Review

The effectiveness of occupational health and safety management system interventions: A systematic review

Lynda S. Robson a,*, Judith A. Clarke a, Kimberley Cullen a, Amber Bielecky a, Colette Severin a, Philip L. Bigelow a,b, Emma Irvin a, Anthony Culyer a,c, Quenby Mahood a

a Institute for Work and Health, 481 University Ave., Ste. 800, Toronto, ON, Canada M5G 2E9
b Department of Public Health Sciences, University of Toronto, 6th Floor, Health Sciences Building, 155 College Street, Toronto, ON, Canada M5T 3M7
c Department of Economics and Related Studies, University of York, Heslington, York YO1 5DD, England, UK

Received 23 December 2005; received in revised form 12 May 2006; accepted 6 July 2006

Abstract

A variety of OHSMS-based standards, guidelines, and audits has been developed and disseminated over the past 20 years. A good understanding of the impact of these systems is timely. This systematic literature review aimed to synthesize the best available evidence on the effects of OHSMS interventions on employee health and safety and associated economic outcomes. Eight bibliographic databases covering a wide range of fields were searched. Twenty-three articles met the study’s relevance criteria. Thirteen of these met the methodological quality criteria. Only one of these 13 original studies was judged to be of high methodological quality; the remainder had moderate limitations. The studies’ results were generally positive. There were some null findings but no negative findings. In spite of these promising results, the review concluded that the body of evidence was insufficient to make recommendations either in favour of or against OHSMSs. This was due to: the heterogeneity
of the methods employed and the OHSMSs studied in the original studies; the small number of studies; their generally weak methodological quality; and the lack of generalizability of many of the studies.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Management system; Systematic literature review; Audit; Intervention; Effectiveness

Contents

1. Introduction .......................................................... 331
   1.1. Defining OHSMSs ........................................... 331
   1.2. Voluntary and mandatory OHSMS initiatives .......... 332
   1.3. Background literature and research justification .... 333
   1.4. Conceptual framework of the review ................. 334
2. Methods .............................................................. 335
   2.1. Literature search ......................................... 335
   2.2. Selection of studies ...................................... 335
   2.3. Quality appraisal (QA) .................................. 336
   2.4. Data extraction (DE) ..................................... 337
   2.5. Evidence synthesis ....................................... 337
3. Results .............................................................. 338
   3.1. Publications identified through the literature search and quality assessment ... 338
   3.2. Quality of the literature investigating OHSMSs ..... 338
   3.3. Findings on the implementation and effectiveness of voluntary OHSMSs .... 340
       3.3.1. Description of the studies on the implementation and effectiveness of voluntary OHSMSs .... 340
       3.3.2. Findings on the implementation of voluntary OHSMSs ................ 341
       3.3.3. Findings on the intermediate outcomes of voluntary OHSMSs ....... 341
       3.3.4. Findings on the final outcomes of voluntary OHSMSs ............. 342
   3.4. Findings on the implementation and effectiveness of mandatory OHSMSs .... 342
       3.4.1. Description of the studies on the implementation and effectiveness of mandatory OHSMSs .......... 342
       3.4.2. Findings on the implementation and effectiveness of the IC legislation in Norway .................. 342
   3.5. Evidence synthesis ....................................... 344
       3.5.1. Evidence for the effectiveness of voluntary OHSMS interventions ...... 344
       3.5.2. Evidence for the effectiveness of mandatory OHSMS interventions ..... 346
4. Discussion ........................................................... 347
   4.1. Identifying and addressing research gaps ............ 347
   4.2. Strengths of the review .................................. 348
   4.3. Limitations of the review ............................... 348
5. Conclusions ........................................................ 349
Acknowledgements .................................................... 350
References ............................................................ 350
1. Introduction

The concept of an occupational health and safety management system (OHSMS) has become common over the past 20 years. A variety of OHSMS-based standards, guidelines, and audits have been developed (British Standards Institution, 1996, 1999; Dalrymple et al., 1998; Frick et al., 2000; Gallagher et al., 2003; Grayham and del Rosario, 1997; HSE, 1997; ILO, 2001; Standards Australia and Standards New Zealand, 1997) within the public, private and not-for-profit sectors and many have been adopted by workplaces. Some countries, including Canada, are in the process of developing management standards for occupational health and safety. An understanding of the impact of these systems is therefore timely.

This systematic literature review investigates the effectiveness of mandatory and voluntary OHSMS interventions on employee health and safety and associated economic outcomes. The scope of the topic was developed through formal and informal consultations with representatives of employers, labour, and several public and not-for-profit institutions. A secondary aim of the review is to characterize the content and methodology of the existing research literature on OHSMSs. The purpose here is to identify gaps and weaknesses in the literature, whose addressing through future research could help inform the primary research question.

There are no other systematic literature reviews on the topic of OHSMSs, although some narrative reviews exist (Frick et al., 2000; Gallagher et al., 2003; Saksvik and Quinlan, 2003; Walters, 2002). A systematic literature review uses explicit, thorough methods to identify, select, appraise and synthesize a set of research studies on a well-defined topic. This methodology makes the findings of a systematic review less vulnerable to the biases of a single researcher than those of a narrative review. Systematic reviews aid decision-makers by sifting through an enormous literature to find the high-quality studies and to synthesize them.

This paper will first discuss some of the concepts and background literature relevant to the review.

1.1. 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.”

It is not clear from the definition whether the management system includes only management components or technical/operational components as well. This problem of demarcating the scope of a management system has also been noted by Nielsen (2000):

“OHSMS systems are not, of course, a well-defined set of management systems. Indeed there are not clear boundaries between OHS activities, OHS management, and OHSMS systems.”
OHSMSs, commonly understood, are distinguishable from traditional OHS programs by being more proactive, better internally integrated and by incorporating elements of evaluation and continuous improvement. Some OHSMS documents (e.g., Chemical Industries Association, 1995; HSE, 1997; ILO, 2001) explicitly ascribe their basic source as the Plan-Do-Check-Act model of continuous quality improvement made famous by W. Edwards Deming (Tortorella, 1995). In contrast, traditional OHS programs have relatively less activity corresponding to the Check and Act domains of the Deming Model. Furthermore, action tends to be in response to workplace accidents, legislation, or enforcement, rather than proactive.

