 who had been contacted), one cannot generalize from these results to the larger population. The influence of the networking component was not analyzed in this study.	

Table H.1.7.

Study #7: Saksvik, P. O. and Nytro, K. Implementation of internal control (IC) of health, environment and safety (HES) in Norwegian enterprises, Safety Science. 1996; 23(1): 53-61.

Research Question:	
Was the rate of absenteeism different after implementation of Internal Control regulations? Was the rate of accidents...followed by one or more days of absenteeism...different after implementation of Internal Control regulations?	
Experimental Design:	
BEFORE-AFTER STUDY Comparisons of absenteeism and accident rates were made before the intervention (1990) and after the intervention (1991 and 1st half of 1992).	
Sample Characteristics:	
<u>INDUSTRY:</u> not specified, but likely represents all private and public industries in Norway, based on information contained in related studies. <u>SAMPLE SIZE:</u> n= 2092. <u>WORKPLACE SIZE:</u> Average # of employees=44. <u>OCCUPATIONS:</u> Survey respondents were not explicitly described although authors report they were 'typically' manager representatives. <u>SAMPLING METHOD:</u> Randomized quota sampling using computer-aided telephone interviewing.	
Maintenance of Sample:	
Not applicable. Cross-sectional self-report survey with administrative data sources - not subject to loss of follow-up.	
Intervention:	
MANDATORY OHSMS Regulations on <u>internal control (IC)</u> of Health and Safety (HES) introduced in Jan. '92. Mandatory requirements for all workplaces in Norway, to introduce IC - defined by regulations as 'systematic actions (at the enterprise level) to ensure and document that the activities of health and safety control are performed in accordance with requirements specified in Acts of regulations. e.g., the Working Environment Act of 1977. The systematic actions must be described as administrative procedures.' (p. 54). IC regulations require workplaces to arrange for systematic follow-up of requirements laid down in the following acts: 1) Working Environment Act 2) Pollution Control Act 3) Legislation relating to fire and explosion hazard and fire prevention 4) Product Control Act 5) Civil Defence Act 6) Act relating to Electrical Installations and Electrical Equipment. Changes in absenteeism and accident rates (of more than 1 day of ill-health absence) were compared pre-intervention (1990) and post (1991 & first half of 1992).	
Data Collection, Data Transformation, and Measurement Properties:	
<input checked="" type="checkbox"/> Implementation	Changes as a result of introduction of IC: more clear lines of responsibility, more better risk assessment, better documentation, new strategic plans. Used a cross-sectional survey consisting of 45 questions about HES activities and workplace demographics. The questionnaire was delivered via telephone and had been piloted with a group of 500 enterprises prior to the introduction of IC. No information is provided on validity or reliability of instrument. Representatives of enterprises answered the questionnaire (the rep was usually a manager). Most data provided by the rep was based on archival data, however in a “few” cases the data was based on recall

	(no indication of how many a few was given). Questions were answered by n=915 from the full sample (n=2092), though only n=711 reported having at least started implementation.
<input checked="" type="checkbox"/> Intermediate Outcomes	Changes in HES awareness in workplaces due to introduction of IC: no statistical analyses. Used the same survey as for the implementation measure.
<input checked="" type="checkbox"/> Final OHS Outcomes	<p>Predictor variables: 1) Size of enterprise (# of employees), 2) Sector (public vs. private), 3) HES competency (Having worker HES representatives and an Occupational Health Service), 4) HES activity level prior to IC implementation (4 point scale looking at activities in four different HES areas over previous 3 years prior to IC implementation such as assessing psychosocial work environment factors, risk assessment analyses, having worked out action plans to improve work environment, etc), 5) HES result level prior to IC implementation (absenteeism and accident rates prior to IC - for 1990), 6) time pressure (relation between productivity development and changes in size of staff - using 4 point scale). 7) status of implementation of IC (defined in three ways: subjective measure (rated from 'not started' to 'finished' using five point scale), quantitatively (# of activities completed to implement IC), from a list of nine activities), qualitatively (4 point scale derived from authors own recommendations of content of implementation process - 3 items: better risk assessment, new action plans, better documentation).</p> <p>Outcome variables: 'HES results' - 1) absenteeism, 2) accident rate (accidents followed by more than 1 day of ill health) for 1991 and 1st half of 1992 compared with 1990. Changed to categorical variable: higher, stable lower. Data drawn from n=2092 in same survey as described above.</p>
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	
Statistical Tests:	
TWO Multiple regression analyses with following predictors: organizational background variables (size of enterprise; public vs. private sector; HES competency; absenteeism and accident rate level prior to IC introduction; time pressure), status of implementation of IC. Outcome variables: HES results (change in absenteeism and accident rate)	
Results:	
Implementation	Effect estimates: Descriptive stats: % change in HES practices due to IC introduction: (n=915, though only n=711 from total sample of n=2092 had at least started to implement). 58% reported clearer lines of responsibility; 48% reported more/better risk assessment; 58% reported better documentation; 42% reported new strategic plans.
Intermediate OHS Outcomes	Effect estimates: Descriptive stats: % change in HES practices due to IC introduction: (n=915, though only n=711 from total sample of n=2092 had at least started to implement). 69% change in increased HES awareness.

