York Health Economics Consortium

LLOYD'S REGISTER FOUNDATION

Scoping Review of the Effectiveness of Leading Indicators for Improving Safety Outcomes in Occupational Safety and Health Settings: What is the State of the Art?

Final Report

DR LAVINIA FERRANTE DI RUFFANO, Project Director

DEBORAH WATKINS, Research Consultant

PAUL MILLER, Information Specialist

EMMA BISHOP, Research Assistant

10/04/2024

Investors in People | Health & Wellbeing Award

UNIVERSITY of York

Table of Contents

Executive Summary **1**
Acknowledgements **8**
Abbreviations **9**
1 Introduction **11**
1.1 Background 11
1.2 Review Objectives 11
2 Methods **12**
2.1 Eligibility Criteria 12
2.2 Identification of Relevant Studies 13
2.3 Selection of Relevant Studies 13
2.4 Data Charting 13
2.5 Risk of Bias Assessment 13
2.6 Synthesis and Reporting 14
3 Results **15**
3.1 Included Study Characteristics 15
3.2 Effectiveness of Leading Indicators 44
4 Discussion **48**
4.1 Summary of Key Findings 48
4.2 Strengths of the Review 51
4.3 Limitations of the Review 52
4.4 Implications for Practice 53
4.5 Future Research 53
4.6 Conclusion 55
5 References **56**
Appendix A: Review Methods **61**
A.1 Eligibility Criteria 61
A.2 Identification of Relevant Studies 63
A.3 Selection of Relevant Studies 66
A.4 Data Charting 66
A.5 Risk of Bias Assessment 68
Appendix B: Search Strategies **69**
Appendix C: Identification of Studies **73**
C.1 Literature Search Results 73
C.2 Study Selection 75
Appendix D: Excluded Studies **76**
Appendix E: Full Risk of Bias Assessment **90**
Appendix F: Included Studies **114**
Appendix G: Detailed Results Tables **117**
Appendix H: Summary of Brandt 2023 ..... **175**

Executive Summary

1. INTRODUCTION

The International Labour Organization estimates that globally 395 million workers sustain non-fatal work injuries and there are almost three million work-related deaths per year [1]. Safety performance has traditionally been assessed using lagging indicators, such as work-related illnesses and fatalities, which measure the occurrence of past events [2]. Since lagging indicators are measures of past occurrences, they cannot be used to identify which factors contribute to safety performance.

Leading indicators have been identified to be precursors of lagging indicators and are considered to be conditions, events or measures that precede undesirable events and have value in predicting or preventing the event's arrival [3]. They are defined as proactive, preventative and predictive measures that inform how effective health and safety practices are [2, 4]. While implementing successful leading indicators should lead to improvements in lagging indicators of safety performance, existing studies have struggled to find any significant relationships and there are no known literature reviews on the topic [5, 6].

Lloyd's Register Foundation (the Foundation) is supporting research into leading indicators as part of its charitable mission to enhance the safety of workers globally, and across industries. In order to address the absence of reviews in this area of research, the Foundation has commissioned a rapid evidence assessment of the available evidence base for leading indicators.

This scoping review aimed to map the key characteristics of studies that assess the impact of leading indicators on lagging indicators in order to understand:

- For which leading indicators has effectiveness been evaluated?
- For which industries and in which countries does this evidence come from?
- Are different leading indicators evaluated in lower (LI) to middle income (MI) countries, compared to high income (HI) countries?
- How has effectiveness been evaluated?
- Is there any evidence for the effectiveness of leading indicators, and what is the nature of that evidence?
- How robust is this evidence base, and how can it be improved?

2. METHODS

Studies published in the English language from 2010 onwards and evaluating the impact of any leading indicator on workplace safety outcomes were eligible. For the purpose of this scoping review we defined a safety leading indicator as a proactive measure that can be used to predict current or future safety performance (e.g. safety audits, training, corrective action measurements).

We searched Scopus, Web of Science Core Collection and the National Institute for Occupational Safety and Health (NIOSH) NIOSHTIC-2 database on 1st August 2023. Study selection, data extraction and risk of bias (RoB) assessment were carried out by a single reviewer, with a second reviewer checking 10% of study selection decisions and the data extraction and RoB assessments for 20% of included studies. Studies were summarised in tables and discussed narratively to explore the quality of the studies, the relationship between studies and patterns that we discerned in the data.

3. RESULTS

Study selection

Following deduplication, 4,339 individual records were identified and assessed for relevance. Of these, 4,073 were excluded at first pass and title and abstract review. The full texts of the remaining 266 records were sought for retrieval but 50 were unobtainable. The remaining 216 full texts were examined and 48 studies were considered eligible and included in this review: 5 cohort studies, 12 before-after studies, 6 case series, 17 cross-sectional studies and 8 case studies.

Study characteristics

The included studies were undertaken in the following industries: construction (n=18), energy (n=13), mining (n=3), automotive (n=1), dairy manufacturing (n=1), logging (n=1), maritime (n=1), newspaper publishing (n=1), mixed (n=8), not reported (n=1). Studies included data from between 1 and 1,180 companies, although this detail was not always reported. Details on worker characteristics were scarcely reported. Twenty-seven studies were carried out in HI countries, 4 in upper middle income (UMI), 4 in lower middle income (LMI) and 13 studies were carried out in multiple countries or did not report the country.

Leading and lagging indicators

Following categorisation, five different types of leading indicator were identified as having been evaluated in the included studies: new guideline, tool or process (n=15); safety climate/culture (n=5); audits and inspections (n=2); monitoring of safety (n=1); mixed (n=25).

Each included study evaluated between 1 and 11 lagging indicators, with 36 studies reporting more than one. The lagging indicators were grouped into 10 categories: injuries (n=25), accidents (n=18), incidents (n=14), near misses (n=8), lost time (n=6), fatalities (n=5), compensation (n=4), sickness/illness (n=3), safety (n=1) and ‘other’ (n=4 studies). Though multiple studies evaluated the same lagging indicators, they were usually measured using different methods and therefore could not be collated.

Study validity

At least one study design weakness was identified in each study, and so no studies were considered at very low RoB. Seven studies were determined to be at low RoB, 27 at moderate RoB and 14 at high RoB. Poor reporting was a common issue in the included studies, and incomplete information on study methods posed challenges to identifying and assessing how studies were conducted.

External validity of the studies was difficult to assess due to limited reporting of key information, but overall, generalisability is likely to be weak. Common reasons for this were studies reporting data from only one company or evaluating leading indicators that were specific to the company or industry.

Effectiveness of leading indicators

Overall, 27 studies found that at least some of the leading indicators evaluated were favourably associated with lagging indicators (in 20 studies it was unclear whether there was an effect and in only one was there not an effect). However, studies were too heterogeneous in the methods used and lagging indicators measured to present a meaningful distribution of leading indicator effects. All 15 studies evaluating a new guideline, tool or process found an association with lagging indicators however only three reported on the statistical significance of this association. All five studies assessing safety climate/culture reported a favourable association with lagging indicators but only four reported statistical significance. While two studies evaluated the impact of audits and inspections and both reported statistical significance, one concluded that there was no relationship between leading and lagging indicators. One study reported on the relationship between monitoring of safety and lagging indicators and found a statistically significant relationship. Sixteen of the 25 studies that evaluated mixed leading indicators reported statistically significant relationships with leading indicators.

4. DISCUSSION

Overall, the studies identified by this review were found to be very heterogeneous in every study characteristic examined. Reporting of study characteristics was also variable, with key information missing from studies, particularly concerning worker characteristics.

For which leading indicators has effectiveness been evaluated?

Due to variations in terminology, it was challenging to identify whether studies had measured the same leading indicator. We assessed that the included studies evaluated a large number of different leading indicators, representing the wide range of safety practices. The most commonly evaluated leading indicator was the implementation of some type of guideline, framework, tool or process (n=15). More than half of the included studies (n=25) evaluated multiple leading indicators either individually or through looking at composite scores. Across all studies, leading indicators were often specific to either the industry or the company being evaluated. The leading indicators evaluated in this evidence base did not all fully meet previously published definitions of leading indicators from the Campbell Institute 2015 and Xu 2021.

For which industries and in which counties does this evidence come from?

Studies were carried out across eight different industries (most commonly construction [n=18] or energy [n=13]) with a small number evaluating leading indicators across multiple (or in some cases, any) industry. Country settings were inconsistently reported but of the 39 studies providing this information, 22 different countries across all five continents were represented.

Are different leading indicators evaluated in lower to middle income countries, compared to high income countries?

The World Bank classification of country income status could not be assessed for 12 studies due to lack of reporting. Of the remaining studies, the majority (27/48) took place in HI settings, four UMI, four LMI and one study took place in eight countries ranging from LMI to HI. There did not appear to be any clear association between the type of leading indicator evaluated and country income status.