Redinger and Levine (1998) considered what constitutes an OHSMS in detail. After reviewing 13 publicly available management system documents for occupational health and safety, environment, or quality, they selected four of the most comprehensive from which to construct an integrative, universal OHSMS model containing 27 elements. The 16 primary elements of their model are:

- management commitment and resources,
- employee participation,
- occupational health and safety policy,
- goals and objectives,
- performance measures,
- system planning and development,
- OHSMS manual and procedures,
- training system,
- hazard control system,
- preventive and corrective action system,
- procurement and contracting,
- communication system,
- evaluation system,
- continual improvement,
- integration,
- management review.

1.2. Voluntary and mandatory OHSMS initiatives

OHSMS initiatives are either mandatory or voluntary. Mandatory OHSMSs arise from government legislation and their use is enforced 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. Their use is not required by governments; instead, incentives are sometimes offered by governments or insurance carriers to organizations that voluntarily adopt particular OHSMSs.

Many voluntary OHSMSs, especially those marketed through commercial industries, are most frequently observed in large companies. They are characterized by being more thoroughly specified, and as a result, are considered to be too complex for the majority of (smaller) employers (Frick and Wren, 2000). Voluntary OHSMS schemes marketed through public agencies, however, target not only large companies but also smaller ones (Frick and Wren, 2000). These schemes 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. Mandatory OHSMSs are simpler in terms of the demands placed on organi-
zations, since they are intended for all or most workplaces, including small ones.

1.3. Background literature and research justification

There is a body of research on the correlates of low injury rates in organizations (e.g.,
Cohen, 1977; Habeck et al., 1998; LaMontagne et al., 1996; Reilly et al., 1995; Shannon
et al., 1996). These studies suggest which potential elements of an OHSMS are important
by identifying those that are correlated with low injury rates. Other studies (Mearns et al.,
2003; Simard and Marchand, 1994) have developed researcher-defined measures of
OHSMSs and shown that a more developed OHSMS is correlated with a lower injury rate.

While these studies are a valuable contribution to the literature, they cannot tell us what
the likely effect is of a particular type of OHSMS intervention in a particular type of work-
place. They are limited by their cross-sectional design and the lack of an OHSMS interven-
tion variable.

A separate stream of research and practice related to safety management systems in
high hazard and high reliability operations, such as in the nuclear, chemical process,
and airline, rail and marine transportation industries, has also developed (e.g., Figuera
of a safety management system overlaps with that of an OHS management system, but
is generally distinct. For instance, the scope of concern for a safety management system,
unlike that of an OHSMS, extends beyond workers to include the physical work environ-
ment and the surrounding community. Furthermore, the scope of an OHSMS covers a
broad range of workers health concerns, in contrast to that of a safety management system
which focuses on preventing traumatic injuries related to the loss of control of processes.

Many OHS practitioners presume that OHSMS interventions will be effective in lowering
injury and illness. After all, OHSMS standards and guidelines synthesize expert knowledge,
much of which is consistent with the research cited in the first paragraph of this section.
However, the effectiveness of interventions cannot be presumed in all cases. Indeed, the fail-
ure rate of quality management systems has been documented as ranging from 67% to 93%
(Gardner, 2000). There is reason to expect that the failure rate of OHSMSs would be at least
as high. Typically, the level of management commitment to high product or service quality is
higher than to employee health and safety. The effectiveness of mandatory OHSMS strategies
has also been doubted by Quinlan and Mayhew (2000) in light of the current trends of the
globalization of business, the casualization of the labour force and declining unionization.

Some criticisms of OHSMSs have also emerged in more academic circles. Suggested del-
eterious effects of OHSMSs in general or of particular types of OHSMSs have included:
the weakening of external regulatory approaches (Bennett, 2002); a false sense of security
derived from the presence of a formal OHSMS (Gallagher et al., 2003); the development of
blame-the-worker attitudes (Nichols and Tucker, 2000; Wokutch and VanSandt, 2000);
and a shift in the power balance away from workers and toward management (Lund,
2004; Nichols and Tucker).

The success of OHSMSs is likely to be dependent on the nature of the intervention,
characteristics of the workplace and characteristics of the external environment. This
review synthesizes the research results on the effectiveness of a wide range of OHSMS
interventions under a wide range of conditions.
1.4. Conceptual framework of the review

The conceptual framework underlying the review is depicted in Fig. 1. The review included interventions directed at developing an OHSMS in one or more workplaces. It therefore included studies of extra-workplace initiatives arising from 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 or internally.

A minimalist operational definition of an OHSMS intervention was adopted to screen the research literature. For a study to be included, 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.

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 rather than of its effects in workplaces. Second, implementation information would allow one to distinguish between two possible explanations for an absence of effect, namely poor intervention content or poor implementation of the intervention. It is possible for a well-conceived intervention to fail through poor implementation.

One could measure implementation in extra-workplace OHSMS initiatives (e.g., legislative initiatives) through indicators of extra-workplace implementation (e.g., labour inspector orders pertaining uniquely to the legislation). In this review, however, implementation was considered only at the workplace level; specifically, a change in the state of the workplaces’ OHSMSs.

Final outcomes were identified by considering the ultimate purpose of OHSMS interventions. This would be improved employee health and safety for many stakeholders. For others, the associated economic benefits were also of interest. On these grounds, two types of final outcomes were studied in the review: OHS outcomes and economic outcomes. Examples of final OHS outcomes are changes in rates of employee injury or illness. Examples of economic outcomes are changes in workplace workers’ compensation premium rates and workplace productivity.