Final OHS Outcomes	<p>Effect estimates: The R2ed values are presented for the two regression models (absenteeism-focused and accident-focused). NOTE: These R2ed values represent how much variance explained by the ENTIRE model and NOT just the IC status variable. R2ed value of 1st Regression Model (focusing on absenteeism): 0.05. R2ed value of 2nd Regression Model (focusing on accidents) = 0.12.</p> <p>-Point estimates are provided for the variables of interest to the review (i.e., IC status as well as any significant predictors that are not the direct interests of the review). Beta coefficients of 1st Regression Model (focusing on absenteeism): IC status = 0.09, HES result level prior to IC implementation = -0.19. Beta coefficients of 2nd Regression Model (focusing on accidents): IC status = 0.05, HES result level prior to IC implementation = -0.38.</p> <p>Test statistics: : Test statistics (t values) are provided here for the variables of interest to the review (i.e., IC status) as well as any significant predictors that are not the direct interests of the review). T values for 1st Regression Model (focusing on absenteeism): IC status = 2.94, HES result level prior to IC implementation = -6.52. T values for 2nd Regression Model (focusing on accidents): IC status = 1.66, HES result level prior to IC implementation = -13.90. F-statistic for the 1st regression model: 10.2. F-statistic for 2nd regression model: 29.9.</p> <p>Statistical hypothesis: Beta coefficients for all potential predictor variables (including IC status) will be zero. Variance explained by the models will be zero.</p> <p>P-value: P values are provided here for the variables of interest to the review (i.e., IC status) as well as any significant predictors that are not the direct interests of the review). In addition, the significance of the overall model is reported. P values for 1st Regression Model (focusing on absenteeism): IC status = less than 0.01, HES result level prior to IC implementation = less than 0.001, entire model = less than 0.001. P values for 2nd Regression Model (focusing on accidents): IC status = more than 0.05, HES result level prior to IC implementation = less than 0.001, entire model = less than 0.001.</p>
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
No. 2092...if you assume you need 10 n per independent variable...still have plenty	
Were Any Harms of the Intervention Identified?	
Not applicable.	
IWH Reviewers' Comments:	
An additional outcome was identified that did not fit in to any of the other type of evidence categories we currently have available (i.e., implementation, effectiveness, FB). This outcome was the integration of IC and TQM.	

Table H.1.8.

Study #8: Torp, S., Riise, T., and Moen, B. E. Systematic health, environment and safety activities: do they influence occupational environment, behaviour and health?, *Occupational Medicine (Oxford)*. 2000; 50(5): 326-333.