How has effectiveness been evaluated?

Eligible studies used one of five observational study designs to assess the impact of leading on lagging indicators, most commonly cross-sectional studies, but also cohort studies, before-after studies, case series and case studies.

Methods of analysis varied across the 48 studies. Linear correlation and simple comparisons of frequencies were the most commonly used methods. Other methods, such as modelling approaches and interrupted time series, were each used by a single study. Leading indicators were evaluated against a wide variety of lagging indicators covering 10 categories (injuries, accidents, incidents, near misses, lost time, fatalities, compensation claims, sickness/illness, safety and other outcomes). The methods of defining and measuring these indicators differed considerably between studies, restricting the meaningful synthesis of effects in future systematic reviews.

Is there any evidence for the effectiveness of leading indicators, and what is the nature of that evidence?

All but one included study reported a positive effect of leading indicator on lagging indicator outcomes. However, results were very varied and while most studies reported multiple analyses, none found all analyses to produce evidence of an effect. Additionally, claims of impact were not always verified through appropriate statistical tests, therefore this review concluded that evidence of impact of leading on lagging indicators was produced by 27 out of 48 studies.

How robust is this evidence base, and how can it be improved?

Overall, the quality of the included studies means that their ability to demonstrate that leading indicators have causal impacts on lagging indicators is weak, for three reasons:

1. The evidence base for this review consisted entirely of observational studies which are not designed to demonstrate causality or draw conclusions as robust as a controlled, comparative study design could.
2. The internal validity of included studies (whether the results of the study are likely to reflect true differences) was determined to be moderate to low, indicating findings might be explained by problematic study design. Over 80% of studies included in the review were deemed to be of a moderate or high RoB and therefore are more likely to report distorted estimates of effects.
3. The evidence base was considered to be poorly generalisable across occupational settings due to limited and inconsistent reporting of key study information, and the common evaluation of a single company or company-specific leading indicator.

Strengths and limitations

A strength of this review is the use of explicit, systematic methods captured in a predefined review protocol that was registered on the Open Science Framework (OSF). Furthermore, we consulted with occupational safety and health (OSH) experts from varying backgrounds during the development of the review protocol to ensure validity of research aims, objectives and review methods. The results of our search underscored the absence of existing reviews and therefore makes a unique contribution to the research literature.

Search terms were limited to retrieve records which referred to the concept of leading indicators. Studies that only referred to specific indicators would not have been retrieved. Date and language restrictions were applied in the review. In addition, a large proportion of documents could not be obtained for assessment of relevance at full text stage (50 of 266) which may have resulted in a number of missed eligible studies. Single reviewer study selection and data extraction were used in this review, increasing the possibility of missing studies or incorrectly extracting information. However, both were performed in duplicate for the first 20% of studies to ensure consistency of approach.

The review focussed on published studies, which may have missed relevant unpublished data that demonstrate positive impacts of leading indicators, such as internal company benchmarking data. While we did obtain and examine some unpublished reports, those not known or available to our expert group could not be considered or included. The impact of this to the conclusions in this review cannot be quantified. It is also not possible to determine whether organisations that were represented by the included studies would be judged to have a good safety performance, since this review did not set out to examine the interdependencies between programmes and leading indicators.

The number of included studies that evaluated multiple or mixed leading indicators could suggest that our approach of looking for associations between individual leading indicators and safety outcomes does not reflect real practice where a range of indicators, appropriate to the company's health and safety needs, are tracked in tandem.

Finally, this review was not designed to consider how leading indicators impact occupational health outcomes (including wellbeing), which could in turn affect safety outcomes through improved decision-making and situational awareness. This review can therefore not contribute to an understanding of whether occupational health outcomes could be valid surrogates for downstream impacts to safety.

Implications for practice

This review shows a huge amount of variation in the current evidence base for leading indicators which makes it difficult to identify recommendations for practice. This reflects the findings of previous studies and suggests that practitioners should exercise caution when evaluating leading indicators.

The biggest challenge in evaluating the impact of leading indicators was the inability to compare findings across studies. If comparability across sites, units, companies or industries remains a goal of future research, then more standardised data collection and wider data sharing should be encouraged.

Currently, the greatest value of leading indicators may be for organisations to tailor them specifically for their own setting and perform evaluations that set out to assure the organisation’s key risk control systems. The challenge for this more introspective goal will be to identify a set of indicators that demonstrate the ongoing integrity of that safety and management system.

Future research

There is a clear need for research to address whether leading indicators can improve occupational health, and to investigate whether this outcome could be a valid surrogate for downstream safety performance.

To understand whether specific leading indicators improve safety outcomes across settings, syntheses with statistical pooling of effects would ideally combine the results of prospective, comparative studies that have controlled for confounders, selection bias and measurement bias. To achieve this, standardised metrics to measure and report leading and lagging indicators should be developed.

Primary studies should also report study methods more completely to ensure their results can be applied in practice, and to help understand the reliability of their results. Future research is also needed to understand whether the tool used to appraise the methodological quality of studies in this review captures the most important potential sources of bias in OSH studies of leading indicators.

Creating a more comparable evidence base would also enable researchers to identify and investigate reasons for not observing expected associations with lagging indicators. Further research is needed to determine which lagging indicators are the most valid to reach conclusions on safety performance. Clarity around the definition of a leading indicator is also needed.

The heterogeneity of evidence included in this review may also reflect the complexity of how leading indicators are typically used in practice, where multiple indicators covering a range of safety management functions are selected and tracked together. A key challenge to future research will be to identify methods by which more controlled evaluations of groups of indicators can be performed, and what level of heterogeneity might be acceptable when seeking to compare results across multiple studies.

Considering the differences in practice and decision-making between healthcare and OSH, it may be valuable to consider developing evidence standards that are specific to the OSH context. Such standards would help to ensure that future research in this area is useful and contributing to improvements in leading indicator practice and implementation. Developing the standards through active engagement with a range of stakeholders will also be important to ensure that the standards produce the most appropriate evidence, and to maximise the potential for it to impact on OSH practice.

Conclusion

This review has identified a substantial, though disparate, evidence base evaluating the impact of leading indicators on safety lagging indicators. Almost all studies reported a positive impact, though the degree to which these findings are reliable indicators that leading indicators cause changes to lagging outcomes remains unknown for several reasons.

The overriding characteristic of the evidence base is the heterogeneity of topics evaluated, and the methods used to evaluate them. This research effort provides an optimistic signal to the discipline that leading indicators are being empirically evaluated. Yet the variation may also prevent generalisations to be made regarding the ability of individual (or specific groups of) leading indicators to reliably improve the safety performance across industries.

To gain an understanding of the general utility of leading indicators and which leading indicators are best to deploy in different setting, future studies should improve their approaches to minimising bias, and identify common tools to measure both leading and lagging indicators to facilitate the comparison of studies. Developing an OSH-specific evidence standards framework is likely to guide and assist this process.

Acknowledgements

We thank the Lloyd's Register Foundation for funding this research, and particularly Ruth Frankish for her support throughout. We are also grateful to the experts who have taken the time to share their expertise and knowledge with us, and helped guide and review this work:

- John Dony – National Safety Council, US.
- David Dowe – Lloyd's Register Foundation.
- Nick Fahy – RAND Europe, UK and Belgium.
- Katherine Mendoza – National Safety Council, US.
- Steven Naylor – Health and Safety Executive, UK.
- James Pomeroy – Arup, UK.
- Malcolm Staves – L'Oreal, France.