![Fig. 1. Conceptual framework for the review on OHSMSs.](image-url)
Intermediate OHS outcomes were considered to be outcomes of secondary interest, though they are proxies potentially for final OHS outcomes. They are changes in mediators between the OHSM and final OHS outcome, such as safety climate; employee knowledge, beliefs, values or perceptions; employee behaviours; or OHS hazards.

2. Methods

The methodological steps were: (1) search of the literature; (2) selection of relevant studies by applying inclusion and exclusion criteria; (3) appraisal of the quality of the research evidence in the studies; (4) extraction of the higher quality evidence from the studies; and (5) a synthesis of the higher quality evidence.

2.1. Literature search

Eight 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), CCInfoWeb (consisting of NIOSHTIC-2, HSELINE, and OSHLINE), Safety Science and Risk Abstracts (SSRA, from 1981), EconLit (from 1969), and American Business Inform (ABI, from 1918). 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.

The search strategy combined two sets of keywords using an “AND” strategy. The first set of keywords focused on a set of 30 OHSM terms (e.g., health and safety management system(s), safety management system(s), systematic occupational health and safety management, OHS program(s), OHSAS 18001, OHSMS, BS8800, International Safety Rating System(s), safety program(s), occupational health/safety standard(s), safety and health legislation). The keywords in the second set included 35 evaluation or OHS effect terms (e.g., evaluat(ing/ion/ions/e/es/ed), program/me evaluat(ing/ion/ions/e/es/ed), implementat(ing/ion/ed/es), effect(s), impact(s), climate, culture, perception(s), behavio(u)r, workplace injur(y/ies), injury, occupational health, occupational exposure, occupational accident, compensation cost(s), compensation claims cost(s), time loss/lost). The full search strategy is available in Robson et al. (2005) or from the authors. The terms within each group were combined using an “OR” strategy. When possible, the titles, abstracts, case registry, and subject headings were searched for keywords. The search strategy was simplified for CCInfoWeb and ABI, which had simpler search capacities.

After merging the citations from the electronic search of the eight databases and removing duplicates, 4837 remained for inclusion in the review. Four additional sources were used to identify citations of potential relevance: the reference lists of all journal articles meeting the eligibility criteria; the reference lists of pertinent reviews published in the form of books, book chapters, or journal articles; bibliographies requested from experts in the area of OHSMs; and the personal files of the authors.

2.2. Selection of studies

Titles, key words and abstracts of each article were independently screened by two reviewers. Full text articles were retrieved for those studies appearing to meet the eligibility criteria, and for those where the information in the title, abstract, and key words was
insufficient for exclusion. Upon retrieval of the full text article, the eligibility of a study was determined again through a pair of reviewers. A consensus method was used to resolve any disagreements between the reviewers about inclusion. A third reviewer was used if the pair could not reach a consensus on eligibility.

The full set of inclusion and exclusion criteria used to determine the eligibility of studies is described below:

Publication type. Only peer-reviewed journal articles were included.
Population of interest. Workplaces could be located anywhere in the world.
Nature of intervention. An OHSMS intervention, initiated at either the workplace level or extra-workplace level was required. An OHSMS intervention was 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 initiatives (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 demonstrated that it was 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 had to have been reported or referenced and to have qualified as that of an OHSMS (i.e., two or more system elements (Redinger and Levine, 1998) were specified, at least one of which was in the management domain rather than the activity/technical domain).

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) with primary prevention as the main focus. Management system interventions focusing on disability or health services were excluded. 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 included.

Type(s) of evidence. Studies had to examine either OHSMS implementation or the effectiveness of OHSMS interventions.

Implementation studies were required to have a quantitative measure of change in the level or intensity of the OHSMS. Effectiveness studies were required to have a quantitative measure of one of the following outcomes: intermediate OHS outcomes (e.g., changes in knowledge, beliefs, values, perceptions, behaviours, hazards, or risks); final OHS outcomes (e.g., changes in injury/illness statistics or employee quality of life); or economic outcomes (e.g., changes in the costs associated with employee illness/injury).

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.

2.3. Quality appraisal (QA)

Studies meeting the eligibility criteria were assessed for methodological quality using a process developed by the authors based on previous work (Abenhaim et al., 2000; Côté
et al., 2001; Drummond and Jefferson, 1996; Franche et al., 2005; Jadad, 1998; Kuhn et al., 1999; Schulpher et al., 2000; Tompa et al., 2004; van Tulder et al., 2003; Zaza et al., 2000). The method developed for this review emphasized parsimony with an aim to streamlining the consensus procedure. The questions focused on internal validity.

The five quality appraisal questions were:

(i) Are you confident that the means of selecting and maintaining the sample minimized bias?
(ii) Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?
(iii) Are you confident that the measurement methods did not introduce bias to the corresponding findings?
(iv) Were appropriate statistical tests applied to the data?
(v) Are you confident that there are no additional potential sources of bias in the estimate of implementation/effectiveness not already captured in the previous questions?

Multiple choice response options were provided (e.g., yes, partially, no, unclear/unknown from information provided).

Reviewers gave an overall summary rating of the quality of evidence provided by the study. They selected one of the following four categories:

- very low (serious limitations),
- low (major limitations),
- moderate (moderate limitations), and
- high (no or minor limitations).

The methodological quality of the evidence of each study was rated independently by two reviewers, who then met for consensus. If consensus could not be reached an additional reviewer was consulted. The quality assessment was carried out separately for implementation, intermediate, final OHS and economic outcomes in each study.

2.4. Data extraction (DE)

Evidence rated during the quality appraisal step as moderate or high was extracted for the evidence synthesis. Pairs of reviewers independently extracted data from the included studies, using a standard form, and then met to reach consensus. Data were extracted on the research question, intervention, study design, study population, results, and statistical analyses.