Research Question:
“This study investigated the relationships between the level of implementation of IC as a predictor variable and the satisfaction of workers at motor vehicle repair garages with the HES activities and working environment of the garages, participation in HES activities, coping with musculoskeletal symptoms, occurrence of musculoskeletal symptoms and sick leave as separate outcome variables.”
Experimental Design:
CROSS-SECTIONAL STUDY
Cross-sectional study with data collected in 1996.
Sample Characteristics:
<p><u>SAMPLE SIZE:</u> 1567 managers and garage workers in 237 garages</p> <p><u>TYPE OF INDUSTRY:</u> Industry code 8111 - Automotive repair</p> <p><u>SIZE OF WORKPLACES:</u> < 10 employees: 41% 10-30 employees: 42% >30 employees: 17% Mean 18; median 12; SE 20; range 2-140 employees</p> <p><u>NUMBER OF WORKPLACES:</u> 311 initially sampled; 237 agreed to participate</p> <p><u>OCCUPATIONS:</u> mechanics 77% panel beaters or sprayers 17% supervisors or foremen 13% other 13%</p> <p><u>SAMPLING METHOD:</u> Convenience sample plus other participants, some of whom were matched group. 130 member garages of the Norwegian Association of Motorcar Dealers and Service Organization that were scheduled to participate in a management training course in health and safety were recruited, as well as 181 other Association members including 130 who had been matched to the training course participants in terms of size and geographic location. Questionnaires were distributed to supervisors/managers./ safety directors at the 311 selected garages, 250 were returned, and 237 companies agreed to participate. A total of 2174 questionnaires were distributed to managers and garage workers at the 237 firms, and 1567 were returned.</p> <p><u>OTHER:</u> 98% were male 36% were union members</p>
Maintenance of Sample:
Maintenance of the sample is not an issue due to the cross-sectional design.
Intervention:
MANDATORY OHSMS
In January 1992, <u>Internal Control Regulation</u> , a new regulation on health and safety management, came into force in Norway. All private and public companies were required to act systematically to ensure health, environment and safety (HES) activities. IC was defined as: systematic action (at the company level) to ensure and document that the activities of health, environment and safety control are performed in accordance with requirements specified in acts or regulations (such as the Working Environment Act of 1977). The systematic action must be described as administrative procedures.” Regulation emphasizes that the employer is responsible for the company’s HES activities, that the workers must be encouraged to participate in those activities, and that each company may adjust its HES management system to its needs and special risk factors. Each company is required to

document HES assessments regularly, arrange meetings between management and workers at which HES is discussed, and regularly audit its IC system. Note: in 1997, the IC was revised to give more emphasis to management commitment, employee participation, and systematic environmental assessment as opposed to content factors. - this was after the data collection

Data Collection, Data Transformation, and Measurement Properties:

<input type="checkbox"/> Implementation	
<input checked="" type="checkbox"/> Intermediate Outcomes	<p>Independent variables: Measures of IC (from self-administered managers’ questionnaires). Two such measures of IC: (1) an overall question on how the manager subjectively rated the status of the implementation of IC at the garage (5-point scale, from “do not know what IC is” (=1), “not started” (=2) to “implemented IC (=5). (2) Internal Control Index: 16 questions on IC factors relevant to most companies, e.g., “does the garage have a health and safety deputy?”(3-point scale – yes, partial, no). Missing values for variables included in IC index were given the mean score of other variables in the index. All garages had at least 50% of values filled in.</p> <p>Dependent variables: Data from garage workers’ questionnaires, all scored on 7-point scale, poor=1 to very good=7. Satisfaction with HES activities; satisfaction with the physical working environment; satisfaction with the psychosocial working environment; psychological job demands; decision authority; social support; HES-related management support; health-related support and control; workers’ participation in HES activities.</p>
<input checked="" type="checkbox"/> Final OHS Outcomes	<p>Independent variables: Measures of IC (from self-administered managers’ questionnaires). Two such measures of IC as described above.</p> <p>Dependent variables: The workers’ questionnaire contained a total of 7 questions related to musculoskeletal symptoms experienced in the last 30 days. Six of the questions were derived from the Subjective Health Complaints Questionnaire, and asked whether they had experienced pain in the head, neck, upper back, lower back, shoulders or arms; one question on knee pain was added. The severity of pain was rated on a four-point scale for all 7 items, and the scores for each body part were added together. One question on sick leave asked respondents whether they had been away from work in the last 30 days (yes or no).</p>
<input type="checkbox"/> Economic Outcomes	
<input type="checkbox"/> Facilitators/Barriers	

Statistical Tests:

Cronbach’s alpha coefficients were calculated for each index. Pearson’s correlation analysis was performed to test correlations between the two measures of IC, the size of the company, and the percentage of unionized workers at the garage. Multiple regression analyses, adjusted for company size and unionization, were performed to investigate the relationships between each IC variable and each of the intermediate and final outcome variables. For these analyses, workers from each garage were assigned the IC score or the subjective rating of IC implementation from their managers’ responses.