Abbreviations

AARM-CMS	Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious
AARM-CMSL	Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious, light
ACES	Assessments of contractor safety
AR	Autoregressive
BIM	Building information modelling
BMI	Body mass index
BS	Bachelor of science
CAUL	Council of Australian University Librarians
CDM3	Construction disability management maturity model
CMS	Catastrophic, Major, Serious
CMSL	Catastrophic, Major, Serious, Light
DART	Injuries involving days away, restricted or transferred
DKK	Danish Crone
EAG	Expert advisory group
FCT	Portuguese foundation for science and technology
FTE	Full-time equivalent
HI	High income
HSE	Health, safety and environment
HSSE	Health, safety, security and environment
JHA	Job hazard analysis
LMI	Lower middle income
LTI	Lost time incidents
LTIF	Lost time injury frequency
LTIR	Lost time incident rate
LTSA	Long-term sickness absence
MA	Meta analysis
MS	Master of science
NA	Not applicable
NASA	National aeronautics and space administration
NIOSH	National Institute for Occupational Safety and Health
NOSACQ-50	Nordic occupational safety climate questionnaire
NR	Not reported
NSERC	Natural Sciences and Engineering Research Council of Canada
OIR	Occupational incident rate
OPM-MU	Organisational performance metric - Monash University
OSF	Open Science Framework
OSH	Occupational safety and health
OSHA	Occupational safety and health administration
PCC	Population, concept, context
PPE	Personal protective equipment
PhD	Doctor of philosophy
PICO	Population, intervention, comparators, outcomes
PPE	Personal protective equipment
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
RIR	Recordable injury rate

RJMC	Regional journal management centre,
RoB	Risk of bias
ROP	Research opportunities program
SD	Standard deviation
SH-26	Safety management self-assessment questionnaire
SPE	Society of petroleum engineers
SSHE	Safety, security, health and environment
TIR	Total incident rate
TR	Talonrakennus (residential construction)
TRCFR	Total recordable case frequency rate
TRIFR	Total recordable injury frequency rate
TRincFR	Total recordable incident frequency rate
TRincR	Total recordable incident rate
TRIR	Total recordable injury rate
UK	United Kingdom
UMI	Upper middle income
US	United States
USG	United States gypsum corporation
VAR	Vector autoregression
VEC	Vector error correction
WSIB	Work safety and insurance board
YHEC	York Health Economics Consortium

1 Introduction

1.1 Background

The International Labour Organization estimates that globally 395 million workers sustain non-fatal work injuries and almost three million die due to work related accidents and diseases per year [1]. The performance of safety has traditionally been assessed using lagging indicators which measure the occurrence and incidence of events in the past [2]. Lagging indicators, most commonly including work-related injuries, illnesses and fatalities, show safety has improved in some industries [7]. However, because they measure past occurrence, they are not considered to provide any indication of which factors contributed to improvements in safety.

Leading indicators have more recently been identified as precursors to lagging indicators, including for example conditions, events or measures that precede undesirable events that have some value in predicting or preventing the arrival of the event [3]. Leading indicators are defined as proactive, preventive, and predictive measures that inform how effective the performance of implemented health and safety activities is [2, 4]. Implementation of leading indicators by definition would suggest that the lagging indicators would subsequently decline. However, previous studies examining the interaction of leading and lagging indicators have struggled to find significant relationships and no reviews of this evidence are known to exist [5, 6].

Lloyd's Register Foundation (the Foundation) is supporting research into leading indicators as part of its charitable mission to enhance the safety of workers globally, and across industries. In order to address the absence of reviews in this area of research, the Foundation has commissioned a rapid evidence assessment of the available evidence base for leading indicators. York Health Economics Consortium (YHEC) designed and undertook the review, consulting at all key stages with a group of experts in occupational safety and health (OSH) including a representative from the Foundation, in order to ensure a useful contribution to the field.

1.2 Review Objectives

This scoping review aimed to map the key characteristics of studies that assess the impact of leading indicators on lagging indicators in order to understand:

- For which leading indicators has effectiveness been evaluated?
- For which industries and in which countries does this evidence come from?
- Are different leading indicators evaluated in lower (LI) to middle income (MI) countries, compared to high income (HI) countries?
- How has effectiveness been evaluated?
- Is there any evidence for the effectiveness of leading indicators, and what is the nature of that evidence?
- How robust is this evidence base, and how can it be improved?

2 Methods

The review was conducted according to the methods for scoping reviews, outlined by the Joanna Briggs Institute [8]. The following methods were captured in a review protocol, which was registered on the online Open Science Framework (OSF) in advance of data extraction. Review methods were discussed with the expert advisory group (EAG) prior to beginning the review.

2.1 Eligibility Criteria

To ensure that relevant studies were consistently identified, a clear definition of the eligible study participants, concept and context was developed. These eligibility criteria are summarised in Table 2.1 below and described in detail in Appendix A, Section 1A.1.

As per the JBI guidelines for scoping reviews [8], the PCC (Population, Concept, Context) criteria were used as opposed to the PICO (Population, Intervention, Comparators, Outcomes) criteria usually used in standard systematic reviews. The concept of interest refers to the focus of the scoping review and includes elements from standard systematic review eligibility criteria such as “interventions” and “outcomes”. Context involves such factors and geography, industry and setting.

Table 2.1: Eligibility criteria

	Inclusion Criteria	Exclusion Criteria
Population	Studies of workers of any age, including employees and contractors (including people providing casual labour). Studies including both workers and non-workers will only be eligible if data are reported separately for the workers.	Non-workers e.g., members of the public.
Concept	Studies of any leading indicator reporting any safety lagging indicator. A leading indicator is a factor linked to the practice of safety and health within a workplace, which is intended to prevent future accidents or other adverse safety outcomes, or otherwise improve workplace safety.	None.
Context	Any workplace setting. Studies of participants in a mixture of occupational and non-occupational settings will only be eligible if data are reported separately for workers in occupational settings.	Non-occupational settings.
Study design	Any primary study design measuring outcomes in at least 10 workers.	▪ Case studies of fewer than 10 workers. ▪ Case reports. ▪ Reviews*.
Limits	▪ English language studies only. ▪ Conference abstracts. ▪ For primary studies, inclusion will be limited to papers published in and after 2010.	▪ Primary studies published only as pre-prints. ▪ Primary studies published before 2010. ▪ Editorials or news items. ▪ Non-English language papers.

* The included studies lists of systematic reviews published in the last five years were checked for eligible primary studies (systematic reviews were not eligible for inclusion themselves).

2.2 Identification of Relevant Studies

Searches were conducted to identify studies on the impact of leading indicators on safety outcomes in the workplace. The search strategy design and resource selection reflected the pragmatic scoping review context. The search methods were not designed to be exhaustive. They were designed to target a selection of potentially relevant studies, whilst enabling searches to be conducted and results assessed within the context of the project resource and timeline. The searches were conducted on 1 August 2023. Full details of the search methods are provided in Appendix A, Section 1A.2, and the search strategies are reported in Appendix B.

2.3 Selection of Relevant Studies

Following the removal of obviously irrelevant records by a single reviewer, two reviewers independently assessed the titles and abstracts of 10% of the remaining records, and a single reviewer assessed the remaining records. Subsequently, two reviewers independently assessed 10% of the records at full text review and a single reviewer assessed the remaining records.

The number of records included and excluded at each stage is reported in Appendix C, Figure C.1. Studies excluded after assessment of the full document, with the reasons for exclusion are listed in Appendix D, Table D.1.

2.4 Data Charting

A data extraction template was developed in Excel and piloted by two reviewers on 10% of included studies before progressing to full data extraction.

One researcher extracted data from the remaining included studies, and a second researcher checked all data points for another 10% of included studies.

The elements extracted in the scoping review are reported in Appendix A, Section 1A.4.

2.5 Risk of Bias Assessment

One reviewer appraised the internal validity of each included study, and a second reviewer checked 20% of these assessments. This was performed using a simple tool designed by Robson and coauthors for reviews of effectiveness in OSH published as part of a systematic review of OSH management systems [9]. More information on this tool and how it was used in this review is reported in Appendix A, Section 1A.5.

Risk of bias (RoB) results are reported in Section 3.1.6.1 and detailed RoB assessments for each included study are reported in Appendix E, Table E.1.

2.6 Synthesis and Reporting

Following extraction of information, simple coding of study design characteristics was conducted to enable the breadth of the evidence base to be characterised and summarised. This was performed by a single reviewer.

This scoping review report conforms to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [10] and includes detailed transparent descriptions and justifications of all aspects of the review methods, including the review eligibility criteria, full search strategies, a study flow diagram, a table of studies excluded at full text assessment and detailed tables of the data from eligible studies.

The studies were summarised in tables providing data on their methods and results, alongside a narrative summary exploring the quality of the studies, the relationships between studies and any patterns that we discerned in the data. This was accompanied by an overall assessment of the strength of the research evidence in relation to the research question.

In order to ensure that the report's conclusions were useful to the OSH discipline, results of the review were discussed with the EAG who assisted with identifying and describing implications for practice, and for future research. The report does not necessarily reflect the opinions of Lloyd's Register Foundation.

3 Results

A total of 48 studies (reported in 48 documents) were included in the review and are presented in Appendix F, Table F.1. Results of the literature search and study selection, including PRISMA flow diagram, are reported in Appendix C. A list of the 168 studies excluded at full text is provided in Appendix D, Table D.1.

3.1 Included Study Characteristics

Included studies were published between 2010 and 2023, 35 as full text journal articles, 12 were reported in conference papers [11-22], and one was a report published directly by the Campbell Institute [23]. Articles were published in 28 journals, of which seven published more than one included study: Safety Science reported nine studies [5, 6, 24-30], the Journal of Safety Research reported five studies [31-35], Accident Analysis and Prevention reported three studies [36-38], the Journal of Construction Engineering and Management reported two studies [39, 40] and Resources Police reported two studies [41, 42]. Two conference papers were published in Society of Petroleum Engineers (SPE) [20, 22] and two in SPE Latin American and Caribbean Health / Safety / Environment / Social Responsibility Conference 2013: Sustainable Solutions for Challenging Health, Safety, Security and Environment (HSSE) Environments in Latin America and the Caribbean [14, 18].