2.5. 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. (Briss et al., 2000; Franche et al., 2005; GRADE Working Group, 2004; Kuhn et al., 1999; Tompa et al., 2004; van Tulder et al., 2003). Criteria for these algorithms are customarily based on study design, quality of research, consistency of the results, and number of studies.
This review did not adopt an explicit algorithm at the outset, since there is a lack of consensus in the field as to which synthesis algorithm is best. In addition, it was thought premature to base an algorithm upon a newly developed quality assessment tool. Instead, a summary statement was synthesized in the style of a traditional narrative review, after considering the same aspects of evidence as do the algorithms. Such an approach is permissible in best-evidence syntheses (Slavin, 1995) and has been used in other systematic reviews in this field (see special issue of American Journal of Preventive Medicine, 2000).

3. Results

3.1. Publications identified through the literature search and quality assessment

Following a review of titles and abstracts, and initial screening of full papers where necessary, 23 studies were identified which met the inclusion criteria. Their citations were found through the following sources:

- bibliographic databases (18),
- citations in reviews (3),
- authors’ personal files (1),
- suggestion by external reviewer of interim report (1).

The 23 eligible studies were then assessed for methodological quality. Thirteen contained evidence given an overall rating of “moderate” or “high” and proceeded to the next step in the review, data extraction. The other ten studies, whose evidence received an overall rating of “low” or “very low,” (Anonymous, 1993, 1994; Bolton and Kleinsteuber, 2001; Dotson, 1996; Eisner and Leger, 1988; Everley, 1997; Kjellén et al., 1997; Lanoie, 1992; Mikkelsen and Saksvik, 2004; Nichols, 1990) were excluded from further review. (It should be noted that this overall rating was determined from the point of view of this review’s question about OHSMS effectiveness. The studies excluded at this stage were not necessarily low quality studies from the point of view of their own research questions.)

An overview of the characteristics of the 13 studies proceeding to data extraction is shown in Table 1. Seven were of voluntary OHSMS interventions, all of which took place in English-speaking countries. Many of the voluntary interventions were unique to the workplace(s) under study. Six studies were of mandatory OHSMS interventions, with four dealing with the Norwegian Internal Control legislation introduced in 1992. The other two were initiatives of Canadian provincial legislatures in the late 1970s.

While the studies of mandatory OHSMSs were of large numbers of workplaces, the studies of voluntary OHSMSs were of small numbers of workplaces and often of just one workplace. The most common effect studied was implementation but there were also some findings on intermediate OHS outcomes, final OHS outcomes and economic outcomes.

3.2. Quality of the literature investigating OHSMSs

The methodological quality of the OHSMS intervention literature was generally judged to be weak. Only the study of implementation by LaMontagne et al. (2004) was assessed as “high” quality.
Table 1
Summary of the 13 studies providing the highest quality evidence

<table>
<thead>
<tr>
<th>Authors (year of publication)</th>
<th>OHSMS intervention(^a)</th>
<th>Country</th>
<th>Industrial sector</th>
<th>Number of organizations in analysis</th>
<th>Type of data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voluntary interventions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alsop and LeCouteur (1999)</td>
<td>Organization’s own; based on international and national standards</td>
<td>Australia</td>
<td>Public administration</td>
<td>1 organization; multiple sites</td>
<td>Economic</td>
</tr>
<tr>
<td>Bunn et al. (2001)</td>
<td>Organization’s own</td>
<td>United States</td>
<td>Manufacturing</td>
<td>1 corporation; multiple sites</td>
<td>Implementation Final OHS Economic Intermediate OHS</td>
</tr>
<tr>
<td>Edkins (1998)</td>
<td>Organization’s own; based on other programs in industry</td>
<td>Australia</td>
<td>Transportation, air</td>
<td>1 airline; 2 sites</td>
<td>Final OHS Economic</td>
</tr>
<tr>
<td>LaMontagne et al. (2004)</td>
<td>Researcher</td>
<td>United States</td>
<td>Manufacturing</td>
<td>15 work sites</td>
<td>Implementation</td>
</tr>
<tr>
<td>Pearse (2002)</td>
<td>Organization’s own</td>
<td>Australia</td>
<td>Manufacturing</td>
<td>16 companies</td>
<td>Implementation</td>
</tr>
<tr>
<td>Yassi (1998)</td>
<td>Organization’s own</td>
<td>Canada</td>
<td>Health care</td>
<td>1 hospital</td>
<td>Economic</td>
</tr>
<tr>
<td><strong>Mandatory interventions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dufour et al. (1998)</td>
<td>Quebec LSST (1979)</td>
<td>Canada</td>
<td>Manufacturing</td>
<td>All firms in most manufacturing sectors</td>
<td>Economic</td>
</tr>
</tbody>
</table>

\(a\) Abbreviations used in column: AS/NZS 4804, Australian/New Zealand Standard; HSE, Health and Safety Executive (British government).
The prevalence of methodological weakness was mainly a consequence of the experimental designs typically employed. Only three of the 23 eligible studies (Edkins, 1998; LaMontagne et al., 2004; Yassi, 1998) employed a comparison group of any kind; the remaining studies used after-only, cross-sectional, before-after or time series designs.

The quality assessment process rated studies against four main criteria concerned with the study’s internal validity: sampling, confounding, measurement, and statistics. The most common weakness in the 23 eligible studies was poor evaluation of and control for confounders, with 11 studies failing to meet the corresponding criterion at least partially. The second most common weakness, found in six studies, concerned potential bias introduced through measurement methods.

The difficulty in meeting the confounding criterion was inherent in the studies with no comparison group. In some cases, the publications lacked explicit consideration of potential confounders. In single workplace studies, co-interventions were a major concern. For example, some authors mentioned the existence of coincident interventions such as cost control (Bunn et al., 2001) or restructuring (Kjellén et al., 1997). Bolton and Kleinsteuber (2001) stated that “many other variables, including budgetary considerations and the scope and nature of programmatic tasks and activities” could have contributed to the injury rate trend they observed. For other single workplace studies, co-intervention was a concern simply because the authors omitted any explicit consideration and discounting of it. No doubt in some cases, there were no major, relevant changes, but the lack of any report caused the study to fail on the criterion concerned with confounding.

