Impact Evaluation of Health and Safety Management in the National Health Service -

A Literature Review

Prepared by
Occupational Health and Safety Advisory Service (OHSAS)
and University of Aberdeen for the
Health and Safety Executive 2005

Impact Evaluation of Health and Safety Management in the National Health Service -

A Literature Review

Karen J M Niven PhD MSc FFOH FIOSH RSP
Occupational Health and Safety Advisory Service
(OHSAS) and University of Aberdeen
Navy House
Stuart Road
Rosyth Industry Park
Rosyth
Fife KY11 2BJ

The design of an impact evaluation of a health and safety management system in the National Health Service (NHS) was informed by a review of relevant literature. Most occupational health and safety interventions had their origins in positivism and the main recommendation has been that their scientific rigor should be improved. If this were achieved this would provide more specific evidence to support (or not) the general theme of the current evidence that health and safety interventions tend to improve health and safety.

The review identified an optimum design comprising six main aspects. These were a longitudinal design; inclusion of comparison groups; an intervention that was of interest to the NHS; a participative style; multiple measurement methods and multiple indicators of effectiveness.

This report and the work it describes were funded by the Health and Safety Executive (HSE) With additional funding/support from: Scottish Executive Health Department, Directorate of Human Resources, University of Aberdeen, Tayside Primary Care NHS Trust, and Fife Primary Care NHS Trust. Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

HSE BOOKS

ACKNOWLEDGEMENTS

My grateful thanks are extended to the following organisations and individuals without whom the literature review in this report could not have existed.

The funding bodies
The Health and Safety Executive
The Scottish Executive, Department of Health
Tayside Primary Care NHS Trust
Fife Primary Care NHS Trust

Other sources of advice and encouragement
Professor John Cairns – University of Aberdeen, Health Economics Research Unit
Dr John Cherrie – University of Aberdeen, Department of Environmental and Occupational Medicine

CONTENTS

EXECUTIVE SUMMARY

The main objective of the review was to undertake an evaluation of the literature base on intervention studies within the field of health and safety management. The purpose of this was to identify the current best-practice evidence-base so as to inform research design which could then be used to evaluate whether change could be detected in health and safety performance when a health and safety management system was introduced to the healthcare sector[1].

A noticeable general finding was that, despite regular overlapping of the publication years of the articles examined, few of the published reviews evaluated the same studies. Reasons for this are unclear but might be related to the multidisciplinary nature of the subject area and the breadth and depth of the literature.

Results from the evaluation of literature relating to evaluation of health and safety performance identified four main approaches. These included those methods, which could be used in an objective manner (such as those based on audit or incident data; subjective measures of safety climate/culture; economic evaluations; and methods that used more than one type of method.

It was intended to use a method developed by HSE, which was based on the use of incident data to evaluate health and safety performance. This method was used to link analysis of incidents to the health and safety management system of an organisation, an attribute that was not present in any of the other methods reviewed. The most suitable method that used subjective methods was identified as the Loughborough Safety Climate Assessment Toolkit, which had been based on a systems approach to organisational culture, which had been extensively validated. It was judged appropriate to include a costing aspect and also to use more than one type of method.

A systems based management intervention in the form of a workbook had been developed and it was intended to use this as a research intervention. The review considered that four factors should be considered as part of the intervention. There were: that the intervention was not overly generic; that it should be straightforward to use; that measures were taken to ensure that it produced an effect strong enough to be detected; and that consideration be given to the use of methods based on behavioural models.

The review of research designs suitable for measuring change in health and safety performance concluded that there exist in the literature base severe limitations in scientific rigor relating to impact evaluations of health and safety interventions. A possible reason for this was thought to be that the subjective benefits of health and safety intervention programmes have been so impressive and the likelihood of impressive effects so great that there have been few demands for more reliable data[1].

Other factors that may have also contributed to the lack of a robust literature base included:

- A general lack of scientific training for safety practitioners[2, 3];
- A general lack of funding for evaluation research[4, 5];
- Ethical issues relating to research designs that might withhold risk control measures from control groups[6];
- The pragmatic nature of health and safety[7];
- Rule-driven basis of prevention strategies[8, 9].

The future ability to conduct meta-analyses will continue to be affected by these limitations. Hillage *et al*[10] suggested a comprehensive strategy for improving the evidence base, including ensuring evaluation was included in the planning of interventions; developing a stronger evidence base by conducting more systematic reviews and meta-analyses; and adopting minimum standards. Worthen[11]

has predicted that the use of evaluation will increase in years to come and the approaches used will evolve to become more suited to applications within occupational health and safety.

Taking into account the findings of the literature review, it has been concluded that a suitable research design should posses the following features:

- A base of currently accepted UK national guidance on safety management;
- An economic evaluation component;
- A linkage between the causes of incidents and accidents to the health and safety management system;
- An intervention in the form of written guidance or work book format, supplemented by an educational component;
- A quasi-experimental or experimental design with baseline measurements;
- A comparison group;
- Random sampling where possible;
- Validated methods;
- A control strategy for potential inter and intra-observer bias;
- More than one performance evaluation method including both objective and subjective methods;
- An attempt to correlate the results of the performance evaluation methods (triangulation);
- Sample sizes of sufficient size to allow statistical inferences to be made, where possible;
- Validity checks to ensure that pooling of data from different sites was appropriate before proceeding with analysis.

INTRODUCTION

The research described in this report was funded by a joint partnership between HSE and SEDoH. The HSE had invested considerable resources in developing a methodology, which enabled the identification of health and safety management root causes and costs of incidents. They were interested in primary research that used the methodology, so as to inform decisions as to potential future applications. The SEDoH were interested in the improvement of health and safety management within the healthcare sector. The workbook designed by the author was intended to do this by assisting line managers with their health and safety management responsibilities. SEDoH were also interested in whether use of this tool could help to improve health and safety management performance in the healthcare sector.

In this context there were three research objectives for a critical in-depth review of the literature identified. These were:

Literature Review Objective 1:
To identify suitable measures for evaluating health and safety performance;

Literature Review Objective 2:
To evaluate the design of interventions to improve health and safety management performance;

Literature Review Objective 3:
To identify a suitable design for a research study to measure change in health and safety performance.

This report will address these objectives by drawing from the body of published work. The review has three parts:

Section 1: describes the methods used to conduct the literature review;

Section 2: contains the findings of the literature review, grouped against each of the three literature review objectives;

Section 3: contains conclusions about the implications of the findings for research designs most suited to achieving the overall research objectives identified in Chapter 1.

1. LITERATURE REVIEW METHODOLOGY

As observed by others[5, 12-16] an issue for researchers in occupational health and safety management is the breadth and depth of the literature, which is spread across several disciplines. Therefore, the search was carried out across professional and scientific disciplines, and included several different data sources and literature collections (listed below).

The review was conducted during parts of 2002 and 2003. Relevant studies were retrieved, using a systematic approach. Searches were restricted to articles in English, and conducted using combinations of the following search words:

> Accident, Assessment, Change, Evaluation, Impact, Incident, Intervention, Management, Measurement, Meta-analysis, Prevention, Review, Root-cause, and Safety.

The following sources of information were used:

- *Biomedical:* – BioMedNet; Cambridge Scientific Abstracts (CSA); Cochrane Library; Cumulative Index to Nursing and Allied Health (CINAHL) Database; Internet Database Service; MEDLINE; NeLH; OVID.
- *General* – Emerald Fulltext; Emerald Management Reviews; Ingenta Services; NESLI (National Electronic Site Licence Initiative); ZETOC –BL.
- *Government:* – Canada – EU Cooperation on Workplace Safety & Health (www.eu-ccohs.org/); Department of Health (http://www.doh.gov.uk); European Agency for Safety & health at Work (http://europe.osha.eu.int/); Health & Safety Executive (http://www.hse.gov.uk/); International Labour Organisation (ILO) (http://www.ilo.org/); National Institute for Occupational Safety & Health (NIOSH) (http://www.cdc.gov/niosh/homepage.html); *Scottish Executive Publications online*.
- *Health & Safety:* -Directory of Occupational Health and Environmental Hygiene Links (http://www.agius.com/hew/links); Embase (covering occupational health, environmental health and ergonomics databases); International Occupational Safety and Health Information Centre (CIS); The Institution of Occupational Safety & Health (http://www.iosh.co.uk).
- *Life Sciences:* – Edina CAB Abstracts; ISI Web of Science Citation databases.
- *Science & Engineering:* – EDINA Compendex (indexes to engineering articles and conference proceedings); *European Network for Process Safety (SAFETYNET)(http://www.safetynet.de/);* Science Direct.
- *Social science:* – Applied Social Sciences Index and Abstracts (ASSIA); BIDS International Bibliography of the Social Sciences (IBSS); PsychINFO.

Additional information was obtained by checking author names and key publications via citation indices.

2. FINDINGS

The identified publications are discussed in subsequent sections. However, a noticeable general feature was that, despite regular overlapping of the publication years of the articles examined, few of the published reviews evaluated the same studies. For example Livingston *et al*[17] did not cite the same work on accident causation theories published earlier by a study group commissioned by the Advisory Committee on the Safety of Nuclear Installations (ACSNI)[14] despite having clear objectives to review occupational health and safety-related intervention studies. Reasons for this are unclear, but one possible cause could have been the multidisciplinary nature of the subject area and the breadth and depth of the literature. This was supported by the conclusions reached by other researchers (identified earlier in Section 2.1.).