3.1.1 Study methods

Approximately one third of included studies (17 / 48) reported the study design used, and for seven of these the review team considered the reported study designs to differ from standard study designs used in healthcare research [5, 24, 29, 43-46]. Author reported study design for the included studies is reported in Appendix G, Table G.1. The reviewers identified five different types of study design used in the included studies which are defined in the context of this review in Table 3.1. Using these definitions, the review team assigned a standardised design to each study based on its description of methods.

Table 3.1: Study design definition

Study design	Definition	Included studies
Cohort study	A comparison between lagging indicators measured in two or more groups that each used different leading indicators (or where one used a leading indicator and one did not), to determine which leading indicator had an impact on lagging indicators. For example, in Van Derlyke 2022, staff from different companies were surveyed about both leading indicators and safety outcomes to evaluate which leading indicators were associated with which lagging indicators [47].	5 [18, 21, 27, 36, 47]
Before-after study	A comparison of lagging indicators before and after a leading indicator was implemented, such as a new programme or guideline (differs from a case series where both the leading and lagging indicators are measured at the different time points). For example, Choe 2016 calculated trends in safety data for five years before and five years after the revision of the OSHA steel erection standard, and compared these to identify whether changes in trends were due to the leading indicator [24].	12 [12-14, 16, 17, 19, 20, 22, 24, 42, 48, 49]
Case series	Follows multiple companies using the same leading indicators and evaluating leading and lagging indicators over time. For example, Laitinen 2013 collected data from all eligible companies in the region and evaluated the correlation between audit scores and accident rates [30].	6 [28-30, 40, 45, 50]
Cross-sectional study	Measurement of **both** leading and lagging indicators is performed at one timepoint, to assess the relationship between the two at one point in time (rather than examining trends over time). For example, Chen 2017 administered a questionnaire to eligible workers to collect data on leading and lagging indicators simultaneously, and then evaluated the correlation between them [51].	17 [11, 25, 26, 31-33, 35, 37-39, 43, 44, 46, 52-55]
Case study	A study of a single company evaluating leading and lagging indicators over time. For example, Winge 2019 used company data from a large construction company to evaluate the correlation between leading and lagging indicators [34].	8 [5, 6, 15, 23, 34, 41, 51, 56]

Abbreviations: OSHA - Occupational safety and health administration.

Only one study identified itself as a cohort study and was described by the authors as national cohort surveys [27]. The remaining four cohort studies did not report the design used [18, 21, 36, 47].

No study described itself as a before-after study, although Choe 2016 was described by the authors as an interrupted time series [24], which is a common method of analysis employed in before-after designs. Two of the studies that the reviewers classified as before-after studies were described by the authors as case studies [48, 49]. Although both included data from only one company, the reviewers considered them before-after studies because the leading indicators were something that was implemented at a specific timepoint, rather than measured on an ongoing basis, with lagging indicators compared before and afterwards. The remaining nine before-after studies did not report a study design [12-14, 16, 17, 19, 20, 22, 42].

Two of the six case series did not report the design used [40, 50]. Amir-Heidari 2017 was defined by the authors as a framework development and case study but was reclassified in our review as a case series as multiple companies were included and evaluated together [28]. Laitinen 2013 was described by authors as “cross-sectional study design, even though the study covered a 3-year period” which the reviewers classified as a case series [30]. The other two studies that the reviewers classified as case series were reported by the authors as a “real-life long-term evaluation” [29] and “descriptive-analytic applied research” [45].

Twelve of the seventeen cross-sectional studies did not report their design, two were described by authors as cross-sectional studies [52, 53], one was described as a survey [43], one as a "deductive quantitative study" [44] and one as a "cross-cultural validation study" [46].

Of the eight case studies, three were reported as case studies [6, 41, 56], one was described as a retrospective analysis [5], and four did not report their study design [15, 23, 51].

The forty studies reporting information on study dates collected data between 1990 [12, 29] and 2022 [41] and for periods of between two months [11] and 16 years [29], with only six studies collecting data for less than 12 months [11, 23, 25, 42, 44, 46].

Twenty-four studies did not report how the study was funded [12-14, 17-20, 22-26, 28-30, 37, 38, 40, 41, 44, 49, 51, 54, 56]. The remaining twenty-four studies received funding from a variety of research and occupational organisations, as reported in Appendix G, Table G.1.

3.1.2 Population and setting

3.1.2.1 Industry

Studies were most commonly undertaken in construction industries (n=18 studies) [5, 6, 15, 16, 21, 24, 29, 32, 34, 38, 39, 46, 48, 51-53, 55, 56], followed by energy (oil and gas but also energy transportation, n=13) [12-14, 18-20, 22, 25, 26, 28, 35, 40, 54] and mining (n=3) [41, 42, 49]. Five industries were each represented by one study: the automotive industry [45], dairy product manufacturing [47], logging [43], maritime [50] and newspaper publishing [17]. An additional eight studies included participants from multiple (where reported, between three and ten) industries [11, 27, 30, 31, 33, 36, 37, 44], and the final study did not report the industries involved (a paper containing five case studies each based at a different company) [23].

3.1.2.2 Population: companies and workers

Eighteen studies evaluated a single company [5, 6, 12-17, 19, 20, 22, 34, 39, 41, 48, 49, 51, 56] and twelve did not report the number of companies included [18, 21, 24, 27, 29, 31, 32, 37, 42, 43, 45, 46]. The rest included between two [54] and 1,180 companies [38] or up to 2,148 individual contractors [53]. Details on individual workers such as ethnicity and gender, were scarcely reported and the characteristics that were reported varied and were reported to varying levels of detail. Reported worker characteristics are summarised in Appendix G, Table G.3.

While the companies involved and their staff were almost always stakeholders, in some cases managers or supervisors specifically were also stakeholders, such as in Breitsprecher 2019 in which the leading indicator was training for managers [22]. A range of other stakeholders were identified including developers of the DrivingChange programme [13], OSH experts [45], professional organisations such as the National Institute for Occupational Safety and Health (NIOSH) [31] and insurance companies such as the Ohio Bureau of Workers' Compensation [31].

Studies within the construction industry (n=18) included between one [5, 6, 15, 16, 34, 39, 48, 51, 56] and 1,180 companies [38] or 2,148 contractors [53]. Those that reported the setting of the study (n=10) reported it as construction or building sites [5, 16, 29, 32, 34, 38, 46, 48, 52, 56] while the other eight did not report whether they evaluated staff involved in construction sites alone or also those from other industry settings, such as office or transport staff [6, 15, 21, 24, 39, 51, 53, 55].

Each construction industry study evaluated data from between eight and 1,180 sites [15, 29, 32, 38, 46, 48, 52], although this detail was not reported by 11 studies [5, 6, 16, 21, 24, 34, 39, 51, 53, 55, 56]. Zahoor 2017 included data from 40 sites across five cities (in Karachi, Lahore, Islamabad/Rawalpindi, Faisalabad and Hyderabad [46]). Dennerlein 2020 reported that 64 surveys were completed by subcontractors across 24 sites from 43 unique companies from a variety of trades within construction [52]. Seven studies provided information on company size, four describing them as mid-sized [6] or large [15, 29, 34] but without providing further definition; two evaluated companies of a range in sizes [32, 38] and one reported average worksite size in square feet, full time equivalent (FTE) staff and staff costs [52].

The 13 studies in the energy industry included one [12-14, 19, 20, 22], two [54] or three [28, 35] companies, or 261 contractors [40], or did not report the number of companies included [18, 25, 26]. Ten did not report the setting of the study [12, 13, 18-20, 22, 25, 26, 28, 40] and the remaining three were set in driving [14], on ships [54] and on oil platforms [35]. Four studies reported the number of sites included (nine [20] or ten sites [35], 55 facilities [25] and 59 tankers [54]) but the remaining nine did not report number of sites [12-14, 18, 19, 22, 26, 28, 40]. One study described the included organisation as large but did not define this [54] and the remaining 12 did not report company size.

Of the three mining studies, one included one company but did not report number of sites included [41], a second included nine mines but did not report how many companies those represented [42] and the third included one mine from one company [49]. One was set in diamond mines [41], one in surface and underground mines [49] and one did not report the setting [42]. None reported company sizes.