Results:	
Implementation	Not applicable.
Intermediate OHS Outcomes	Effect estimates: Regression coefficients for (1) internal control index; and (2) Internal control status rated by manager: Satisfaction with HES activities at garage: 0.16***; 0.15***. Satisfaction with physical working environment: 0.079*; 0.11***. Satisfaction with psychosocial working environment: 0.044; 0.086**. Psychological job demands: -0.027; 0.011. Decision authority: -0.018; 0.000. Social support: 0.074*; 0.12***. HES-related management support: 0.097**; 0.087**. Health-related support and control: 0.082**; 0.11***. Workers' participation in HES-related activities: 0.071*; 0.13***. (* p=<0.05; **p=<0.01; ***p=<0.001). The results of the multiple regression analysis showed that the IC index based on the 16 items was significantly correlated with 6 of the 9 intermediate OHS variables; these significant correlations all indicated that the IC had positive effects on these outcomes. Higher levels of correlation were obtained in the regression analysis with IC status as rated by managers (7 of 9 were statistically significant).
Final OHS Outcomes	Effect estimates: Employees at garages rated by their manager as having a higher level of IC reported significantly fewer musculoskeletal symptoms (p<0.01). No significant relationship between either of the IC measures and sick leave were found. Standardized regression coefficients for (1) internal control index; and (2) Internal control status rated by manager: Musculoskeletal symptoms: -0.026; -0.076**. Sick leave in last 30 days: -0.048; -0.013.
Cost of Intervention	Not applicable
Economic Outcomes	Not applicable
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
No. Sample size resulted in adequate statistical power.	
Were Any Harms of the Intervention Identified?	
No.	
IWH Reviewers' Comments:	
This is a correlation study of the degree of implementation of internal control regulations for health and safety and intermediate and final OHS outcomes. The authors point out that conclusions about causality are not possible because of the cross sectional design. We note some sources of potential bias, e.g., selection bias, (low response rate); social desirability bias (exclusive use of self-report measures).	

Table H.1.9.

Study #9: Yassi, A. Utilizing data systems to develop and monitor occupational health programs in a large Canadian hospital, *Methods of Information in Medicine*. 1998; 37(2): 125-129.

Research Question:	
To describe the methods used to implement this data system and some of the results attributable to utilizing databases, to guide occupational health programs in this hospital during the last 6 years.	
Experimental Design:	
TIMES SERIES	
Time series involving WCB Assessments for hospital and the overall health care industry in Manitoba. Time series covered yearly intervals 1990-1995.	
Sample Characteristics:	
The sample consists of one worksite, an acute and tertiary care teaching hospital in inner-city, Winnipeg. There are 6000 employees.	
Maintenance of Sample:	
Changes in workforce not mentioned (except increase in payroll \$25 million (17%) from 1990-95); payroll in terms of dollars gradually increased over the period 1990-92 and then leveled from 1992-95. Thus, there seemed to be no major changes.	
Intervention:	
VOLUNTARY OHSMS	
The hospital's Department of Occupational and Environmental Medicine adopted a new risk assessment/risk management approach to occupational hazards (biological/chemical; physical; ergonomic; psychosocial). It consisted of: improved record-keeping (i.e., coding injuries); more systematic data collection and analysis, assisted by databases; hazard identification; risk assessment; planning programs to address risks; defining the programs' objectives and standards; assigning the responsibility for particular programs to particular individuals; program evaluation or surveillance; performance measurement; and continuous improvement; economic evaluation.	
Data Collection, Data Transformation, and Measurement Properties:	
<input type="checkbox"/> Implementation	
<input type="checkbox"/> Intermediate Outcomes	
<input type="checkbox"/> Final OHS Outcomes	
<input checked="" type="checkbox"/> Economic Outcomes	One measure is the hospital's Worker's Compensation Board premium rate (\$ per 100\$ payroll). A second measure is the accumulated savings in worker's compensation payments from 1991-95. This was calculated by first applying the hospital's 1990 premium rate (\$1.51 per \$100 payroll) to each of the payroll values for the five years subsequent, in order to estimate what the premium payments would have been in the absence of the program. The amounts actually paid for each of those years were then subtracted from the respective estimated amounts, and the differences were the estimated annual savings.
<input type="checkbox"/> Facilitators/Barriers	