2.1. Literature review objective 1 (To identify suitable measures for evaluating health and safety performance)

Four main approaches to health and safety performance evaluation were identified. They were:

- Methods that could be used in an objective manner, such as those based on audit or incident data;
- Subjective methods such as safety culture/climate evaluations;
- Economic evaluations;
- Methods that used more than one type of method.

Each will be considered in turn below.

2.1.1. Performance evaluation methods that could be used in an objective manner.

It was intended to use the method developed by HSE, which was based on the analysis of incident data. There was a large literature base identified on the use of incident data to evaluate health and safety performance. There was also a large literature on methods that were based on the use of audit. Advantages and limitations of both approaches are discussed below.

Methods based on the analysis of incident data.
To develop the method to be used in this research, HSE used the findings of a review by Livingston *et al*[17] as the basis of their methodology, which used incident investigation to link root cause analysis with failures in the health and safety management system[18].

Root cause analysis was at the core of the clinical risk management agenda of the UK NHS[19-22]. The use of methods based on root cause analysis would therefore be complementary to other NHS initiatives and a potentially useful research tool for evaluating health and safety performance in the NHS.

Based on events and causal factors charting[17], the HSE method was designed to be applicable to a wide range of industries. The method represented a potentially powerful novel tool with which to evaluate health and safety performance, so far as it related to the UK system for health and safety management. Other evaluation methods were identified[17, 23-26] but none included this aspect.

So far as limitations to using incident investigation as a performance indicator was concerned, a major potential source of bias was the degree of intra- or inter- investigator consistency.

In 2002 the National Patient Safety Agency (NPSA)[a] published an evaluation of a six-month feasibility study into the key requirements for managing, reporting, analysing and learning from adverse patient

[a] The NPSA was established as a Special Health Authority in July 2001, with a remit for the co-ordination, reporting and assessment of adverse events occurring in the NHS in England and Wales.

incidents[19, 20, 27, 28]. Cowan[29] highlighted the practical challenges of achieving consistency in the investigation of adverse clinical incidents and commented that this factor remained an unresolved issue between and within NHS organisations. Other researchers have supported this conclusion. Examples are given below (listed alphabetically by author):

- The ACSNI Study Group[14] concluded that investigation outcome was linked to the beliefs of each investigator. A simplistic concept of a single primary accident cause as opposed to belief in a multiple causation model were thought likely to affect the approach taken by the investigator, and affect the conclusions reached;
- Chilton *et al*[13] identified an extensive literature base from the 1980s and 1990s, supporting the theory that both experts and novices exhibited systematic biases in their health and safety assessments, with strong evidence that experts may have become over-confident. The outcome of these assessments was also thought to be dependent on the background and experience of the assessor;
- Hartley and Booth[30] evaluated a range of occupational health and safety specialists from different professional disciplines (e.g. safety, occupational hygiene, ergonomics, occupational health nursing) as they undertook a simulated risk assessment task. They found that the professions used different approaches and formed widely different conclusions. There were also significant variations within each professional group;
- Henderson *et al* [31] found wide variation in approaches to incident investigation and concluded that the current UK national standard of incident investigation was poor. They examined whether these differences were likely to affect the outcome of investigations and concluded that the main limitation was the likelihood that underlying causes would be identified inconsistently;
- Rakel *et al*[32] highlighted problems in achieving consistency of findings between UK HSE field inspectors following health and safety inspections. The research team attempted to guide and standardise the approach taken by the inspectors but this led to frustrations and tensions between the team and the inspectors. The issue remained unresolved and prevented detailed analysis of the data;
- Scherer *et al*[33] found statistically significant differences ($p<0.01$) between both the approach and outcomes of two different approaches to safety inspections conducted by enforcing officers from State and Federal agencies in the USA.

In addition, some authors have cautioned against using rates within safety incentive programmes because of the difficulty of benchmarking created by different reporting cultures within organisations, especially where frequency targets are imposed or safety incentives exist[34, 35].

It was clear that evaluation methods that used techniques of incident investigation or workplace inspection should attempt to control for sources of inter- and intra-investigator bias.

Another significant limitation of the use of incident data as a performance measurement method was the potential for differences in incident reporting conventions between and within organisations. Kjellen suggested four obstacles (i.e. economic, methodological, psychological, and organisational) that could affect whether an incident is reported or not[36]. This could be further confounded by the possibility that, as staff became more aware of health and safety, they might tend to report more incidents[37]. This effect could have a profound confounding effect on performance evaluation. The main method to control for this risk would be to use more than one evaluation method, preferably linked via triangulation (see section 2.2.3.)

Because of the risk of inconsistent identification of underlying causes[31] ensuring a consistent approach was recognised as important. Since it was widely accepted that the causes of incidents were related to health and safety management[38-40], it followed that reliable identification of causal factors linked with aspects of health and safety management systems could potentially offer the possibility for identifying management remedial actions which might reduce the occurrence of incidents.

None of the identified methods attempted to link analysis of incidents to the health and safety management system of an organisation, a finding also noted by Hale *et al*[41]. Within the NHS this was thought to be a negative feature since there were strong drivers to ensure robust health and safety management systems were developed[38, 42-45]. An evaluation method that was linked to the health and safety management system would also be more likely to be understood by staff working in the NHS. Therefore the HSE method was the only method, which could be used for this purpose.

Methods based on the use of audit

The use of audit has been well established since the 1970s as a means to evaluate health and safety management performance[37, 46]. Audit has been described as a systematic process whereby a comparison is made between the audit standard and the actual situation in a workplace[46]. They are predictive in nature (a classic "leading" indicator, described further in Section 2.1.4[47]) and readily acceptable by management because they are familiar with the concept being used in other areas such as finance and clinical practice.

A number of commercial audit systems existed. These included:

- The International Safety Rating System (ISRS)[48, 49]. Developed by Det Norske Veritas (DNV), was designed to be used by trained managers[50] and included nearly 600 audit questions grouped into 20 elements relating to health and safety management;
- The TRIPOD system, was an approach to safety developed by the Universities of Leiden and Manchester in collaboration with the Shell Petrochemical Company[51, 52]. The system used checklists to identify "latent" failures[53], seen as the potential cause of accidents, for the offshore industry;
- The Complete Health and Safety Evaluation (CHASE) system[54] was developed by HASTAM Ltd, an independent company of the Aston University health and safety department. Based on accepted health and safety management models[40, 55], the system was designed to be tailored to the needs of the organisation.

Other researchers have developed non-commercial audit tools. Examples include:

- Scotney[56] who developed an audit tool for HSE inspectors to measure health and safety performance in Small and Medium Enterprises (SMEs);
- Fuller[57] who used HSE[58] and BSI[59] key health and safety management performance indicators to assess safety culture via benchmarking audits in a series of intra- and inter- company comparisons.

Although audit systems have been used effectively to evaluate health and safety management performance they have a number of limitations. They are prone to investigator bias and may not be relevant to the organisation being audited[37]. This was noted by Hale *et al*, who concluded that wide variations in audits have led to a lack of consistency between audit systems few of which are linked to underlying management models[41].

A major strength has been in their ability to be used to develop motivation for change (i.e. as an intervention *per se*)[46], although some authors have questioned their use in performance evaluation because of a lack of evidence that they brought about an improvement in safety by reducing injury rates[37, 50, 60]. However, any change that they might bring about may not be desirable in a research

project, which seeks to use the technique to measure change of an intervention. This effect would however, be minimized with the use of comparison groups in the research design (see section 3).

Because they are so comprehensive, and require a high level of commitment at all levels in an organisation, they can be time consuming to implement and were most commonly used in organisations that had mature health and safety management systems, such as the petrochemical industry[46].

Nevertheless, using audit to evaluate health and safety management performance was identified as a potentially powerful method. This was also acknowledged by HSE who ensured that the pilot work for their method was validated against an HSE inspection audit[18]. This study found good agreement with the findings of the HSE root cause analysis method and the audit. This meant that resources could be devoted to the root cause investigation of incidents. Had this finding not been present it would have been necessary to have considered the inclusion of an audit component as part of the evaluation.

2.1.2. Subjective performance evaluation methods.

Guldenmund[61] and Hale and Hovden[62] have reviewed the theory and research base on the nature of safety culture and concluded that the scientific approach was immature given that there was no single accepted model or definition of safety climate or safety culture.

Cooper[63] suggested that safety culture could be described as an interactive relationship between psychological, situational and behavioural factors. Safety climate, on the other hand, was regarded as a component of this, which could be measured. This definition and distinction between the two terms was adopted for the research project.

The assessment of safety climate has been seen as having considerable potential as an alternative safety performance indicator. Cooper[63] suggested that safety climate, defined by internal psychological factors, such as attitudes and perceptions of safety, could be assessed using interviews, checklists or questionnaires.

Given there was no prerequisite by the sponsoring bodies for the research project for a subjective performance evaluation method, the literature search therefore sought methods that were:

- Fully developed and ready for use. This was because the research objectives were not intended to include the development and validation of a qualitative method;
- Applicable to a range of occupational groups across a diverse industry such as the NHS;
- Applicable to general health and safety management rather than a specific hazard or group of hazards.