Insufficient information was also a reason why some studies failed to satisfy the criterion regarding measurement methods. In other studies, self-report methods were used to measure OHSMS implementation without any consideration of the bias arising when people have a vested interest in portraying compliance.

3.3. Findings on the implementation and effectiveness of voluntary OHSMSs

3.3.1. Description of the studies on the implementation and effectiveness of voluntary OHSMSs

The review’s search for relevant literature resulted in 14 studies of voluntary OHSMSs, seven of which remained after quality assessment (Alsop and LeCouteur, 1999; Bunn et al., 2001; Edkins, 1998; LaMontagne et al., 2004; Pearse, 2002; Walker and Tait, 2004; Yassi, 1998). Only one study (Bunn et al., 2001) examined multiple outcomes (implementation, final OHS outcomes, and economic outcomes). The other studies focused on only one – implementation (LaMontagne et al., 2004; Pearse, 2002; Walker and Tait, 2004), intermediate OHS (Edkins, 1998), or economic (Alsop and LeCouteur, 1999; Yassi, 1998).

The studies involved a variety of interventions and samples. Four studies (Alsop and LeCouteur, 1999; Bunn et al., 2001; Edkins, 1998; Yassi, 1998) reported on OHSMSs developed and implemented in a single organization – respectively, a municipal government, an international manufacturer, a regional airline, and a hospital. The study by Alsop and LeCouteur (1999) concerned an OHSMS developed in the presence of and integrated with management systems for quality and the environment. The intervention reported by Bunn et al. (2001) consisted of the creation of a Health, Safety and Productivity Department, which integrated all functions concerned with employee health, safety and associated productivity. Key features of the intervention were goal-setting, performance measurement and senior management commitment. Edkins (1998) investigated a new system of hazard
identification, safety information management, and safety communication, overseen by a new safety manager. The hospital intervention (Yassi, 1998) consisted of the systematic identification, measurement, and control of OHS risks, as well as program evaluation.

The three other studies of voluntary OHSMSs involved interventions delivered to multiple organizations. Pearse (2002) described a community intervention with 20 small and medium-sized metal fabrication companies. It consisted of the development and dissemination of OHSMS guidelines adapted from the Australian/New Zealand voluntary standard (AS/NZS 4804), group networking meetings and audits. LaMontagne et al.’s (2004) research group targeted large manufacturers that were likely to use hazardous substances. Fifteen worksites completed the intervention, which included a baseline OHS program assessment and tailored consultation and educational sessions for managers and committees. Sessions emphasized upper management commitment, employee participation, improvement in the resources directed towards hazard control, and OHS training for all employees. Walker and Tait (2004) determined the impact of 90-min consultation sessions delivered by two non-profit information centres upon the OHSMSs of small- and medium-sized organizations.

3.3.2. Findings on the implementation of voluntary OHSMSs

The four studies of implementation all showed the sizeable development of OHSMSs, as determined by audit methods. The Walker and Tait study (2004) showed marked self-reported changes in terms of a policy statement being present and a risk assessment having taken place. Forty-six percent of organizations added a policy statement and 79% added a risk assessment as a result of the intervention. It must be noted that these organizations had requested assistance with establishing an OHSMS, and that these two OHSMS elements were emphasized in the information provided to the organizations.

The other three implementation studies reported positive changes in the scores from a quantitative audit. Bunn et al. (2001) reported an increase from 63% to 79% of available audit points over a 2 years period for the organization studied. Pearse (2002) and LaMontagne et al. (2004) included data in their publications that allowed effect sizes to be calculated (0.48 and 0.68, respectively). These are considered to be medium and medium–large, respectively (Cohen, 1977).

Only one of the four studies of implementation contained information relating these changes in audit scores to changes in final OHS or economic outcomes (Bunn et al., 2001). Thus, the ultimate practical significance of most of these reports remains unknown. In the study by Bunn et al. (2001), described below, there were changes in final OHS and economic outcome indicators, but it is difficult to attribute these solely to the changes in the OHSMS because of the concurrent cost containment initiatives that were also underway.

3.3.3. Findings on the intermediate outcomes of voluntary OHSMSs

As for intermediate outcomes, Edkins (1998) reported greater positive changes in the intervention group than in the comparison group, both for some self-reported measures (e.g., safety culture index) and for some objective measures (confidential hazard reports, organizational actions). Thirteen new organizational actions in safety were attributed to the intervention group, while none were attributed to the control group. As with the implementation findings, the significance of these results for final outcomes remains unknown. Furthermore, the reader is left wondering whether the large achievements in organizational actions were largely attributable to the personal qualities of the new safety manager.
rather than the entirety of the new OHSMS. Indeed, the reader is told that many of the actions addressed long standing issues.

3.3.4. Findings on the final outcomes of voluntary OHSMSs

Bunn et al. (2001) was the only study of voluntary OHSMSs to report on final OHS outcomes. They found a 24% decrease in illness/injury frequency and a 34% decrease in lost-time case rate over 3 years. It should be noted that management became accountable to the board of directors for improvement in these safety indicators, coincident with the intervention. One effect of this would have been to ensure critical management commitment for the OHSMS intervention. It could also have led to a shift in the reporting practices in the organization, such that incidents were less likely to be documented. The authors refer to “aggressive return to work programs,” which could have contributed to the reduction in lost-time case rate.

Workers’ compensation costs also decreased markedly in studies of voluntary OHSMSs. Bunn et al. (2001) found a 13% decrease in workers’ compensation cost per employee. Yassi (1998) and Alsop and LeCouteur (1999) found decreases of 25% and 52% in premium rates, respectively. However, in each of these studies, there were initiatives beyond primary prevention which likely contributed to these results. Yassi (1998) mentioned a return to modified work program. Alsop and LeCouteur (1999) referred to “targeted and sustained effort in claims and injury management” during the 2 years prior to the OHS initiative. Bunn et al. (2001) listed various cost-cutting measures including better case management and aggressive return to work programs.