The five studies evaluating different industries also varied widely in reported worker and company characteristics. The study in the automotive industry included 11 experts but did not report the number of companies that they represented (or the companies' size or number of sites) or the setting [45]. The study in the dairy product manufacturing industry included 82 companies (five small [11 to 19 employees], 13 medium [20 to 99 employees] and 64 large [100 or more employees]), but the setting and the number of sites were not reported [47]. The study in the logging industry reported neither the number of companies or sites, nor their size or setting [43]. The study in marine transport included 102 vessels from three companies in both the shoreside and shipboard setting and measured size by the weight of the ship (ten vessels were less than 25,000 deadweight tonnage, 34 vessels were between 25,000 and 50,000 deadweight tonnage, and 38 vessels were more than 50,000 deadweight tonnage) [50]. The study in the newspaper publishing industry evaluated one company with more than 2,200 employees but reported no further details on study setting or size [17].

The eight studies conducted in multiple industries evaluated between 23 [36] and 1,240 companies [33] although three did not report this detail [27, 31, 37]. Company size was defined by the number of employees [30, 31, 37], using definitions based on European and American standards [44], or was not defined [11]. Two studies did not report the size of the companies [27, 36], and none of the eight studies in multiple industries reported the study settings or number of sites. The study that did not report the industry included five companies (and named them) but did not report further population or setting characteristics [23].

3.1.2.3 Countries

Twenty-seven studies were conducted in HI countries:

- Nine in the United States (US) [13, 17, 24, 31, 39, 43, 49, 52, 56], six in Canada [6 , 15, 16, 32, 33, 55], and two in both [21, 53].
- Two were conducted in Australia [5, 37].
- Two in Finland [29, 30].
- One each in Chile [38], Denmark [27], Norway [34], Portugal [11], Qatar [12] and Singapore [51].

Four studies took place in UMI countries: one study included companies in South Africa, Botswana and Namibia [42], one study in just Namibia [41] and two in Malaysia [26, 35].

Four studies were performed in Iran [28, 45, 48] or Pakistan [46], classified as LMI countries by the World Bank 2022-2023 classification [57].

Of the remaining 13 studies, three included the US (HI) as well as other countries that were not reported [23, 36, 40]. One study included eight countries ranging from LMI to HI [25]. Four were international studies but did not list the included countries [18, 20, 47, 54]. Three reported the continents but not countries (Africa [22], Middle East [14] and Europe, Asia, Australia, South America, North America and Africa [44]). Two studies did not report the country at all [19, 50].

Table 3.2: Study population and setting

Study	Setting	Location of studies (income level classification)	Number of sites	N of companies	Company or worksite size	N of workers
Construction						
Alarcón 2016 [38]	Construction companies	Chile (HI)	1,180	1,180	Micro (≤9 employees), small (10 to 49), medium (50 to 199), and large companies (≥200 or more). Number of each NR.	NR
Cao 2019 [51]	NR	Singapore (HI)	NR	1	NR	NR
Chen 2017 [32]	Construction sites	Canada (HI)	112	NR	Micro (1 to 4 employees)=5.1% Small (5 to 99)=55.7% Medium (100 to 499)=25.7% Large (500+)=13.5%	783
Choe 2016 [24]	NR	US (HI)	NR	NR	NR	NR
Dadashi Haji 2023 [48]	Construction site	Iran (LMI)	35	1	NR	NR
Dennerlein 2020 [52]	Building sites	US (HI)	24	43	Average worksite was 245,850 (SD 358,790) square feet; involved 116 FTE staff (SD 124), and cost $116.80 million (SD $215.82 million).	1,426
Hinze 2013 [39]	NR	US (HI)	57 projects	NR	NR	NR
Laitinen 2010 [29]	Building sites	Finland (HI)	426 (in 2006) sites in the region, of which 310 (73%) participated in the contest.	NR (80% of construction companies in the region).	"Large companies" (not defined by the authors).	NR
Lingard 2017 [5]	Building site	Australia (HI)	NR	1	NR	NR

Study	Setting	Location of studies (income level classification)	Number of sites	N of companies	Company or worksite size	N of workers
Manjourides 2019 [53]	NR	US and Canada (both HI)	NR	2,148	NR	NR
Mohammed 2019 [15]	NR	Canada (HI)	8 construction projects	1	Described as large company, definition NR	NR
Pereira 2017 [16]	Building sites	Canada (HI)	8 projects	1	NR	NR
Quaigrain 2023 [55]	NR	Canada (HI)	NR	8	NR	NR
Rajendran 2013 [56]	Construction sites	US (HI)	NR	1	NR	NR
Versteeg 2019 [6]	NR	Canada (HI)	48 projects	1	Company described as 'mid-sized', but no numbers given.	NR
Wei 2020 [21]	NR	Canada and the US (both HI)	NR	NR	NR	587
Winge 2019 [34]	Construction sites	Norway (HI)	12 projects	1	Described as one of Norway's biggest construction clients but size not reported.	NR
Zahoor 2017 [46]	40 under-construction multi-storey building projects in Pakistan (at least 70 metres high).	Pakistan (LMI)	40 Karachi: 28 Lahore: 7 Islamabad/ Rawalpindi: 3 Faisalabad: 1 Hyderabad: 1	NR	NR	426
Energy						
Amir-Heidari 2017 [28]	Three oil and gas drilling companies.	Iran (LMI)	NR	3	NR	NR
Bitar 2018 [25]	An upstream organisation with an oil and gas company.	55 facilities in 8 countries: Angola (LMI) Azerbaijan (UMI) Georgia (UMI) Indonesia (LMI) Norway (HI) Trinidad (HI) UK (HI) US (HI)	55 facilities	NR	NR	3,514

Study	Setting	Location of studies (income level classification)	Number of sites	N of companies	Company or worksite size	N of workers
Breitsprecher 2014 [22]	NR	NR ("Africa region")	NR	1	NR	NR
Doherty 2010 [12]	NR	Qatar (HI)	NR	1	NR	NR (reported by authors that at its peak the company had >35,000 workers onsite).
Gale 2011 [13]	NR	US (HI) HI	NR	1	NR	NR
López 2013 [14]	Driving	NR ("Middle East")	NR	1	NR	NR
Merrick 2014 [54]	Shipboard	Multinational (countries NR)	59 tankers	NR (2 shipping fleets)	One organisation described as large (the authors do not define 'large'). The other was described as global.	915
Salas 2016 [40]	NR	Both the US and internationally (US – HI)	NR	261 contractors	NR	261
Stough 2012 [18]	NR	Global (NR)	NR	NR	NR	NR
Tang 2017 [26]	NR	Malaysia (UMI)	NR	172	NR	172
Tang 2018 [35]	Platforms	Malaysia (UMI)	10	3	NR	NR
Tauseef 2012 [19]	NR	NR	NR	1	NR	NR
Thananan 2014 [20]	NR	International (NR)	9	1	NR	NR
Mining						
Coetzee 2023 [41]	Mines	Namibia (UMI)	NR	1	NR	>900
Govender 2022 [42]	NR	South Africa, Botswana and Namibia (all UMI)	9 mines	NR	NR	NR
Haas 2018 [49]	Mines (surface and underground).	US (HI)	1 mine	1	NR	>450
Automotive						
Vosoughi 2021 [45]	NR	Iran (LMI)	NR	NR	Reported to be conducted in the largest automobile companies in Iran.	11

Study	Setting	Location of studies (income level classification)	Number of sites	N of companies	Company or worksite size	N of workers
Dairy product manufacturing						
Van Derlyke 2022 [47]	NR	International (NR)	NR	82	Small (11 to 19 employees) (n=5 participants from small companies), medium (20 to 99) (n=13), large (>100) (n=64)	NR
Logging						
Lagerstrom 2019 [43]	NR	US (HI)	NR	NR	NR	743
Marine transport						
Grabowski 2010 [50]	Shoreside and shipboard	NR	102 vessels	3	Ten vessels were less than 25,000 deadweight tonnage, 34 vessels were between 25,000 and 50,000 deadweight tonnage, and 38 vessels were more than 50 000 deadweight tonnage.	Shipboard n=1,599 Shoreside n=157
Newspaper publishing						
Schiavi 2013 [17]	NR	US (HI)	NR	1	>2,200 employees. Categorized as medium manufacturing facilities. Circulation is 250,000 daily and 350,000 on Sunday.	>2,200
Multiple industries						
Brandt 2023 [27]	NR	Denmark (HI)	NR	NR	NR	63,500
Laitinen 2013 [30]	NR	Finland (HI)	NR	126	Average company size: 186 employees (range: 4 to 993 employees).	23,399
Moore 2022 [31]	NR	US (HI)	NR	NR	Employer size (number, %): 1 to 10 (265, 4%) 11 to 49 (3,030, 48%) 50 to 99 (1,700, 27%) 100 to 249 (1,163, 18%) 250 to 999 (197, 3%) >1,000 (7, <1%)	2,295
Mousavi 2020 [44]	NR	Europe, Asia, Australia, South America, North America, Africa (NR)	NR	112	Small and medium (n=57), large (n=55). Definitions of sizes based on European	112

Study	Setting	Location of studies (income level classification)	Number of sites	N of companies	Company or worksite size	N of workers
					and American standards but specifics NR.	
Robson 2017 [33]	NR	Canada (HI)	NR	1,240	NR (all >20 employees)	NR
Sá 2023 [11]	NR	Portugal (HI)	NR	59	"The vast majority are medium and large sized" - sizes not defined	NR
Sheehan 2016 [37]	NR	Australia (HI)	66	NR	Workplace (site) size ranged from 4 to 532 employees with a mean size of 54 employees.	3,578
Wachter 2014 [36]	NR	US and multinational (US – HI)	NR	Study 1: 330 Study 2: 23	NR	Study 1: mean number of employees per establishment was 632 Study 2: 650
Industry not reported						
Campbell Institute 2015 [23]	NR	Mixed (not all reported but included US [HI])	NR	5	NR	NR

Abbreviations: CDM3 – construction disability management maturity model, FTE – Full time equivalent, HI – High income, LMI – Lower middle income, NASA - National aeronautics and space administration, NR – Not reported, SD – Standard deviation, UMI – Upper middle income, US – United States, USG – United States gypsum corporation.