The findings are grouped under these three headings below.

Methods not fully developed and ready for use.
Flin *et al*[64] reviewed the common features of eighteen different methods used to assess safety climate across five different sectors (energy/chemical, manufacturing, transport, construction and generic). The authors concluded that none were ready for use (by 2000). The one that had most potential for the research project was the generic method. This was HSE's Health and Safety Climate Survey Tool[65, 66] but at the time of the evaluation was not yet supported by published field data.

Davies *et al*[67] reviewed six published general tools designed to measure safety climate. Although the independence of the review could be questioned (the review had been funded by HSE, and was not peer-reviewed), all methods were assessed as likely to benefit from case study examples to establish their usefulness in assisting with planning improvement actions.

The HSE Health and Safety Climate Survey Tool method was assessed as suitable but its main purpose was as a method to promote employee involvement in health and safety rather than as an evaluation method *per se*. It also required the participation of all staff in an organisation, something that was impractical for the research contemplated.

Other examples of methods, which were not included in either of the reviews of Flin *et al*[64] or Davies *et al*[67], are listed below:

- Bailey and Petersen[68] developed a questionnaire instrument for the railway industry in the USA. Although extensively researched over a nine year period the method had been developed using early software and programming language, so was not ready for use;
- Donald and Canter[69] developed a set of scales to measure attitudes to safety in the chemical industry. This method was also acknowledged as requiring further development;
- Hayes *et al*[70]described the development and validation of a method to measure a scale of perceptions of workplace safety – the "Work Safety Scale", using a five-dimension model of work safety perception. However, the validation was not complete and the authors identified that further development work was still required;
- Some authors [7, 71, 72] have advocated the use of models more traditionally associated with business or market research. Examples identified included in-depth interviews, or approaches such as cognitive mapping using "focus groups". These techniques potentially offer an alternative for conducting intervention research in the field of health and safety risk management. However, examples of applied research in health and safety management were needed.

Methods unsuitable for use across a range of staff groups.
Examples are given below of methods that were rejected because they were not thought to be applicable to the wide range of different staff groups within the NHS:

- Niskanen[73] developed scales for use in measuring safety climate in workers and supervisors in the Finnish National Road Administration;
- Rundmo and Hale[74] investigated attitudes of senior executives and managers to health and safety management at the start of a training workshop on management safety behaviour;
- Smallman and John[75] sought indicators of corporate performance and health and safety, which could be used to benchmark British directors' attitudes to health and safety;
- Vassie and Lucas[76] used semi-structured one-to-one interviews to gauge the opinion on safety performance of the health and safety professional, employed within a random sample of each of thirty-five UK manufacturing companies;
- Zohar[77], Brown and Holmes[78], and Dedobbeleer and Beland[79] developed and refined safety climate scales for production and construction workers.

Methods not suited to general health and safety management
Examples are given below of methods that were rejected because they were not thought to be sufficiently generic:

- Basen-Enquist *et al*[80] developed, using factor analysis, health and safety climate scales for use in a self-reporting questionnaire. The scales were developed to measure organisational change as a result of a health promotion programme;
- Gershon *et al*[81]devised a safety climate questionnaire to measure hospitals' commitment to blood borne pathogens risk management programmes;
- Pilkington *et al*[82, 83] used interviews to identify the steps employers were taking to manage the risk of work-related stress.

Therefore a method could not be identified that required no further validation prior to use. The method that offered most potential was that of Cox and Cheyne[84] who published a nine-dimension

methodology for assessing general health and safety management culture in all staff groups in the offshore industry (the Loughborough Safety Climate Assessment Toolkit). The method was based on a systems approach to organisational culture, which had been refined using focus groups, factor analysis, and field trials. The questionnaire, along with a "toolkit" for use in data analysis were freely available via the Internet[85]. This was assessed as potentially adaptable for use as an evaluation method in the field of health and safety performance evaluation in the NHS.

Although developed for the offshore industry its question set was sufficiently generic to offer potential for use in the NHS, subject to confirmatory factor analysis, such as that conducted by Brown and Holmes[78] and Dedobbeleer and Beland[79].

2.1.3. Economic Evaluations.

The author has reviewed economic evaluations in the field of health and safety in healthcare[1]. As well as suffering from reporting bias the other main limitation within economic evaluations was in benefit evaluation, which has been acknowledged as difficult to estimate[13, 86, 87]. Warner *et al*[88]described this as a "wish bias" whereby researchers underestimate costs and overestimate benefits in their economic evaluations thereby "finding what they seek". Typical examples included:

- Cooper *et al*[89] experienced difficulties with quantification and analysis of financial consequences of workplace stress within the confines of company financial routines. There was also a lack of basic data with which to construct a financial baseline to measure the outcome of stress interventions;
- Kemmlert[90] reported impressive savings for ergonomic preventive strategies. This was based on an assumption that benefits accrued by one individual (expressed as a percentage reduction in sickness absence), as a result of the introduction of a control measure, also applied to other staff working in the same area. This presumed that all staff would all develop the same health problem and be off work for the same lengthy period of time as the index case, a highly unlikely occurrence.

There was also a risk that organisations such as the NHS, with a short-term outlook on financial management, attempting to balance the books within a fiscal year may not be convinced by the realistic potential for achieving impressive savings.

For instance, Monnery[91] reported a cost analysis of accidents and work-related ill health to a cheque clearing department in a financial services organisation. A method developed by HSE[92] was used to gather data on accident and ill health costs. The resultant costs were not regarded as substantial (0.5% of the annual salary costs), while those relating to ill health (mainly due to upper limb disorders) were negligible.

Nevertheless, including a costing aspect in the research design, with other performance evaluation indicators was regarded as potentially useful. Justification for this would be to investigate whether cost could be linked to other performance evaluation methods. This was a potential use of costing data that has not been widely explored, except by HSE and the author[18].

2.1.4. The use of more than one method.

Many authors advocated that more than one method should be used to measure health and safety performance (e.g. [26, 32, 47, 61, 63, 64, 72, 84, 93-97]).

Easterby-Smith *et al*, in discussing the philosophy of investigating human and social behaviour, suggested the relevance of the two traditions of positivism and phenomenology[72]. The former (objectivist) paradigm was characterised by research designs that seek external causes and fundamental laws to explain behaviour. The latter (subjectivist) approach attempted to understand and explain why people

have different experiences. It was suggested that, to effectively analyse health and safety performance, studies should, ideally, address both concepts.

Within health and safety research the general approach has been to use both quantitative (objective) and qualitative (subjective) data collection methods[95, 96], various data sources; and different observers[72] (e.g. Cox and Cheyne[84], Jacobs and Haber[95], and Shannon *et al* [98]).

Various authors[47, 64, 93, 94] have also recommended an approach that used both lagging indicators (e.g. incident data) and leading indicators (e.g. audits or measurements of safety climate).

Examples of published research that used multiple measurement evaluations are given below:

- Cooper[63] recommended multi-level analysis and their three-level reciprocal safety culture model was based on this concept. The model generated data on the individual from a safety climate questionnaire. The degree of safe behaviour was assessed using checklists, and audits of the safety management system provided data on the safe situation;
- Cox *et al* [99]described a phenomenological impact assessment of a variety of interventions to control stress in hospital staff in five different case studies. Evaluation was carried out by three different comparisons (i.e. awareness, tangible impact and positive impact) between test and comparison groups. The data were obtained by a combination of questionnaires and interviews. Results were descriptive;
- Jacobs and Haber[95] and Haber *et al*[96] used various methods including structured interviews, behavioural checklists, assessments of worker performance using rating scales or observation, and questionnaire surveys to assess safety within the nuclear industry. Both qualitative and quantitative data were generated;
- Mearns *et al*[47] investigated whether a favourable safety climate was associated with lower incident rates. They measured safety climate by postal questionnaire, in thirteen offshore oil and gas installations on two occasions, with a year in between each measurement. Another questionnaire was used to obtain data on management safety practice and both of these were compared with incident reporting data. Associations were assessed between safety climate and performance and safety management practice and performance. Results showed suggested associations between some management practices and incident rates. However the sample sizes were small and there was no comparison group, thus making the results indicative than definitive;
- Using a non-random matched case-control design, Rakel *et al*[32] collected information on change in expressed opinion on health and safety which occurred after a seminar or mail shot conducted by HSE. The authors used both qualitative and quantitative data collection methods, various data sources, and different observers[32]. They found that seminars were effective in stimulating change in opinion but mail shots less so. However, the data from each source were too dissimilar to combine and statistical inferences could not be made.

Guldenmund[61] identified that few studies have attempted to establish correlation ("triangulation") between safety performance measures (such as accident/incident data or safety audits) and safety culture or climate assessments, and called for these relationships to be explored in more detail. Two examples of primary research that attempted this were identified:

- Hurst *et al*[93] combined a safety attitude questionnaire, a management systems audit, and accident rates at six major hazard sites. Attempts were made to examine the variation in accident performance data and safety attitudes on different sites. Incident rates varied by a factor of 100 between the six sites. The audit tool was useful for individual sites but time consuming and difficult to standardise across all six sites. The questionnaire produced mixed results and could not be correlated against self-reported accident rates;
- Cabrera *et al*[94] combined the use of a questionnaire to assess safety climate, a safety compliance audit, and one-to-one interviews to assess the level of safe behaviour in the aviation industry. The

purpose of the research was to evaluate whether there was a relationship between safety climate, audit score and behaviour. The results were largely descriptive of the many variables identified and failed to establish a relationship between the three measures of performance.