3.4. Findings on the implementation and effectiveness of mandatory OHSMSs

3.4.1. Description of the studies on the implementation and effectiveness of mandatory OHSMSs

Nine studies of mandatory OHSMS interventions were identified by the review. Six remained after the quality assessment screen. Four of these were based on the Norwegian regulations on Internal Control (IC) of health, safety, and the environment (Nytro et al., 1998; Saksvik and Nytro, 1996; Saksvik et al., 2003; Torp et al., 2000). Two studies focused on regulations in the Canadian provinces of Quebec and Ontario, respectively (Dufour et al., 1998; Lewchuk et al., 1996).

The Canadian regulations were progressive for their time (the late 1970s), but in terms of contemporary approaches to integrated health and safety management are more limited and less system-oriented in their requirements. The Norwegian IC regulations, on the other hand, incorporated systematic management concepts, which were increasingly being found in best practice models in business at that time.

3.4.2. Findings on the implementation and effectiveness of the IC legislation in Norway

The IC studies demonstrated most clearly the effects of the legislation on the use of OHSMSs. There was an increase from 8% to 47% of workplaces fully implementing the IC requirements over the period 1 year to 7 years post-intervention (Nytro et al., 1998; Saksvik and Nytro, 1996; Saksvik et al., 2003). Consistent with these findings, Saksvik and Nytro (1996) found that, in 1993, a sizeable portion of workplaces credited the legislation with improvements in various aspects of an OHSMS (e.g., 25% attributed clearer lines of responsibility). Researchers also found an impact on awareness of health, safety
and the environment, an intermediate OHS outcome, with 39% of workplaces attributing an increase to the legislation.

The attempts by Saksvik and Nytro (1996) and Torp et al. (2000) to find any impact of the legislation on other intermediate OHS outcomes and on final OHS outcomes were less convincing. Saksvik and Nytro (1996) carried out multi-variable regression analyses to relate the degree of self-reported IC implementation after 1 year to self-reported changes in absenteeism and injury rates. The model for absenteeism was statistically significant, but explained only 5% of the observed variance. The model for injury rates was not statistically significant. These relatively weak results might have been partly due to the short observation period and the crude measure of change used. Respondents were asked to report on whether rates were lower/stable/higher in 1991 and the first half of 1992, compared with 1990. (The legislation came into effect on January 1, 1992.)

Torp et al. (2000) took a similar analytic approach to survey data gathered 4 years post-implementation. They found that the majority of intermediate variables tested (e.g., satisfaction with HES activities in the garage, perceived physical working environment), had a statistically significant relationship with IC status. For the two final OHS outcomes tested, musculoskeletal symptoms and sick leave, the relationship was in the expected direction but weak in both cases (standardized beta coefficients of $-0.026$ and $-0.048$, respectively), and statistically significant in only the first.

The evidence provided by Saksvik and Nytro (1996) and Torp et al. (2000) on the effects of IC implementation is weak because of the cross-sectional designs. Whether the IC implementation preceded the outcomes remains unknown and unmeasured factors related to both the degree of IC implementation and final OHS outcomes (e.g., management commitment to OHS) are alternative conceivable explanations for the findings.

### 3.4.3. Findings on the effectiveness of simple OHSMS legislation in Canada

Lewchuk et al. (1996) examined the effect of the 1979 introduction of Bill 70 legislation in Ontario on lost-time injury claim rates. They analyzed claims data from 1976 to 1989 in six manufacturing and two retail sectors. These two sectors were differently affected by the legislation, with the requirements for manufacturing being more intensive. The reviewers therefore categorized the research design as two distinct time series.

Results from regression analyses of the manufacturing sector data (controlled for employment, union status, time and sector) showed that Bill 70 had a significant effect, equivalent to an 18% decrease in lost-time injury rate. A separate analysis showed that the effects were progressively larger for 1980, 1981, and 1982+, suggesting a phasing in of implementation and perhaps a lag in effects. In contrast, in the retail sector, the effect of the legislation was small and statistically not significant.

An explanation for the differences observed between sectors might be a difference in the OHSMS requirements, especially the requirement for a joint-health-and-safety committee in only the manufacturing sector. Conflicting with this interpretation, however, is the observation from separate regression analyses that joint-health-and-safety committees formed in 1980 or later appeared to have no effect on injury rates – and yet Bill 70, introduced in 1979, appeared to have an effect. Thus, other unmeasured, industry-specific contextual factors or co-interventions might have played a role instead. For example, it seems likely that the retail sector would have received less attention from the Ministry of Labour’s inspectorate because of retail’s lower level of risk.
In the Quebec study (Dufour et al., 1998), pooled time series regression analyses based on the 3 years of annual data across the 19 manufacturing sectors were used to develop six different specifications of the authors’ theoretical equation explaining variation in total productivity growth. All of the specifications had good explanatory power (R-squared values from 0.54 to 0.77) and regression coefficients were relatively consistent. The regression coefficient for the variable “percentage of companies having prevention programs” (the OHSMS variable) was positive and statistically significant. The authors interpreted this to mean that the prevention programs reduced injuries, which in turn enhanced firm productivity.

However, there are plausible alternative explanations for the findings. 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 at the aggregate level with both (a) the more rapid development of a prevention program in response to a legislative change and (b) productivity growth. Additionally, the data on the prevalence of prevention programs in the study relied on the report of the prevention program by the workplace to the public authorities; its accuracy was unknown.

3.5. Evidence synthesis

Sections 3.3 and 3.4 described the findings of the individual studies of voluntary and mandatory OHSMSs. This section gives a higher level synthesis of each group of studies and assesses the body of evidence for OHSMS effectiveness.

The heterogeneity of the studies precluded a quantitative pooling (i.e., meta-analysis), especially because they are so few in number (van Tulder et al., 2003). The evidence was therefore synthesized qualitatively. The following discussion focuses separately on voluntary and mandatory OHSMSs and assesses the evidence available for each.