3.1.3 Leading indicators

Following categorisation, a total of five different types of leading indicator were identified as having been evaluated in the included studies. Table 3.3 summarises the leading and lagging indicators in each study, as well as the relationship being assessed in the study, and more detailed descriptions of the leading indicators and the way they were measured are reported in Appendix G, Table G.2.

3.1.3.1 New guideline, tool or process

The most common leading indicator category was the implementation of a new guideline, tool or process, evaluated in 15 studies including all 12 before-after studies [12-14, 16, 17, 19, 20, 22, 24, 42, 48, 49], one cross-sectional study [11], a case series [29] and a case study [41]. Leading indicators were measured using company records [41], publicly available data [29] or by questionnaire capturing the self-reported use of the leading indicator [11]. In the 12 before-after studies, leading indicators were not measured, rather the lagging indicators were measured before and after the leading indicator was implemented.

These leading indicators varied considerably in terms of which guideline, tool or process was being evaluated; the majority of these studies (8 / 15) [12, 13, 16, 19, 20, 41, 42, 49] set out to assess the impact of introducing a new safety programme, such as the '4Cs framework' in the two diamond mining studies [41, 42]. Others evaluated the impact of a national guideline, [24], a regional contest of safety behaviours and performance [29], or the use of 'lean tools' (not further described) [11]. In the remaining studies the leading indicator appeared to be specific to the company such as the RasGas Elements of Excellence [12], and evaluating its effect on safety outcomes.

Studies evaluating a new guideline, tool or process varied in the extent to which these leading indicators met the definitions posed by Xu et al 2021 or the Campbell Institute. Seven studies measured safety performance while proactively improving it and so met the Xu et al 2021 definition [12-14, 16, 17, 20, 29]. The rest were considered to partially meet the definition because they were either a proactive safety intervention that did not include the measurement of current performance [11, 22, 24, 41, 42] or it was a new way to measure safety performance but not in itself a proactive action to improve safety [19, 48, 49]. Ten studies met the Campbell Institute definition of leading indicators and five only partially met the definition due to the indicators not including the monitoring of current safety performance [11, 22, 24, 41, 42].

Most studies (n=8) were conducted in HI countries [11-13, 16, 17, 24, 29, 49], with a minority in UMI (n=2) [41, 42] or LMI (n=1) countries [48]. Four did not report the country or countries in which they took place [14, 19, 20, 22].

3.1.3.2 Safety climate/culture

Five studies evaluated safety culture or climate, all using a survey or questionnaire to measure this leading indicator [21, 27, 32, 43, 46]. Two used the Nordic Occupational Safety Climate Questionnaire (NOSACQ-50), including a cross-sectional study in logging [43] and a cohort study comparing companies with different numbers of safety climate problems in multiple industries [27]. Three studies (two cross-sectional [32, 46] and one cohort [21]) used questionnaires adapted from previous studies and modified by the study authors.

An assessment of safety culture partially meets the Xu et al definition of a leading indicator because it is an assessment of current safety that can identify the strengths, weaknesses and risks of a system, but it is not a proactive action to correct risks. Safety culture meets the Campbell Institute definition of a leading indicator because it provides information about current safety performance.

Four of these studies were conducted in HI countries [21, 27, 32, 43] and one in a LMI country [46].

3.1.3.3 Audits and inspections

Two studies evaluated the use of audits and inspections as leading indicators, a case series [30] and a cross-sectional study [33]. The case series included multiple industries in Finland (a HI country) and looked the Elmeri+ method which was developed as an easy and simple tool to measure occupational health and safety [30]. In the cross-sectional study, firms were audited as part of the WorkWell programme in Ontario, Canada (also a HI country) [33].

Although not in themselves a corrective action, audits and inspections do identify the strengths, weaknesses and risks in a workplace and therefore meet the Campbell Institute’s definition of a leading indicator and partially meet Xu et al’s definition.

3.1.3.4 Monitoring of safety

One cross-sectional study in multiple industries in the US (HI) used the Safety Management Self-Assessment Questionnaire (SH-26) by the Ohio Bureau of Worker's Compensation to assess safety hazards and management practices, which was also judged to meet the Campbell Institute’s definition of a leading indicator and to partially meet Xu et al’s definition.

3.1.3.5 Mixed

Twenty-five studies reported multiple and varied leading indicators, either assessing each one individually or all together [5, 6, 15, 18, 23, 25, 26, 28, 34-40, 44, 45, 47, 50-56].

Alarcón 2016, a cross-sectional study in the construction industry in Chile (a HI country), assessed different prevention management practices, alone or in combination with each other [38]. Amir-Heidari 2017, a case series of three oil and gas drilling companies in the US (HI), evaluated a composite score of twelve leading indicators [28]. Dennerlein 2020, a cross-sectional study in the construction industry in the US (HI), used a survey designed by the study authors to assess safety performance through different leading indicators [52]. Stough 2012, an international cohort study in the energy industry, compared lagging indicators in companies with different numbers of leading indicators [18]. Van Derlyke 2022, an international cohort study in dairy product manufacturing, compared outcomes between companies that did and did not implement the eight leading indicators [47].

The remaining 20 studies evaluated multiple leading indicators and reported the results for each one separately. They include one cohort study [36], three case series [40, 45, 50], nine cross sectional studies [25, 26, 35, 37, 39, 44, 53-55] and seven case studies [5, 6, 15, 23, 34, 51, 56]. Each of these 20 studies reported data for between two and 38 leading indicators.

Eleven of these twenty studies were set in HI countries [5, 6, 15, 23, 34, 37, 39, 51, 53, 55, 56]. Two took place in Malaysia, an UMI country [26, 35] and one took place in Iran, a LMI country [45]. One study included eight countries: four HI (Norway, Trinidad, United Kingdom (UK) and US), two UMI (Azerbaijan and Georgia) and two LMI (Angola and Indonesia) [25]. Four studies were multinational and did not report the included countries [36, 40, 44, 54], although two did report that they included the US (HI) [36, 40]. One study did not report where it was set [50].

More studies fully met the Campbell definition of a leading indicator than the Xu definition (12 and seven respectively), with the remaining studies partly meeting both definitions.