The sample sizes in both papers were too small to allow statistical inferences to be made. However these publications represented the only research to be identified, which attempted to triangulate three different performance measurement methods. Therefore, there existed many advocates for the benefits of a triangulated approach, but no definitive proof of its effectiveness.

To address this some authors had identified a need for quantitative weightings of the importance of different influences and evaluation methods[95, 100].

Nevertheless multiple measurement methods represent a strong feature of an evaluation method. HSE linked their root cause analysis approach with previously published guidance on health and safety management[58] and the costs of incidents[92]. Their new methodology therefore linked incident investigation, root cause analysis and costing[18]. This method, however, did not include any aspect of measurement of safety climate. If this were included the potential for triangulation of data would be greatly strengthened.

2.2. Literature review objective 2 (To evaluate the design of an intervention to improve health and safety management performance)

To fulfill the project research aim required a formal health and safety intervention to be introduced into the NHS. For this to be realistically achievable across the whole sector it was regarded as needing to possess the following features:

- Acceptable to NHS Trusts, otherwise willingness to participate in the research would be unlikely to be secured;
- Able to be implemented within current resources (i.e. cost-effective). This was because cost and spending targets and priorities tended to be set at National level and Trusts had comparatively restricted freedom to act outwith these targets;
- Generic, so as to be applicable across the wide variety of types of services and staff groups within the NHS;
- In keeping with other health and safety drivers within the NHS (such as from HSC[101, 102], DoH[19, 103] and SE[45]).

The intervention workbook[b] used for the research was developed to ensure that it met the above criteria. In practice this meant taking a general systems approach, which had a long history in the NHS.

For example, in 1994 the HSC[38] issued information for directors and managers on management of health and safety in the health services. One of the objectives in an ambitious ten-year plan was to fully integrate health and safety management with all aspects of management within the NHS. This was intended to harmonise with other legal requirements and HSE guidance[39, 104].

Another advantage of taking a generic systems approach was that it was in keeping with established conventions within the health and safety profession.

[b] The workbook can be viewed by contacting the author.

Hale and Hovden[62] have reviewed the literature on systems-based intervention approaches in health and safety. They concluded that most models were derived from principles of quality management (the "plan-do-check-modify" loop), which had proven effectiveness. Extensive pioneering research in the 1970s established the importance of systematic and pro-active health and safety management in leading to lower injury rates (e.g. [105-109]). Since then the position has been consolidated with increasing acceptance of the model both in the UK and internationally (e.g.[110-112])

This paradigm has also been supported and advocated by HSE[39, 40] and, importantly, has found widespread acceptance within the NHS. Most Trusts were thought to be working, to a greater or lesser extent, towards a systems-based model of this sort. Therefore although other approaches were available (such as those developed by the British Standards Institution (e.g.[55, 59, 113, 114]) or the nuclear industry (e.g.[115, 116])), the NHS at the time of the research was thought to be more readily accepting of an intervention tool that was based on the systematic approach to management of health and safety advocated by HSE[40].

Nevertheless, it was acknowledged that there were limitations to a systems-based approach of this type. These risks were:

- The intervention might be overly generic to make an impact on health and safety performance at local departmental level;
- The intervention might not be sufficiently simple to be used intuitively by line managers as a relatively "stand alone" resource;
- The intervention might not produce an effect strong enough to be detected;
- The effectiveness of the method might be reduced because it did not make use of health and safety management programmes based on psychological or behavioural models.

Each is considered in turn below.

2.2.1. The risk that the intervention might be overly generic to make an impact on health and safety performance at local departmental level.

Some groups of researchers used an approach tailored to the specific needs of an "at risk" group. For example:

- Swuste *et al*[117] used prioritised safety work plans to improve performance. The approach was iterative, designed to identify interventions suited to specific hazards;
- Cumming *et al*[118]found a statistically significant reduction in falls by elderly people in the home following an occupational therapist-led intervention (P=0.50), which was targeted at the specific needs of the individual;
- Pelletier *et al*[119] found a positive effect ($P<0.05$) for an intervention aimed at reducing stress within financial sector employees. This study used written material supported by telephone access to advice that was tailored to specific needs of individuals.

Although an iterative approach may have been desirable for the research it was regarded as insufficiently generic for application within the NHS as a whole. The risk was addressed by ensuring that examples, relevant to the NHS were included in the workbook and local conventions and terminology were used.

2.2.2. The risk that the intervention might not be sufficiently simple to be used intuitively by line managers as a relatively "stand alone" resource.

Daltroy *et al*[120] found, in a large-scale RCT of an educational programme to prevent work-related low back injury, that no significant long-term benefits were associated with training alone in a group of postal workers. Therefore, simply training managers in health and safety management techniques alone was unlikely to have an effect.

This was supported by De Roo[12] who undertook a systematic review of farm safety interventions. The approaches identified were education and training programmes, booklets, workbooks and guidance documentation. Methods tended to be used in combination with each other. For example, most written material was supplemented with an educative component such as a briefing session.

It seemed logical that if written material was used that, to ensure its effect was optimised, some form of educative component supports it. It was therefore regarded as important to use a combined approach, using a written guidance/workbook approach supplemented by education and training as a cost-effective delivery mechanism.

2.2.3. The risk that the intervention might not produce an effect strong enough to be detected.

Spangenberg *et al*[121] used a multi-faceted safety campaign, with five main components, to improve safety performance on a construction site. The effect of the campaign, a reduction in the number of reported injuries resulting from accidents, was only marginally significant. This illustrated the challenge of achieving an effect, despite considerable and sustained effort. Goldenhar and Schulte[122] suggested that, for an intervention to have an impact on safety performance, it had to be interesting and workers needed to use it regularly.

To address these issues the workbook was aimed at line managers who had acknowledged legal responsibilities for health and safety management. Additionally it was written in a non-technical language and Trusts were encouraged to ensure that managers appreciate the need to use it regularly.

2.2.4. The risk that the effectiveness of the method might be reduced because it did not make use of health and safety management programmes based on psychological or behavioural models.

HSE have advocated[123] that consideration of "human factors" is a key ingredient of effective health and safety management, and had recommended that behavioural strategies be incorporated into a general health and safety management system.

The use of behavioural approaches has been well documented over four decades (e.g.[123-134]). However there has been little consensus on standardised approaches. Most of the primary research has been typified by a lack of methodological rigor[131] and programmes that tended to be complex and labour-intensive to implement[133, 134].

Nevertheless, a psychological approach may have been useful to include in an intervention. Its relatively under-developed nature need have not been a barrier to its adoption since it had potential to affect health and safety behaviour.

The main disadvantage was that it was, at the time, insufficiently generic for use as the sole approach in a general intervention.

Therefore, on balance, a systems-based approach[62] based on currently accepted guidance on health and safety management[40] was most likely to harmonise with other NHS drivers and therefore represent a potentially effective tool for the NHS.

2.3. Literature review objective 3 (To identify a suitable design for a research study to measure change in health and safety performance)

The purpose of a research design has been defined as ensuring that there are measures in place to organise research activity, including the collection of data, in ways that are most likely to achieve the research aims[72]. Therefore the overarching strategy for the literature review was to search for research designs that had the potential to do this.

Meta-analysis was known to be a powerful method for assessing the validity of research by statistically combining results of comparable studies[135-137]. Given their potential for identifying strong research designs, meta-analyses and systematic reviews in occupational health and safety evaluation and intervention studies were specifically sought.

Only four meta-analyses of performance change evaluation [5, 50, 138, 139] and four systematic reviews of intervention studies [12, 16, 97, 140] were identified. This was thought to be a low incidence of these types of publication, particularly considering the wide-ranging multidisciplinary professional and scientific basis of the search strategy.

A possible explanation was that the primary research base on which the reviews were based was insufficiently robust to support this type of analysis. For example, Shannon *et al*[97] were unable to make quantitative comparisons between ten studies. This was supported by Rivara and Thompson[141] who attempted to conduct a meta-analysis to assess the effectiveness of different strategies to prevent falls from heights in the construction industry. Their review methodology was based on Cochrane Collaboration guidance[135, 141] and identified only three studies suitable for review. However, the methodological quality was of such poor quality (no controls or appropriate multivariate analyses to control for potential confounding factors) that no attempt was made to undertake the meta-analysis and few conclusions about the effectiveness of the intervention strategies were made.

Oliver *et al* [139] conducted a meta-analysis of twenty-one papers, which met strict pre-determined inclusion criteria for hospital fall prevention programmes. The authors concluded that there was a tendency for the most rigorously controlled trials to produce the smallest effects. For example, Lingard and Rowlinson conducted a behaviour-based approach to safety management in the construction industry in Hong Kong[142]. In common with Harper *et al*[143], the authors used a "multiple baseline across groups" design but also included a reversal aspect, cited as being particularly suited to measurement of safety behaviour[144]. The results were mixed in terms of detected improvements in performance.