3.5.1. Evidence for the effectiveness of voluntary OHSMS interventions

The seven studies on voluntary OHSMS interventions all showed positive effects (see Table 2). One study also had some null findings; there were no negative findings. The positive effects included:

- increased OHSMS implementation over time;
- intermediate effects (e.g., better safety climate, increased hazard reporting by employees, more organizational action taken on OHS issues);
- decreases in injury rates; and
- decreases in disability-related costs (e.g., workers’ compensation costs, short- and long-term disability costs).

The falls in injury rates and workers’ compensation costs would be of practical importance to stakeholders (declines of 24–34% in injury rates and of 13–52% in workers’ compensation costs).

However, six of the seven studies had moderate methodological limitations. The one high-quality study (LaMontagne et al., 2004) provided evidence for only implementation. It is unknown how the implementation effects observed in this particular study would have affected injury rates and economic costs.

Overall, the evidence provided by these seven studies was sparse. Only four studies had findings for any of intermediate OHS, final OHS, or economic outcomes. Each involved a
Table 2
Summary of the evidence

<table>
<thead>
<tr>
<th>Authors (year)</th>
<th>Method. quality</th>
<th>Study design</th>
<th>Number of workplaces in analysis</th>
<th>Effects of the intervention</th>
<th>Implementation</th>
<th>Intermediate</th>
<th>Final OHS</th>
<th>Economic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voluntary interventions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alsop and LeCouteur (1999)</td>
<td>Moderate</td>
<td>Time series</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bunn et al. (2001)</td>
<td>Moderate</td>
<td>Time series</td>
<td>1</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edkins (1998)</td>
<td>Moderate</td>
<td>Non-randomized trial</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LaMontagne et al. (2004)</td>
<td>High</td>
<td>Randomized trial</td>
<td>15</td>
<td>+/0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearse (2002)</td>
<td>Moderate</td>
<td>Before-after</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walker and Tait (2004)</td>
<td>Moderate</td>
<td>After-only</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yassi (1998)</td>
<td>Moderate</td>
<td>Time series with comparison group</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandatory interventions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dufour et al. (1998)</td>
<td>Moderate</td>
<td>Time series, aggregated data</td>
<td>19&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lewchuk et al. (1996)</td>
<td>Moderate</td>
<td>Time series</td>
<td>636</td>
<td>+, 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saksvik and Nytro (1996)</td>
<td>Moderate</td>
<td>Cross-sectional</td>
<td>2092</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nytro et al. (1998)</td>
<td>Moderate</td>
<td>Before-after</td>
<td>1182–2092</td>
<td></td>
<td></td>
<td></td>
<td>+, 0</td>
<td></td>
</tr>
<tr>
<td>Saksvik et al. (2003)</td>
<td>Moderate</td>
<td>Time series</td>
<td>1182–2092</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Torp et al. (2000)</td>
<td>Moderate</td>
<td>Cross-sectional</td>
<td>267</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
<td>+, 0</td>
</tr>
</tbody>
</table>

<sup>a</sup> “+” indicates that the OHSMS had a beneficial effect; statistical significance was not necessarily determined; “−” indicates an effect in the opposite direction, statistical significance was not necessarily determined; “0” indicates that a statistical test was conducted and the results were not significant (n.s.; $p > 0.05$).

<sup>b</sup> For the Dufour study, the number of sectors is shown, since sector was the unit of analysis.
single enterprise, making the direct applicability of the results to other workplaces uncertain.

The single enterprise studies must also be regarded cautiously from the point of reporting bias (one favouring positive results). Researchers are known to be reticent to publish findings that demonstrate an intervention has no significant effect. Workplace representatives who champion interventions (and thus have a vested interest in their success) are likely to have an even greater tendency not to publish such findings. All four 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%) is markedly different than the rate of 67–93% 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 barriers to implementation can be presumed to be the same (e.g., management commitment, culture change).

Generalizability is also an issue in the three multiple workplace studies, all of which focused on implementation. The interventions in both Pearse (2002) and LaMontagne et al. (2004) were delivered by researchers and showed high refusal rates in the recruitment phase of 67% (Pearse, 2001) and 59%, respectively. The effects seen in the sample of firms studied are likely to be a biased estimate of the effectiveness that would be seen if the intervention had been delivered to all eligible firms. Firms willing to participate in the study would also have been likely to have had a higher degree of management commitment to OHS, which would also be predictive of successful OHSMS implementation. A similar consideration is at play for the Walker and Tait (2004) study. They reported on an intervention in the “real world”, but all of the workplaces in their sample had invited the intervention by seeking help from information centres.

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.

3.5.2. Evidence for the effectiveness of mandatory OHSMS interventions

The studies on mandatory OHSMS interventions also indicated consistently positive effects, although there were also null findings on some of the final OHS outcomes (see Table 2). The studies 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 lost-time injury rates; and
- increases in workplace productivity.

The size of the observed changes in OHSMS development (an increase from 8% to 47% of all workplaces with full implementation of OHSMS regulations) and the decline in injury rate (18%) observed are likely to be of practical importance to stakeholders.

On the other hand, all studies in this group had moderate methodological limitations. These limitations arose largely from the simple study designs employed, especially the two cross-sectional studies measuring final OHS outcomes where the direction of causality was uncertain (Saksvik and Nytro, 1996; Torp et al., 2000). In addition, it was hard
to exclude the possibility that there had been confounding or co-intervention in some studies. The most convincing finding was that of Saksvik et al. (2003), demonstrating the implementation of IC regulations in Norway (an increase from 8% to 47% of workplaces with full implementation). Unfortunately, there is a lack of credible evidence about how the Norwegian achievements in compliance affected injury rates and other outcomes.

As well as internal validity weaknesses, two of the six studies were concerned with Quebec and Ontario legislation from the late 1970s and are therefore not applicable to developed countries now.