Table 3.3: Leading and lagging indicators

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
New guideline, tool or process				
Breitsprecher 2014 [22] Xu et al 2021 Campbell Institute	Before-after study	HSE Leadership Academies.	TRincR and motor vehicle accident rate.	Rates of lagging indicators before and after the leadership academies for senior and middle managers.
Choe 2016 [24] Xu et al 2021 Campbell Institute	Before-after study	Revision of the Steel Erection Standard.	Fatality rate, days away injury rate, normalised fatality rate and normalised days away injury rate.	Trends of safety data for 5 years before and 5 years after the revision of the steel erection standard.
Coetzee 2023 [41] Xu et al 2021 Campbell Institute	Case study	Integrated 4C framework.	Loss of life, all injury frequency rate, total recordable case frequency rate, lost time injury frequency rate, MPI and HPI, MPH and HPH, lost time injuries medical treatment cases, first aid cases, days lost.	The relationship between the framework and safety performance indicators over time.
Dadashi Haji 2023 [48] Xu et al 2021 Campbell Institute	Before-after study	A tool integrating BIM and knowledge base.	Accidents, injuries and fatalities.	The rates of accidents, injuries and fatalities before and after the framework was implemented.
Doherty 2010 [12] Xu et al 2021 Campbell Institute	Before-after study	RasGas Elements of Excellence.	TRincR and also heat injury rate during Ramadan.	The effect that the health and safety management tools had on incident rates at the company.
Gale 2011 [13] Xu et al 2021 Campbell Institute	Before-after study	DrivingChange program.	Motor vehicle incidents and field worker safety.	Relationships between leading and lagging indicators over time after DrivingChange was implemented.
Govender 2022 [42] Xu et al 2021 Campbell Institute	Before-after study	Integrated 4C safety framework.	Loss of life, total recordable case frequency rate, lost time injury frequency rate, lost time injury severity rate, all injury frequency rate, high potential incident.	Impact of the Integrated 4Cs framework evaluated through lagging indicator trends over time.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
Haas 2018 [49] Xu et al 2021 Campbell Institute	Before-after study	Field-level risk assessment program.	Incidents.	Days lost scores over time but no statistical relationship is reported.
Laitinen 2010 [29] Xu et al 2021 Campbell Institute	Case series	Implementation of a contest based on the standardised TR-observation method.	Accidents per cubic meter of construction.	Accidents over time before and since the implementation of the contest. The index and subindex scores are also compared.
López 2013 [14] Xu et al 2021 Campbell Institute	Before-after study	Implementation of a RJMC.	Automotive Accidents CMS (industry recognised) AARM-CMS (industry recognised) Total rollover (company and contractors) Automotive Accidents CMSL (industry recognised) AARM-CMSL (industry recognised).	The number of accidents before and after the first RJMC was implemented.
Pereira, 2017 [16] Xu et al 2021 Campbell Institute	Before-after study	Behaviour-based safety program.	TRincR and TIR.	TRincR and TIR before and after the implementation of the program. Correlations between specific report types and incident rates.
Sá 2023 [11] Xu et al 2021 Campbell Institute	Cross-sectional study	Implementation of Lean tools.	Accident rates.	Accident rates following the implementation of Lean tools.
Schiavi 2013 [17] Xu et al 2021 Campbell Institute	Before-after study	"Resurrected safety process".	Percentage reduction in total work-related accidents, lost time accidents and musculoskeletal diseases.	Change in numbers of accidents and injuries over time.
Tauseef 2012 [19] Xu et al 2021 Campbell Institute	Before-after study	Observation intervention program.	TRIFR and automotive accident rate.	Trends in TRIFR and accident rates over time.
Thananan 2014 [20] Xu et al 2021 Campbell Institute	Before-after study	Step Change roadmap.	LTIF, TRIR and process safety event.	SSHE performance after the implementation of the Step Change program.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
Safety climate/culture				
Brandt 2023 [27] Xu et al 2021 Campbell Institute	Cohort study	Safety climate (NOSACQ-50 questionnaire).	LTSA.	Hazard ratios of LTSA for different numbers of safety climate problems.
Chen 2017 [32] Xu et al 2021 Campbell Institute	Cross-sectional study	Safety climate and individual resilience.	Physical safety outcomes.	Six hypothesized safety climate factors in relation to physical safety outcomes.
Lagerstrom 2019 [43] Xu et al 2021 Campbell Institute	Cross-sectional study	Safety climate (NOSACQ-50 survey).	Presence of musculoskeletal symptoms or missed work due to musculoskeletal symptoms.	Safety climate (leading indicator) and musculoskeletal symptoms (lagging indicator).
Wei 2020 [21] Xu et al 2021 Campbell Institute	Cohort study	Safety culture factors, as well as number of projects in previous 3 years and work hours per week.	Physical injuries and unsafe events.	Relationship between safety culture and safety performance.
Zahoor 2017 [46] Xu et al 2021 Campbell Institute	Cross-sectional study	Safety climate. Also 2 of the 3 elements of safety performance.	Safety performance included three broad indicators of which one was eligible for this review: number of self-reported accidents/injuries and near-misses in past 12 months.	The relationships assessed that are relevant to this review are the correlations between safety climate and accidents/injuries and near misses.
Audits and inspections				
Laitinen 2013 [30] Xu et al 2021 Campbell Institute	Case series	Elmeri+ score.	Accidents/10^6 working hours, blue collar accidents/10^6 working hours and lost hours/blue collar worker.	Correlation between Elmeri+ score and accident rates, with subgroup analyses looking at the same relationships in specific industries as well as the correlation between sub index scores and accidents.
Robson 2017 [33] Xu et al 2021 Campbell Institute	Cross-sectional study	Audit score.	Lost time claims and non-lost time claims.	Association between audit scores and claims.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
Monitoring of safety				
Moore 2022 [31] Xu et al 2021 Campbell Institute	Cross-sectional study	Safety hazards and management practices.	Worker's compensation claims.	Relationship between the SH-26 assessment ratings and worker's compensation claims.
Mixed				
Alarcón 2016 [38]	Cross-sectional study	221 practices were identified, grouped into 7 categories: Accidents & incidents investigation Safety planning & resources Management commitment Workers' safety training Management safety training Audits & certifications Safety incentives & rewards.	Accident rate.	The relationship between the different prevention management strategies and accident rate in different companies.
Amir-Heidari 2017 [28]	Case series	A composite score of 12 leading indicators.	A similar composite score including: rate of occupational accidents which lead to fatality, rate of recordable occupational accidents, rate of occupational illness/ health problem reports and amount of legal fines/ costs related to HSE.	Composite scores for leading and lagging indicators over 5 years are reported for each company, but no statistical tests were done to assess the significance or strength of the relationships.
Bitar 2018 [25]	Cross-sectional study	Operating discipline communication, operating discipline implementation, leadership expectations and trust index.	Recorded injury frequency and near miss frequency (both Control of Work related and not).	Relationship between operating discipline communication, operating discipline implementation, leadership expectations and trust index and HSE outcomes including recorded injury frequency and near miss frequency.
Campbell Institute 2015 [23]	Case study	Training hours, safety observations, incident investigation, site audits, leadership engagement.	Incident rates, accident rates, hazard rates, near miss reporting and stop-work authority.	Correlations between leading indicators and incident rates.
Cao 2019 [51]	Case study	Project delay, project man hours, PPE, overhead protection, excavation work, machine safety guarding, safe means of access, operating crane/lifting, scaffold, tower/mobile scaffold, mech elevated work platform, falling hazard/opening, electrical hazard, first aid facilities, emergency preparedness, handling and storage of hazardous materials, safe work procedures, power tool safety, earth control measures, noise/vector and others.	Accident rates.	Time series analyses to identify how well health and safety management is being implemented by assessing relationships between leading indicators and accidents over different time periods to establish if they can be considered predictors.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
Dennerlein 2020 [52]	Cross-sectional study	ACES	OSHA reportable injuries, injuries resulting in DART.	Correlation between worksite level ACES scores and safety climate scores. Regression coefficients between worksite ACES score and OSHA recordable injuries (per 100 FTEs) and DART injuries (per 100 FTEs). Regression coefficients between subcontractor ACES score and OSHA recordable injuries (per 100FTEs) and DART injuries (per 100FTEs).
Grabowski 2010 [50]	Case series	Safety factors.	Safety performance.	Correlations between safety factors and safety performance.
Hinze 2013 [39]	Cross-sectional study	Implementation of various safety practices (e.g., health and safety manual, safety prequalification, subcontractors' participation in general contractors' orientation and training).	Recordable injury rate.	Correlations between safety practices (presence of certain safety practices as well as the number of safety practices) and RIR.
Lingard 2017 [5]	Case study	Management activity (toolbox meetings, pre-brief meetings/ pre-start meetings, safety observations, site surveillance inspections carried out, penalties/ infringements, occupational health and safety audits, non-compliances, hazards reported, hazards closed out, statutory authority inspections carried out, alcohol tests, drug tests, safe work method statements/ JSA documents review and amended, site inductions).	TRincFR.	Assessing temporal relationships between leading and lagging indicators and the causal relationships between them.
Manjourides 2019 [53]	Cross-sectional study	Safety Management System, Safety Program Elements Hazards, Safety Program Elements Programs, Special Elements Non-drug and alcohol, Special Elements Drug & Alcohol Screening and OSHA Citations.	Recordable injury cases and DART.	Correlation between leading (Safety Management, Safety Programs [e.g. falls, hearing protection], and Special Elements [drug testing, return to work]) and lagging indicators (recordable injury cases and days away, restricted, or transferred, both calculated per 100 full-time equivalent person-hours billed).
Merrick 2014 [54]	Cross-sectional study	Organisational and crew member decision frames.	Safety performance.	Interactions between different decision objectives and their individual contribution to the model to look at the relationship between the leading and lagging indicators.
Mohammed 2019 [15]	Case study	Project performance safety and scheduling data	Accidents.	Correlations between project data and safety performance data.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
Mousavi 2020 [44]	Cross-sectional study	Lean maturity.	OHS performance.	Modelling carried out to establish the significance of the different indicators in being able to predict OHS performance. Combined effects of variables on OHS performance is also assessed.
Quaigrain 2023 [55]	Cross-sectional study	CDM3.	Recordable injury rate, lost time case rate and severity rate.	Relationship between maturity of disability management indicators and safety performance.
Rajendran 2013 [56]	Case study	Worker safe behaviour observation, pre task plan and site safety audits.	Near miss incident rates, first aid injury rates, OSHA recordable injury rates, and total incident/injury rates.	Correlation between the three leading indicators and the four lagging indicators.
Salas 2016 [40]	Case series	Contractor safety data.	TRIincR and Severity rate.	The robust ordinary least-square stepwise multiple regression analysis to identify factors with predictive power and to determine the most predictive model.
Sheehan 2016 [37]	Cross-sectional study	Aggregate OPM-MU score and safety leadership aggregate score.	Reported occupational health and safety incidents, unreported occupational health and safety incidents, and near misses.	The association between leading and lagging indicators in occupational health and safety and the moderating effect of safety leadership on that relationship.
Stough 2012 [18]	Cohort study	Operating assets and proactive activities.	Total recordable injury rates and severity-weights total recordable injury rates, incidents, near misses, investigations and corrective actions.	Relationship between proactive safety activities and safety outcomes.
Tang 2017 [26]	Cross-sectional study	Safety performance (inspection and maintenance, management and work management on safety, number of incidents and near misses, personal safety, constrictor's safety, management of plant changes, plant operation and operating procedures, competence, plant design, instrumentation and alarms, hazard identification and risk assessment, documentation, start-ups and shutdown, emergency management).	Incident occurrence.	Relationships between safety factors and incidence occurrence (Z values, R^2 values, p values).
Tang 2018 [35]	Cross-sectional study	Safety performance (inspection and maintenance, emergency management, management and work management, number of incidents and near misses, personal safety, constrictor's safety, management of change, operation and operating procedures, competence, hazard identification and risk	Fatality, fatal accident rates, TRincR, LTIR and reported near-misses.	Correlation between the scores of safety factors and the actual performance of the offshore oil and gas platforms.