Conversely, it was noticeable that the studies most likely to quote impressive effects on health and safety performance following interventions were usually of the non-experimental type (e.g. [14, 145]). Non-experimental before-and-after designs were not regarded as suitable for evaluating change in health and safety performance because of methodological limitations of the research design resulting in uncontrolled sources of bias.

It was identified (in 2.1.4) that performance evaluation should be based on multiple outcome methods. Although it has been acknowledged that this does not solve the problem of bias[146], the use of multiple methods can aid interpretation of results.

Other potential source of bias are identified below.

2.3.1. Lack of control groups and randomisation (e.g. [14, 74, 75, 93, 94, 138, 147-152]).

Two examples are given below:

- The ACSNI Study Group reviewed seventeen examples of before-and-after intervention studies, published between 1973 and 1989[14]. All of the examples cited showed impressive positive effects on health and safety performance. However, the reviewers concluded that widespread lack of comparison groups and lack of random selection of participants reduced the credibility of the results;
- Krause *et al* [138]carried out a longitudinal meta-analysis of accident rates in seventy-three companies over five years. They used a quasi-experimental time series design. A significant improvement in recorded injury rate was reported but there was no control group with which to compare to the test group. This was a particularly important omission because the study took place over a relatively long time period (5 years) and it was foreseeable that other factors could have had an influence on injury rate in that time.

2.3.2. Lack of control of other aspects of internal validity

Placebo effects can occur as a consequence of psychological factors that resulted in performance improvement[144]. For example, McGrail[149] reported a significant decrease in injuries and illnesses following the introduction of an ergonomic intervention in a hospital. However, the intervention was only applied to individuals who had sustained an injury or illness and were off work. It was not possible from the study design to ascertain whether there were any psychological effects of the intervention.

The involvement of the researchers or other outsiders can produce an effect on the measured outcome by virtue of their presence or involvement ("Hawthorne" effects[153]). In the study reported by McGrail[149], subjects were exposed to a wide range of professionals (such as specialists in orthopaedics, neurosurgery, rehabilitation medicine, physiotherapy and occupational therapy). The effect of these individuals on the outcome was not discussed or controlled. The Curtin Industrial Safety Trial[143, 154] was a before-and-after evaluation of the effectiveness of a behaviour-based safety programme. Baseline observations of behaviour were taken prior to the introduction of the intervention followed by measurements taken over consecutive weeks. Control groups were not used - the design would have been less prone to researcher influence if some groups received the baseline and weekly measurements but not the intervention[143].

2.3.3. Lack of exploration of the statistical significance of the findings (e.g. [131, 133, 134, 147]).

Where statistical analysis was conducted the tendency was toward multivariate analysis (chi-square), which was appropriate for nominal measurement scale data generated[26, 149]. Analysis of Variance (ANOVA) was also used for ordinal incident data[149]. Independent *t*-tests were used to compare occupational groups[143, 149]. For example, Harper *et al*'s[143] data for observed safe practice and housekeeping rates were treated as interval and analysed, after empirical logistic transformation, using *t*-tests. This approach was justified by fitting trend lines to the proportions to confirm that many were close to one.

These choices of analysis methods was generally suitable although they would have been enhanced had initial descriptive statistics (such as measures of distribution and spread) been explored to assess the validity of treating data as interval as opposed to ordinal, before estimations of statistical significance were attempted. Only one publication took this approach[155].

There was also a lack of information on sample sizes and associated degrees of freedom (e.g. [97, 156]).

2.3.4. Existence of selection or resentment threats

An example was Yassi *et al*[145] where the stated aim was to investigate the effect of an early rehabilitation-based intervention for nurses with back injuries in a hospital. The nurses assigned to the test group were selected from the highest risk wards in the hospital, whereas those nurses assigned to the control group, which did not receive the intervention, were all from wards which had not previously

reported high incidence of back injuries. The impressive results (decrease of between 23% and 43% in rate of incidence of injuries) could not be justified, given the bias between the two groups. Resentment threats occur when the individuals in the control group react to not receiving the intervention. In the study described above[145] this possibility could not be ruled out because the nurses hospital was on a single site and many of the nurses knew one another and the details of the study.

2.3.5. Small sample sizes (e.g. [93, 94, 122, 148, 157]).

To overcome the limitation of small sample sizes, analysis of data was frequently conducted by first combining data from multi-site samples (e.g. [158-161]). However, Mearns *et al*[47] identified a variable response rate in their results, and advocated caution when considering combining data from different sites. Cooper[63] also concluded that the disadvantages of this approach, by introducing errors into the data, outweighed the advantage of the larger sample size.

In the light of the above potential limitations, a series of research design criteria was developed, based on maximising the potential for provision of reliable data. These criteria were derived from various sources[32, 72, 98, 135, 137, 144, 162-164], but largely based on guidance from the Cochrane Collaboration[135]. Although mainly aimed at reviews of randomised controlled trials (RCTs) in healthcare, the general principles of the Cochrane reviews were well established and found by the author to be readily adaptable for health and safety research.

The research design criteria that should be included were therefore as follows:

- Inclusion of an adequate description of the work with stated intervention objectives;
- Inclusion of justification of the conceptual basis of the work (qualitative or quantitative) and the study design used (experimental, quasi-experimental or non-experimental[144]);
- Inclusion of justification of how external validity had been addressed (i.e. with a control group or randomisation);
- Inclusion of justification of how internal validity (i.e. control threats) had been addressed;
- Inclusion of appropriate outcome analysis (including statistical analysis);
- Inclusion of conclusions that addressed the study objectives and were supported by the analysis;
- Inclusion of a discussion of the limitations of the research;
- Inclusion of a discussion of any practical significance of the results.

In the context of the research planned, a major source of potential bias was the ability to secure a random sample of NHS Trusts. In the experience of the author, for Trusts to be willing to participate in research, they were likely to be interested in the work in order to consent to take part and commit to providing local support. In the Curtin Safety Trial[143, 154] the sample selection process was non-random because the companies involved could not be selected at random[143]. Although this approach introduced a source of bias the concept of a pragmatic “convenience sample” was nevertheless potentially useful. Initial screening criteria were used to select participants. These included management having expressed commitment to improving safety and to collaborating with the trial and the worksite and workers being readily observable. This approach was likely to be useful in the NHS where, given the small number of Trusts available, random allocation was not a practical option.

3. DISCUSSION AND CONCLUSIONS OF REVIEW

The main finding of the review has been that there exist in the literature base severe limitations in scientific rigor relating to impact evaluations of health and safety interventions.

A possible reason for the above limitations was thought to be that the subjective benefits of health and safety intervention programmes have been so impressive and the likelihood of impressive effects so great that there have been few demands for more reliable data[1].

Other factors that may have also contributed to the lack of a robust literature base:

- A general lack of scientific training for safety practitioners[2, 3];
- A general lack of funding for evaluation research[4, 5];
- Ethical issues relating to research designs that might withhold risk control measures from control groups[6];
- The pragmatic nature of health and safety[7];
- Rule-driven basis of prevention strategies[8, 9].

The future ability to conduct meta-analyses will continue to be affected by these limitations. Hillage *et al*[10] suggested a comprehensive strategy for improving the evidence base, including ensuring evaluation was included in the planning of interventions; developing a stronger evidence base by conducting more systematic reviews and meta-analyses; and adopting minimum standards. Worthen[11] has predicted that the use of evaluation will increase in years to come and the approaches used will evolve to become more suited to applications within occupational health and safety.

Taking into account the findings of the literature review, it has been concluded that a suitable research design should posses the following features:

- A base of currently accepted UK national guidance on safety management;
- An economic evaluation component;
- A linkage between the causes of incidents and accidents to the health and safety management system;
- An intervention in the form of written guidance or work book format, supplemented by an educational component;
- A quasi-experimental or experimental design with baseline measurements;
- A comparison group;
- Random sampling where possible;
- Validated methods;
- A control strategy for potential inter and intra-observer bias;
- More than one performance evaluation method including both objective and subjective methods;
- An attempt to correlate the results of the performance evaluation methods (triangulation);
- Sample sizes of sufficient size to allow statistical inferences to be made, where possible;
- Validity checks to ensure that pooling of data from different sites was appropriate before proceeding with analysis.