Statistical Tests:	
Not applicable: No statistical tests.	
Results:	
Implementation	Not applicable
Intermediate OHS Outcomes	Not applicable
Final OHS Outcomes	Not applicable
Cost of Intervention	Not applicable
Economic Outcomes	Effect estimates: Premium rate declined from 1.51 (pre-intervention 1990) to 1.13 (post-intervention 1992-5 average). Accumulated savings 1991-1995. \$2,866,000.
Facilitators/Barriers	Not applicable
Did the Design Lack Statistical Power?	
Not applicable.	
Were Any Harms of the Intervention Identified?	
Increase in staff abuse during study period, but cause not established.	
IWH Reviewers' Comments:	
Health care industry's premium rate declined to a lesser extent, thereby controlling for any compensation system-related changes. The estimate of financial benefits could have been improved by including a sensitivity analysis on the premium rate assumed for the absence of intervention for the 1991-95 period. Other data in the paper and supporting papers support the interpretation that the financial benefit is attributable to prevention efforts.	

Table H.2: Presence of OHSMS Elements

OHSMS ELEMENT	VOLUNTARY				MANDATORY				
	<i>own</i>	<i>own</i>	<i>SafetyMap</i>	<i>WorkCover</i>	<i>IC</i>	<i>IC</i>	<i>IC</i>	<i>LSST</i>	<i>Bill 70</i>
	Yassi 1998	Edkins 1998	Alsop* 1999	Pearse 2002	Nytro* 1998	Saksvik* 1996	Torp* 2000	Dufour* 1998	Lewchuk* 1996
<i>INITIATION (OHS Inputs)</i>									
Management Commitment and Resources	▲	●	●	●	●	●	●	▲	▲
•Regulatory Compliance and OHSMS Conformance	▲	○	●	○	●	●	●	●	●
•Accountability, Responsibility, and Authority	?	▲	●	●	●	●	●	●	●
Employee Participation	?	●	●	●	●	●	●	●	●
<i>FORMULATION (OHS Process)</i>									
Occupational Health and Safety Policy	?	?	●	●	?	?	?	○	○
Goals and Objectives	●	▲	▲	?	●	●	●	○	○
Performance Measures	●	?	●	●	●	●	●	○	▲
System Planning and Development	●	▲	●	?	●	●	●	○	○
•Baseline Evaluation and Hazard/Risk Assessment	●	●	●	●	●	●	●	●	○
OHSMS Manual and Procedures	?	?	●	?	●	●	●	○	○
<i>IMPLEMENTATION/OPERATIONS (OHS Process)</i>									
Training System	?	▲	●	●	▲	▲	▲	●	▲
•Technical Expertise and Personal Qualifications	●	●	?	?	●	●	●	▲	▲
Hazard Control System	●	●	●	●	?	?	?	●	▲
•Process Design	▲	○	?	?	?	?	?	○	▲
•Emergency Response	?	▲	▲	●	?	?	?	▲	○
•Hazardous Agent Management	●	○	●	●	●	●	●	▲	●
Preventive and Corrective Actions	●	●	●	●	●	●	●	▲	●
Procurement and Contractor Selection	?	○	●	●	?	?	?	○	▲
<i>EVALUATION (OHS Feedback)</i>									
Communication System	?	●	▲	?	●	●	●	▲	●
•Document and Record Management System	●	●	?	?	●	●	●	▲	▲
Evaluation System	●	●	●	▲	?	?	?	○	▲
•Auditing and Self-Inspection	▲	▲	●	▲	●	●	●	○	▲
•Incident Investigation and Root Cause Analysis	●	▲	▲	▲	▲	▲	▲	▲	▲
•Health/Medical Program and Surveillance	●	○	▲	○	▲	▲	▲	●	●
<i>IMPROVEMENT/INTEGRATION (Open System Elements)</i>									
Continual Improvement	●	●	●	?	●	●	●	○	○
Integration	?	●	●	○	?	?	?	○	○
Management review	?	?	●	?	●	●	●	○	○