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. Discussion

4.1. 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 OHSMS interventions on employee health, safety and economic outcomes. Moreover, the studies were seldom sufficiently rigorous methodologically to allow for great confidence in the reported findings. Their limitations also inhibit the ability to apply the results 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 reported about measurement methods and potential biases in measurement,
- samples which prevented generalization of findings (i.e., several single workplace studies).

A similar limitation has been found in the quantity and quality of the evidence on the closely related topic of safety management system effectiveness (Hale, 2003).

The scarcity of high-quality published research on the implementation/effectiveness of OHSMSs may relate to the difficulties in carrying out such research. Recruitment of workplaces to both intervention and controls is a challenge. Many workplaces are not willing to make a commitment to a large intervention like an OHSMS, let alone its evaluation. Among those that are, some might not want to risk being allocated to the comparison group. Thus, large refusal rates and some withdrawal after allocation are to be expected. Measurement presents conceptual, logistic and resource challenges because of the complexity of OHSMSs and their environment.

When conducting a controlled trial is not feasible, observational cohort studies are usually considered to be the next best option by those in the epidemiological field. To answer questions about OHSMSs using an epidemiological model, one would need to 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 and
possibly at the worker-level too. However, such research designs are very expensive and complex to implement, and as such, are rarely used. In Canada, such a research design is employed by the national survey organization Statistics Canada (2005) to examine issues such as organizational productivity, but thus far OHS is not within its scope.

4.2. 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 review has clearly eliminated a huge volume of work in finding the relatively few studies of interest and to summarize their findings. Its use of explicit, systematic methods helps ensure that this summary is relatively objective in its appraisal.

Within the parameters set by the review question and the included sources, the review team feels confident that the search has been 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. The review drew from a broad range of academic disciplines. The eight databases used in the search represent the fields of occupational medicine, occupational safety, risk management, management, occupational psychology, and sociology.

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 1). The present review therefore makes a unique contribution to the research literature.

4.3. Limitations of the review

Resource and feasibility constraints limited consideration of the evidence to the published, peer-reviewed literature identified in eight databases. The usual expectation is that the literature of highest quality is in peer-reviewed journals. Only a preliminary search and screen of other literature, i.e., that which is not peer-reviewed and published, was undertaken.

A simple search of thesis dissertations (through the Dissertations Abstract International database) revealed some sources which were eligible (Ford, 1998; Weems, 1998). However, a cursory review of their quality suggested that they would not alter the conclusions of this review in any substantial way and might not in fact pass the quality assessment screen.

Other types of “grey literature” were omitted from the review too (e.g., government reports, conference proceedings, unpublished reports by OHSMS vendors and users). This research team identified two reports of this type (Hanson et al., 1998; European Agency for Safety and Health at Work, 2002) during the course of the review. Accessing the grey literature on OHSMSs in a systematic way would be a great challenge, given the variety of potential sources and the proprietary nature of some of the knowledge.

Although a large set of search terms related to OHSMSs was used to search the bibliographic databases, it remains possible that it failed to capture some relevant studies. For example, an intervention described as a “climate intervention” might not have been detected if other terms like “safety management system,” “safety program,” “safety system”, “occupational safety standard” were not used too (see Section 2.1). Such
omissions are not likely to be many. Not only were the results of the database searches used as a source of potentially relevant citations, but also the following sources: the 26 eligible studies; recent reviews in journals and monographs, including Gallagher et al. (2003), Frick et al. (2000), Hale (2003) and Kogi (2002); personal files; and reference lists requested from experts in the field.

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; Simard and Marchand, 1994), because there had been no intervention.

The review required the study to involve intervention on at least two OHSMS elements, since the review’s focus was on 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., Cohen, 1977; Habeck et al., 1998; LaMontagne et al., 1996; Reilly et al., 1995; Shannon et al., 1996).

By requiring studies to be quantitative in nature and published as a journal articles, publications that were qualitative or theoretical in nature or published in a format outside of a journal article were excluded. The case study of an Esso gas explosion (Hopkins, 2000), based on a Royal Commission on the event, included critical comments about the OHSMS in use. The Frick et al. (2000) monograph contains further critique. Such materials were not included in this review, although some of the issues raised were summarized in Section 1.

5. Conclusions

This systematic review found a relatively small quantity of published, peer-reviewed evidence involving OHSMS interventions, despite the fact that reviewers screened 4837 studies drawn from eight 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 studies’ characteristics.

The review’s synthesis of the evidence showed mostly favourable results. There were a few null findings, but no findings of negative effects.

However, all but one of the studies included in the final synthesis had moderate methodological limitations. The studies were seldom sufficiently rigorous to allow great confidence in the reported findings. In addition, many of the findings concerned only OHSMS implementation, with no accompanying results for outcomes. Finally, the characteristics of workplace samples (e.g., single workplaces, large refusal rates) also prevent certainty about the applicability of the findings to other workplaces.

In conclusion, despite the generally positive results on the effectiveness of OHSMS interventions in the published, peer-reviewed literature, the evidence is insufficient to make recommendations either in favour of or against particular 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.
Acknowledgements

The Institute for Work and Health (IWH) is an independent, not-for-profit research organization. It receives core funding from the Ontario Workplace Safety and Insurance Board. The authors benefited from the assistance of others at IWH: methodological assistance from Jill Hayden, Selahadin Ibrahim and Victoria Pennick; library assistance from Doreen Day and Dan Shannon; and administrative assistance from Krista Nolan. Several external scientists gave generously of their time in reviewing an earlier report on this project: Kaj Frick, Clare Gallagher, Wayne Lewchuk, Michael Quinlan, Jos Verbeek, and Stefan Zimmer. Finally, other external, local stakeholders helped orient LR to the subject matter from their various viewpoints: Andy King, Ron Lovelock, John MacNamara, Ed McCloskey, Ken Taylor.

References


Standards Australia and Standards New Zealand, 1997. AS/NZS 4804: occupational health and safety management systems – general guidelines on principles, systems and supporting techniques. Standards Australia; Standards New Zealand, Homebush (NSW); Wellington.