REFERENCES

1. Niven, K.J.M., *A review of the application of health economics to health and safety in healthcare.* Health Policy, 2002. **61**(3): p. 291-304.
2. Hale, A.R., *Editorial: special issue for the 10th anniversary of the safety science group and the tu Delft.* Safety Science, 1997. **26**(1/2): p. 1.
3. Swuste, P. and F. Arnoldy, *The safety adviser/manager as an agent of organisational change: a new challenge to expert training.* Safety Science, 2003. **41**: p. 15-27.
4. McGuire, T., M. Hanson, and J. Moody, *Management attitudes throughout Scotland towards the funding, provision and use of mechanical aids for moving and handling purposes*. 1995, Lothian NHS Occupational Health Service: Edinburgh. p. 55.
5. Smallman, C., *The reality of "Revitalizing Health and Safety".* Journal of Safety Research, 2001. **32**: p. 391-439.
6. Sheehan, J.J., *Cost-benefit anaysis: a technique gone awry*, in *Legal and ethical dilemmas in occupational health*, J.S. Lee and W.N. Rom, Editors. 1983, Ann Arbor Science: The Butterworth Group. p. 51-215.
7. Gressel, M.G., *An engineer's perspective of the intervention research workshop.* Am J Ind Med, 1996. **29**: p. 382-383.
8. Hale, A. and P. Swuste, *Safety rules: procedural freedom or action constraint?* Safety Science, 1998. **29**: p. 163-177.
9. Hale, A. and P. Swuste, *Avoiding square wheels: international experience in sharing solutions.* Safety Science, 1997. **25**(1-3): p. 3-14.
10. Hillage, J., et al., *The impact of the HSC/E: A review*. HSE Contract Research Report 385/2001. 2001, Sudbury: HSE Books. 74.
11. Worthen, B.R., *Whither evaluation? That all depends.* American Journal of Evaluation, 2001. **22**(3): p. 409-418.
12. DeRoo, L.A. and R.H. Rautiainen, *A systematic review of farm safety interventions.* Am. J. Prev. Med., 2000. **18**(4S): p. 51-62.
13. Chilton, S., et al., *Valuing health and safety controls: A literature review*. HSE Contract Research Report 171/1998. 1998, Sudbury: HSE Books. 135.
14. ACSNI Study Group on Human Factors, *Third report: Organising for Safety*. 1998, Sudbury: HSE Books. 92.
15. Beahler, C.C., J.J. Sundheim, and N.I. Trapp, *Information retreival in systematic reviews - challenges in the public health arena.* Am. J. Prev. Med., 2000. **18**(4S): p. 6-10.
16. Rivara, F.P. and D.C. Thompson, *Systematic reviews of injury-prevention strategies for occupational injuries.* Am. J. Prev. Med., 2000. **18**(4(S)): p. 1-3.
17. Livingston, A.D., G. Jackson, and K. Priestley, *Root causes analysis: Literature review*. HSE Contract Reseach Report No. 325/2001. 2001, Sudbury: HSE Books. 53.
18. Niven, K.J.M., *Accident costs in the NHS.* The Safety & Health Practitioner, 1999. **September**: p. 34-38.
19. Department of Health, *Building a safer NHS for patients*. 2001, www.doh.gov.uk/buildsafenhs/.
20. Department of Health, *An organisation with a memory: report of an expert group on learning from adverse events in the NHS*. 2000, London: The Stationary Office. 89.
21. Neale, G., M. Woloshynowych, and C. Vincent, *Exploring the causes of adverse events in NHS hospital practice.* Journal of the Royal Society of Medicine, 2001. **94**: p. 322-330.
22. Andersen, B. and T. Fagerhaug, *Root cause analysis: simplified tools and techniques*. 2000, Milwaukee, WI: www.amazon.com.
23. Haslam, R.A. and T.A. Bentley, *Follow-up investigations of slip, trip and fall accidents among postal delivery workers.* Safety Science, 1998. **32**: p. 33-47.
24. Jacinto, C. and E. Aspinwall, *WAIT - A new method for the investigation and analysis of accidents at work.* Journal of the Institution of Occupational Safety and Health, 2002. **6**(1): p. 15-37.
25. Reason, J., *Managing the risks of organisational accidents*. 1997, Aldershot: Ashgate.

26. St. Vincent, M., C. Tellier, and T. Petitjean-Roget, *Accidents that occurred in three hospitals in one year.* Safety Science, 1999. **31**: p. 197-212.
27. Department of Health, *Doing less harm: improving the safety and quality of care through reporting, analysing and learning from adverse incidents involving NHS patients - key requirements for health care providers*. 2001: National Patient Safety Agency.
28. National Patient Safety Agency, *NPSA pilot project evaluation report*. 2002: National Patient Safety Agency. 147.
29. Cowan, J., *Achieving consistency in grading adverse clinical incidents: does Doing Less Harm do the job?* British Journal of Clinical Governance, 2002. **7**(1): p. 63-67.
30. Hartley, C. and R.T. Booth, *The performance of health and safety specialists in a simulated risk assessment task.* Journal of the Institution of Occupational Safety and Health, 2000. **4**(2): p. 7-20.
31. Henderson, J., C. Whittington, and K. Wright, *Accident investigation - the drivers, methods and outcomes*. HSE Contract Research Report No. 344/2001. 2001, Sudbury: HSE Books. 432.
32. Rakel, H., et al., *Evaluating the impact of contact techniques*. HSE Contract Research Report 212/1999. 1999, Sudbury: HSE Books. 81.
33. Scherer, R.F., J.D. Brodzinski, and A.L. Canty, *An examination of process and outcome differences in health and safety inspections conducted by state and federal agencies.* Journal of Safety Research, 1997. **28**(3): p. 203-211.
34. Pransky, G., et al., *Under-reporting of work-related disorders in the workplace: a case study and review of the literature.* Ergonomics, 1999. **42**(1): p. 171-182.
35. Herrero, S.G., et al., *From the traditional concept of safety managment to safety integrated with quality.* Journal of Safety Research, 2002. **33**: p. 1-20.
36. Kjellen, U., *An evaluation of safety information systems at six medium-sized and large firms.* J. Occup. Accidents, 1982. **3**: p. 273-288.
37. Budworth, N., *Indicators of performance in safety management.* The Safety & Health Practitioner, 1996. **November**: p. 23-29.
38. Health and Safety Commission Health Services Advisory Committee, *Management of health and safety in the health services - information for directors and managers*. 1994, Sudbury: HSE Books. 29.
39. Health and Safety Executive, *Successful health and safety management (HSG65)*. 1st ed. Health and safety series booklet HS(G)65. 1991: HSE Books. 72.
40. Health and Safety Executive, *Successful health and safety management (HSG65)*. 2nd ed. 1997, Sudbury: Health and Safety Executive. 98.
41. Hale, A.R., et al., *Modelling of safety management systems.* Safety Science, 1997. **26**(1/2): p. 121-140.
42. Department of Health, *The effective management of occupational health and safety services in the NHS*. 2002, London: The Stationary Office. 89.
43. Emslie, S.V. and R.F. Lowe, *Safecode and successful management of health and safety risk in the national health service.* Health Informatics, 1995: p. 91-100.
44. National Audit Office, *A safer place to work - Improving the management of health and safety risks to staff in NHS trusts*. 2003, London: The Stationary Office.
45. Scottish Executive, *Towards a safer healthier workplace*. 1999, http://www.scotland.gov.uk/library2/doc08/shwm-00.htm.
46. Chaplin, R. and A.R. Hale, *An evaluation of the use of the international safety rating system (ISRS) as intervention to improve the organisation of safety*, in *Safety Management. The Challenge of Change*, A.R. Hale and M. Baram, Editors. 1998, Pergamon: London. p. 165-185.
47. Mearns, K., S.M. Whitaker, and R. Flin, *Safety climate, safety management practice and safety performance in offshore environments.* Safety Science, 2003. **41**(8): p. 641-680.
48. DNV, *The International Safety Rating System*. 5th Edition ed. 1990, Rotterdam: Det Norske Veritas.
49. Top, W., *The International Safety Rating System (Part 1)*. 1986, Breda: Loss Control Institute NL.

50. Guastello, S.J., *Do we really know how well our occupational accident prevention programs work?* Safety Science, 1993. **16**: p. 445-463.
51. Wagenaar, W.A., J. Groeneweg, and P. Hudson, *Promoting safety in the oil industry.* Ergonomics, 1994. **37**(12): p. 1999-2013.
52. Hudson, P., et al., *Tripod Delta: Proactive approach to enhanced safety.* Journal of Petroleum Technology, 1994. **46**(1): p. 58-62.
53. Wagenaar, W.A., P. Hudson, and J. Reason, *Cognitive failures and accidents.* Applied Cognitive Psychology, 1990. **4**: p. 273-294.
54. HASTAM, *The Complete Health and Safety Evaluation (CHASE)*. 1994, Birmingham: HASTAM Ltd.
55. British Standards Institution, *OHSAS 18001:1999 Occupational health and safety management systems - Specification*. 1999, British Standards Institution: London.
56. Scotney, V., *Development of a health and safety performance measurement tool*. HSE Contract Research report 309/2000. 2000, Sudbury: HSE Books. 222.
57. Fuller, C.W., *Key performance indicators for benchmarking health and safety management in intra- and inter- company coomparisons.* Benchmarking for Quality Management & Technology, 1997. **4**(3): p. 165-174.
58. Health and Safety Executive, *Successful Health and Safety Management*. first ed. Health and safety series booklet HS(G)65. 1991: HSE Books. 72.
59. British Standards Institution, *Guide to occupational health and safety management systems; BS 8800*. 1996, British Standards Institution: London.
60. Eisner, H. and J. Leger, *The International Safety Rating System in South African Mining.* J. Occup. Accidents, 1988. **10**: p. 141-160.
61. Guldenmund, F.W., *The nature of safety culture: a review of theory and research.* Safety Science, 2000. **34**: p. 215-257.
62. Hale, A.R. and J. Hovden, *Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment*, in *Occupational Injury. Risk, Prevention and Intervention P129-165*. 1998.
63. Cooper, M.D., *Towards a model of safety culture.* Safety Science, 2000. **36**: p. 111-136.
64. Flin, R., et al., *Measuring safety climate: identifying the common features.* Safety Science, 2000. **34**: p. 177-192.
65. Health and Safety Executive, *Health and safety climate survey tool - Information pack*. MISC097. 1997, Sudbury: HSE Books.
66. Health and Safety Executive, *Health and safety climate survey tool - process guidelines*. 1998, Sudbury: HSE Books.
67. Davies, F., R. Spencer, and K. Dooley, *Summary guide to safety climate tools*. HSE Contract Research Report 1999/063. 2001, Sudbury: HSE Books. 45.
68. Bailey, C. and D. Petersen, *Using perception surveys to assesss safety system effectiveness.* Professional Safety, 1989. **February**: p. 22-26.
69. Donald, I. and D. Canter, *Employee attitudes and safety in the chemcial industry.* Journal of Loss Prevention in the Process Industries, 1994. **7**(3): p. 203-208.
70. Hayes, B.E., et al., *Measuring perceptions of workplace safety: development and validation of the work safety scale.* Journal of Safety Research, 1998. **29**(3): p. 145-161.
71. Emory, C.W. and D.R. Cooper, *Business Research Methods (4th Edition)*. 4th Edition ed. 1991: Irwin.
72. Easterby-Smith, M., R. Thorpe, and A. Lowe, *Management research: An introduction*. 1991, London: Sage Publications Ltd. 172.
73. Niskanen, T., *Safety climate in the road administration.* Safety Science, 1994. **17**: p. 237-255.
74. Rundmo, T. and A.R. Hale, *Managers' attitudes towards safety and accident prevention.* Safety Science, 2003. **41**(7): p. 557-574.
75. Smallman, C. and G. John, *British directors perspectives on the impact of health and safety on corporate performance.* Safety Science, 2001. **38**: p. 227-239.