●	present
▲	partially present
○	not present
?	unknown

* Additional documents and/or personal communication were used to determine presence of OHSMS elements

Appendix I - Summary of Inclusion Criteria

Table I.1 Criteria for Inclusion of Studies

	Inclusion	Exclusion
Publication Type ³⁷	<ul style="list-style-type: none"> • Journal articles • Book chapters • Dissertations • Reports • Internet publications containing a report on relevant research. 	<ul style="list-style-type: none"> • Magazine and newspaper articles • Conference proceedings • Newsletters
Population of interest	<ul style="list-style-type: none"> • Workplaces 	<ul style="list-style-type: none"> • Non-workplace settings
Nature of Intervention	<ul style="list-style-type: none"> • OHSMS intervention aimed at workplace or extra-workplace level with the following considerations: <ul style="list-style-type: none"> • Interventions targeted at least two (2) or more OHSMS elements including at least one in the management domain. • Systems focused on either occupational health or occupational safety • Primary prevention was a major component • OHSMSs were indicated by a reference to a specific type of OHSMS (e.g., OHSAS 18001), a specific OHSMS legislation (e.g. Internal Control) or by explicit definition. 	<ul style="list-style-type: none"> • OHSMSs without primary prevention as a major component (i.e., those focusing only on disability or health services). • OHSMS interventions that were not indicated by a reference or explicit definition. • Interventions looking at two or less system elements without at least one element in the management domain.
Types of Evidence	<ul style="list-style-type: none"> • Studies examining one of the three following issues were included: <ol style="list-style-type: none"> i. Implementation of OHSMS ii. Effectiveness of OHSMS interventions iii. Facilitators/Barriers to OHSMS implementation or optimal outcomes following OHSMS implementation. 	<ul style="list-style-type: none"> • Any other type of evidence.
Outcomes	<ul style="list-style-type: none"> • For implementation studies: <ul style="list-style-type: none"> • A quantitative measure of change in the level or intensity of the OHSMS was required. 	<ul style="list-style-type: none"> • OHSMS implementation studies which: <ol style="list-style-type: none"> i. Did not report on a corresponding extra-workplace or workplace intervention. ii. Failed to report a quantitative measure of change in the level or intensity of the OHSMS.

³⁷ Although a variety of publication types were permitted, the seven databases abstracted primarily peer-reviewed, published literature.

	Inclusion	Exclusion
	<ul style="list-style-type: none"> • For effectiveness studies: <ul style="list-style-type: none"> • A quantitative measure of any of the following: intermediate, final, and economic outcomes <ul style="list-style-type: none"> i. <i>Intermediate OHS outcomes</i> include changes in knowledge, beliefs, values, perceptions, behaviours, hazards, or risks ii. <i>Final OHS outcomes</i> include changes in injury/illness statistics, employee quality of life, product quality, productivity, property damage, iii. <i>Economic outcomes</i> include changes in the costs associated with employee illness/injury. 	<ul style="list-style-type: none"> • Effectiveness studies without a quantitative measure of intermediate, final, or financial outcomes.
Comparisons in Outcome studies	<ul style="list-style-type: none"> • A comparison of outcome measures was required. • These comparisons could be made within or across workplaces and could be either: <ul style="list-style-type: none"> i) Between the presence and absence of an OHSMS intervention, or ii) Between OHSMS interventions of different intensities. 	<ul style="list-style-type: none"> • Implementation and effectiveness studies without a comparison of outcome measures.
Facilitator and barrier studies	<ul style="list-style-type: none"> • A quantitative or qualitative study of facilitators or barriers to either <ol style="list-style-type: none"> 1. OHSMS implementation, or 2. Optimal outcomes following OHSMS implementation, • Which may have involved: <ul style="list-style-type: none"> i. The OHSMS itself, ii. The workplace, or iii. The environment external to the workplace. • Studies examining facilitators/barriers of optimal outcomes following OHSMS implementation required a specification of the level of OHSMS implementation. 	<ul style="list-style-type: none"> • Facilitator/Barrier studies reflecting solely one expert's opinion. • Facilitators/barrier studies of optimal outcomes following OHSMS implementation without a specification of the level of OHSMS implementation.