76. Vassie, L.H. and W.R. Lucas, *An assessment of health and safety management within working groups in the UK manufacturing sector.* Journal of Safety Research, 2001. **32**(479-490).
77. Zohar, D., *Safety climate in industrial organisations: Theoretical and applied implications.* Journal of Applied Psychology, 1980. **65**(1): p. 96-101.
78. Brown, R. and H. Holmes, *The use of a factor-analytic procedure for assessing the validity of an employee safety climate model.* Accid. Anal. & Prev., 1986. **18**(6): p. 455-470.
79. Dedobbeleer, N. and F. Beland, *A safety climate measure for construction sites.* J. Safety Res., 1991. **22**: p. 97-103.
80. Basen-Enquist, K., et al., *Worksite health and safety climate: Scale development and effects of a health promotion intervention.* Preventitive Medicine, 1998. **27**: p. 111-119.
81. Gershon, R.M., et al., *Hospital safety climate and its relationship with safe work practices and workplace exposure incidents.* American Journal of Infection Control, 2000. **28**(3): p. 211-221.
82. Pilkington, A., et al., *Baseline measurements for the evaluation of the work-related stress campaign*. HSE Contract Research Report 322/2001. Vol. Prepared by the Institute of Occupational Medicine for the HSE. 2001, Subury: HSE Books. 83.
83. Pilkington, A., et al., *Survey of use of occupational health support*. HSE Contract Research Report 445/2002. Vol. Prepared by the Institute of Occupational Medicine for HSE. 2002, Sudbury: HSE Books. 130.
84. Cox, S.J. and A.J.T. Cheyne, *Assessing safety culture in offshore environments.* Safety Science, 2000. **34**: p. 111-129.
85. Centre for Hazard and Risk Management, *Offshore safety climate assessment toolkit*. 2000, Loughborough University.
86. Drummond, M.F., G.L. Stoddart, and W. Torrance, *Methods for the Economic Evaluation of Health Care Programmes*. Second ed. 1997, Oxford: Oxford University Press. 305.
87. Chilton, S., et al., *Valuation of benefits of health and safety control*. HSE Contract Research report 273/2000. 2000, Sudbury: HSE Books.
88. Warner, K.E., et al., *Health and economic implications of a work site smoking cessation program: A simulation analysis.* J. Occup. Environ. Med., 1996. **38**(10): p. 981-992.
89. Cooper, C.L., P. Liukkonen, and S. Cartwright, *Stress prevention in the workplace: assessing the costs and benefits to organisations*. European Foundation for the Improvement of Living and Working Conditions. 1996, Luxembourg: Office for Official Publications of the European Communities. 110.
90. Kemmlert, K., *Economic impact of ergonomic intervention - four case studies.* J. Occup. Rehabil., 1996. **6**(1): p. 17-32.
91. Monnery, N., *The costs of accidents and work-related ill-health to a cheque clearing department in a financial services organisation.* Safety Science, 1999. **31**: p. 59-69.
92. Health and Safety Executive, *The costs of accidents at work*. Health and Safety Series booklet HS(G)96. 1993, Sudbury: HSE Books. 51.
93. Hurst, N.W., et al., *Measures of safety performance and attitudes to safety at major hazard sites.* Journal of Loss Prevention in the Process Industry, 1996. **9**(2): p. 161-172.
94. Cabrera, D.D., R. Isla, and L.D. Viledla. *An evaluation of safety climate in ground handling activities*. in *Aviation Safety*. 1997. Proceedings of the IASC-97 International Aviation Safety Conference, Netherlands.
95. Jacobs, R. and S. Haber, *Organizational processes and nuclear power plant safety.* Reliability Engineering and System Safety, 1994. **45**: p. 75-83.
96. Haber, S., D. Metlay, and D. Crouch, *Influence of organizational factors on safety.* Proceedings of the Human Factors Society, 1990: p. 871-875.
97. Shannon, H.S., J. Mayr, and T. Haines, *Overview of the relationship between organizational and workplace factors and injury rates.* Safety Science, 1997. **26**(3): p. 201-217.
98. Shannon, H.S., L.S. Robson, and S.J. Guastello, *Methodological criteria for evaluating occupational safety intervention research.* Safety Science, 1999. **31**: p. 161-179.
99. Cox, T., R. Randall, and A. Griffiths, *Interventions to control stress at work in hospital staff*. HSE Contract Research Report 435/2002. Vol. Prepared by the Institute of Work, Health and

Organisations for the HSE, Unison and the Royal College of Nursing. 2002, Sudbury: HSE Books. 160.
100. Hale, A.R., et al., *Relative importance of maintenance management influences on equipment failure and availability in relation to major hazards*. Safety and Reliability, ed. G. Schueller and P. Kafka. 1999, Rotterdam: Balkema. p1327-1332.
101. Health and Safety Commission, *Health and safety targets: How are we doing? A supplement to the HSC annual report and HSC/E accounts 2001/02*. 2002, Sudbury: HSE Books.
102. Health and Safety Commission Strategic Plan 2001 -2004, *http://www.hse.gov.uk/aboutus/plans/hscplans/plan0104.htm*. 2001.
103. Department of Health, *Preventing accidental injury - Priorities for action*. 2002: ISBN 0 11 322477 X, Norwich: The Stationary Office.
104. Health and Safety Commission, *Management of health and safety at work regulations*. 1992: The Stationary Office.
105. Smith, M., et al., *Characteristics of successful safety programs.* J. Safety Res., 1978. **10**(1): p. 5-15.
106. Simonds, R. and Y. Shafai-Sahrai, *Factors apparently affecting injury frequency in eleven matched pairs of companies.* J. Safety Res., 1977. **9**(3): p. 120-127.
107. Cohen, A., *Factors in successful occupational safety programs.* J. Safety Res., 1977. **9**(4): p. 168-178.
108. Shafai-Sahrai, Y., *Determinants of occupational injury experience: A study of matched pairs of companies*. 1973, East Lansing, Michigan: MSU Business Studies.
109. Cohen, A., M. Smith, and C. MH., *Safety program practices in high versus low accident rate companies - An interim report*. 1975, National Institute for Occupational Safety and Health (NIOSH). p. 1-184.
110. Hovden, J. and R. Tinmannsvik, *Internal control: A strategy for occupational safety and health. Experiences from Norway.* J. Occup. Accidents, 1990. **12**: p. 21-30.
111. Ronald, L., *Identifying the elements of successful safety programs: A literature review*. 1999, Workers' Compensation Board of British Columbia: Richmond, British Columbia. p. 49.
112. Saari, J., ed. *Successful accident prevention: recommendations and ideas field tested in the Nordic countries. The final report of the Nordic cooperative program of effective accident prevention methods*. 1987, Institute of Occupational Health: Helsinki.
113. British Standards Institution, *OHSAS 18002:2000 Occupational health and safety management systems - Guidelines for the implementation of OHSAS 18001*. 2000, British Standards Institution: London.
114. Smith, D., G. Hunt, and C. Green, *Managing safety the systems way. BS8800 to OHSAS 18001 - Implementing a cost-effective management system for occupational health and safety*. Second ed. 2000, London: British Standards Institution. 127.
115. Nichols, M. and A. Marcus, *Empirical studies of candidate leading indicators of safety in nuclear power plants: An expanded view of human factors research.* Proceedings of the Human Factors Society, 1990: p. 876-880.
116. Modarres, M., A. Mosleh, and J. Wreathall, *A framework for assessing influence of organization on plant safety.* Reliability Engineering and System Safety, 1992. **38**: p. 157-171.
117. Swuste, P., A. Hale, and F.W. Guldenmund, *Change in a steeel works: learning from failure and particular successes*, in *System Safety: Challenges and Pitfalls of Intervention*, B. Wilpert and B. Fahlbruch, Editors. 2002, Pergamon: Oxford. p. 135-158.
118. Cumming, R.G., et al., *Home visits by an occupational therapist for assessment and modifaction of environmental hazards: A randomised trial of falls prevention.* J Am Geriatr Soc, 1999. **47**(12): p. 1397-1402.
119. Pelletier, K.R., et al., *Managing job strain: A randomised controlled trial of an intervention conducted by mail and telephone.* J Occup Environ Med, 1999. **41**(4): p. 216-223.
120. Daltroy, L.H., et al., *A controlled trial of an educational program to prevent low back injuries.* N Engl J Med, 1997. **337**(5): p. 322-328.

121. Spangenberg, S., et al., *The construction of the Oresund link between Denmark and Sweden: The effect of a multi-faceted safety campaign.* Safety Science, 2002. **40**(5): p. 457-465.
122. Goldenhar, L.M. and P.A. Schulte, *Intervention research in occupational health and safety.* J. Occup. Med., 1994. **36**(7): p. 763-775.
123. Health and Safety Executive, *Reducing error and influencing behaviour*. HSG48. 1999, Sudbury: HSE Books. 88.
124. Komaki, J., K. Barwick, and L. Scott, *A behavioural approach to occupational safety: Pinpointing and reinforcing safe performance in a food manufacturing plant.* Journal of Applied Psychology, 1978. **63**(4): p. 434-445.
125. Cohen, A., M. Smith, and W. Anger, *Self-protective measures against workplace hazards.* J. Safety Res., 1979. **11**(3): p. 121-131.
126. Nasanen, M. and J. Saari, *The effects of positive feedback on housekeeping and accidents at a shipyard.* J. Occup. Accidents, 1987. **8**: p. 237-250.
127. Reason, J., *Human Error*. 1989: Cambridge University Press.
128. Hale, A.R., *Safety rules O.K.? Possibilities and limitations in behavioural safety strategies.* J. Occup. Accidents, 1990. **12**: p. 3-20.
129. Simard, M. and A. Marchand, *A multilevel analysis of organisational factors related to the taking of safety initiatives by work groups.* Safety Science, 1995. **21**: p. 113-129.
130. Petrea, R.E., *The theory of planned behaviour: Use and application in targeting agricultural safety and health interventions.* Journal of Agricultural Safety and Health, 1999. **7**(1): p. 7-19.
131. The Keil Centre, *Behaviour modification to improve safety: Literature review*. HSE Offshore Technology Report No. 2000/003. 2000, Sudbury: HSE Books. 20.
132. Ogden, J., *Health psychology - a textbook*. second ed. 2000, Buckingham: Open University Press. 396.
133. Fleming, M. and R. Lardner, *Behaviour modification programmes: Establishing best practice*. HSE Contract Research Report 2000/048. 2001, Sudbury: HSE Books. 27.
134. Fleming, M. and R. Lardner, *Strategies to promote safe behaviour as part of a health and safety management system*. HSE Contract Research Report 430/2002. Vol. Prepared by the Keil Centre for the HSE. 2002, Sudbury: HSE Books. 74.
135. Clarke, M. and A.D. Oxman, eds. *Cochrane reviewers handbook 4.1.4 [Updated October 2001]*. In: The Cochrane Library, Issue 4, 2001. 2001: Oxford: Update Software: Updated Quarterly.
136. Elvik, R., *Assessing the validity of evaluation research by means of meta-analysis*. 1999, Institute of Transport Economics: Oslo. p. 187.
137. Crombie, I.K. and H.T.O. Davies, *Research in healthcare: Design, conduct and interpretation of health services research*. 1996, Chichester: Wiley. 288.
138. Krause, T.R., K.J. Seymour, and K.C.M. Sloat, *Long-term evaluation of a behaviour-based method for improving safety performance: a meta-analysis of 73 interrupted time-series replications.* Safety Science, 1999. **32**: p. 1-18.
139. Oliver, D., A. Hopper, and P. Seed, *Do hospital fall prevention programs work? A systematic review.* J Am Geriatr Soc, 2000. **48**(12): p. 1679-1689.
140. Rogers, B. and L. Goodno, *Evaluation of interventions to prevent needlestick injuries in health care occupations.* Am. J. Prev. Med., 2000. **18**(4(S)): p. 90-98.
141. Rivara, F.P. and D.C. Thompson, *Prevention of falls in the construction industry: evidence for program effectiveness.* Am. J. Prev. Med., 2000. **18**(4S): p. 23-26.
142. Lingard, H. and S. Rowlinson, *Behaviour-based safety management in Hong Kong's construction industry.* Journal of Safety Research, 1997. **28**(4): p. 243-256.
143. Harper, A.C., et al., *Curtin industrial safety trial: methods and safe practice and housekeeping outcomes.* Safety Science, 1996. **24**(3): p. 159-172.
144. National Institute for Occupational Safety and Health, *Guide to evaluating the effectiveness of strategies for preventing work injuries: how to show whether a safety intervention really works*. 2001, Cincinnati: NIOSH. 124.

145. Yassi, A., et al., *Early intervention for back injured nurses at a large Canadian tertiary care hospital: an evaluation of the effectiveness and cost benefits of a two-year pilot project.* Occup. Med., 1995. **45**(4): p. 209-214.
146. Carter, N. and E. Menckel, *Group routines for improving accident prevention activities and accident statistics.* Int. J. Ind. Erg., 1990. **5**(125-132).
147. Volinn, E., *Do workplace interventions prevent low-back disorders? If so, why?: a methodologic commentary.* Ergonomics, 1999. **42**(1): p. 258-272.
148. Bracker, A., et al., *Industrial hygiene recommendations as interventions: A collaborative model within occupational medicine.* Applied Occupational & Environmental Hygiene, 1999. **14**: p. 85-96.
149. McGrail, M.P., S.P. Tsai, and E.J. Bernacki, *A comprehensive initiative to manage the incidence and cost of occupational injury and illness: report of an outcomes analysis.* J. Occup. Environ. Med., 1995. **37**(11): p. 1263-1268.
150. Barry, E., et al., *Preventing accidental falls among older people in long stay units.* Irish Medical Journal, 2001. **94**(6): p. 172-176.
151. Brown, D., *Initiative evaluation report back in work*. HSE Contract Research Report 441/2002. Vol. Prepared by Corporate Solutions Consulting (UK) Ltd. for HSE. 2002, Sudbury: HSE Books. 185.
152. Horbury, C. and D. Collier, *The effectiveness and impact of the PABIAC initiative in reducing accidents in the paper industry*. HSE Contract Research Report 452/2002. Vol. Prepared by Greenstreet Berman Ltd for HSE. 2002, Sudbury: HSE Books. 122.
153. Shivers, C.H., *Halos, horns and hawthorne: Potential flaws in the evaluation process.* Professional Safety, 1998. **March**: p. 38-41.
154. Harper, A.C., et al., *Curtin industrial safety trial: managerial behaviour and program effectiveness*. Safety Science, 1996. **24**(3): p. 173-179.
155. Spangenberg, S., et al., *Factors contributing to the differences in work related injury rates between Danish and Swedish construction workers.* Safety Science, 2003. **41**(6): p. 517-530.
156. Ludbrook, A., et al., *Effective and cost-effective measures to reduce alcohol misuse in Scotland: A literature review*. Research Findings, ed. S.E.H. Department. 2001, Edinburgh: The Stationary Office. 174.
157. Parkes, K. and T.J. Sparkes, *Organisational interventions to reduce work stress: are they effective? A review of the literature*. HSE Contract Research Report 193/1998. 1998, Sudbury: HSE Books. 52.
158. Marsh, T., et al., *The role of management commitment in determining the success of a behavioural safety intervention.* Journal of the Institution of Occupational Safety and Health, 1998. **2**(2): p. 45-56.
159. Tam, C.M. and I.W.H. Fung, *Effectiveness of safety management strategies on safety performance in Hong Kong.* Construction Management and Economics, 1998. **16**: p. 49-55.
160. McGuire, T., *Managers' attitudes towards mechanical aids.* Nurs. Stand., 1997. **11**(31): p. 33-38.
161. Budworth, N., *The development and evaluation of a safety climate measure as a diagnostic tool in safety management.* Journal of the Institution of Occupational Safety and Health, 1997. **1**(1): p. 19-29.
162. Fagard, R.H., J.A. Staessen, and L. Thijs, *Advantages and disadvantages of the meta-analysis approach.* J. Hypertens. Suppl., 1996. **14**(2): p. S9-12.
163. Crombie, I.K., *The pocket guide to critical appraisal: a handbook for health care professionals*. 1996, London: BMJ Publishing Group. 66.
164. Stufflebeam, D.L., *Evaluation checklists: Practical tools for guiding and judging evaluations.* American Journal of Evaluation, 2001. **22**(1): p. 71-79.