Study	Study design	Leading indicator(s)	Lagging indicator(s)	Relationships assessed
		assessment, plant design, instrumentation and alarm, start-ups and shutdown).		
Van Derlyke 2022 [47]	Cohort study	Safety audits, preventative maintenance, safety training attendance, safety observations, safety inspections, near-miss reporting, stop work authority, JHA/JSA, safety meeting attendance, corrective action completion rate, worker perception survey, and attendance tracking.	OSHA incident rates and DART rates.	Agreement among respondents on the most important leading indicators and the performance of OSHA incident rates and DART rates.
Versteeg 2019 [6]	Case study	Toolbox talks, number of site inspections and number of near misses.	Number of medical injuries, number of first aid injuries.	The relationship between leading indicators and injuries.
Vosoughi 2021 [45]	Case series	Safety indicators and educational indicators.	Total number of work-related lost time injuries, frequency severity index, percentage of total number of work-related lost time injuries reduction compared to previous year.	Identification of the most 'important' leading indicators and their relation to work related lost time injuries.
Wachter 2014 [36]	Cohort study	Safety management practices.	Accident rates.	Between safety management practices and accident rates.
Winge 2019 [34]	Case study	Safety practices.	Safety performance.	Projects were compared and connections between leading indicators and safety performance were identified.

Abbreviations: AARM-CMS - Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious, AARM-CMSL – Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious, light, ACES – Assessments of contractor safety, BIM – Building information modelling, CDM3 – Construction disability management maturity model, CMS – Catastrophic, Major, Serious, CMSL – Catastrophic, Major, Serious, Light, DART – Injuries involving days away, restricted or transferred, HSE – Health, safety and environment, JHA – Job hazard analysis, JSA – Job safety analysis, LTIF – Lost time injury frequency, LTIR – Lost time incident rate, LTSA – Long-term sickness absence, NOSACQ-50 – Nordic occupational safety climate questionnaire, OPM-MU – Organisational performance metric - Monash University, OSHA – Occupational safety and health administration, PPE – Personal protective equipment, RIR – Recordable injury rate, RJMC – Regional journey management centre, SH-26 – Safety management self-assessment questionnaire, SSHE - Safety, security, health and environment, TIR – Total incident rate, TR – Talonrakennus (Residential construction), TRIFR – Total recordable injury frequency rate, TRIR – Total recordable incident rate TRincFR - Total recordable incident frequency rate, TRincR – Total recordable incident rate.
If leading indicators met the Xu et al 2021 or Campbell Institute definition, this is highlighted in green in the study column. If they partially met the definition, it is highlighted in yellow. This has not been done for studies of multiple leading indicators as it depends on the indicator.

3.1.4 Lagging indicators

Lagging indicators and how their relationship to leading indicators was assessed, are summarised in Table 3.3. More detailed descriptions of the leading and lagging indicators are reported in Appendix G, Table G.2.

Each study reported between one and 11 lagging indicators, with 36 reporting more than one. The lagging indicators could be grouped into 10 categories: injuries (n=25 studies), accidents (n=18 studies), incidents (n=14 studies), near misses (n=8 studies), lost time (n=6 studies), fatalities (n=5 studies), compensation (n=4 studies), sickness/illness (n=3 studies), safety (n=1 studies) and 'other' (n=4 studies).

Although most categories of lagging indicator were evaluated by multiple studies, they were usually measured by using different methods. For example, although 18 studies reported accidents, five of these reported accident rates (or used accident rates to correlate against leading indicators) but used different denominators: 'average labour force' [38], 'per 200,000 man-hours' [28], 'per project per month' [51], 'per cubic meter of construction' [29] and 'per 10^6 working hours' [30]. Three used accident rate as part of a collated score with other lagging indicators [28, 44, 46]. Three studies focused on vehicle accident rates, but each one used a different definition of this outcome [14, 19, 22]. In three studies, the authors did not define accident rates [11, 23, 48] and one study measured percentage reduction in total work-related accidents and lost time accidents [17]. Furthermore, eight studies (including five already mentioned) did not report the lagging indicator by itself but the correlation of accidents against different leading indicators [15, 30, 36, 44, 46, 50, 51, 54].

3.1.5 Methods of assessing impact

Methods of assessing the impact of the leading indicators varied both between and within study design. Overall, more than 20 different approaches were used. Table 3.4 shows the method of analysis by study design and this is discussed in more detail in the following sections.

Table 3.4: Method of analysis by study design

Method of analysis	Study design					Total studies n (%)
	Cohort n (%)	Before-after n (%)	Cross-sectional n (%)	Case series n (%)	Case study n (%)	
Comparison of frequencies	3 (60)	10 83)	1 (6)	1 (17)	1 (13)	**16 (33)**
Linear correlation alone	1 (20)	1 (8)	4 (24)	3 (50)	4 (50)	**13 (27)**
Linear correlation and regression	1 (20)	0 (-)	2 (12)	0 (-)	0 (-)	**3 (6)**
Regression alone	0 (-)	0 (-)	4 (24)	1 (17)	1 (13)	**6 (13)**
Interrupted time series	0 (-)	1 (8)	0 (-)	0 (-)	0 (-)	**1 (2)**
Modelling	0 (-)	0 (-)	4 (24)	0 (-)	1 (13)	**5 (10)**
Multiple or other	0 (-)	0 (-)	2 (12)	1 (17)	1 (13)	**4 (8)**
Total	**5 (100)**	**12 (100)**	**17 (100)**	**6 (100)**	**8 (100)**	**48 (100)**

3.1.5.1 Cohort studies

In cohort studies, data from multiple companies were compared to determine whether the use of different leading indicators was associated with differences in lagging indicators. Methods used to perform this evaluation differed between the four cohort studies.

Three compared lagging indicator rates between leading indicator groups [18, 27, 47]. Brandt 2023 obtained worker-reported data from national surveys across industries and performed a time-to-event analysis to calculate the risk (using hazard ratios) of long-term sickness absence (LTSA) for groups of workers reporting no safety climate problems, compared to those reporting between one and five safety problems in the workplace [27]. Stough 2012 compared lagging indicators in companies with more leading indicators with companies with fewer leading indicators [18]. Van Derlyke 2022 examined the perceived effectiveness of leading indicators implemented in the dairy product-manufacturing sector. Survey data were used to compare Occupational Safety and Health Administration (OSHA) incident rates and ‘days away, restricted and transferred’ between companies that did and did not implement all eight of the top leading indicators [47].

The remaining two studies used statistical methods of correlation to examine associations between leading and lagging indicators [21, 36]. Wei 2020 used correlation coefficients to assess whether self-reported physical injuries and unsafe events were correlated with perceptions of safety culture in construction workers, and to make comparisons between two different regions [21]. Wachter 2014 performed both simple correlations and multivariate regression modelling to evaluate the relationship between safety management practices and accident rates across multiple industries [36].

The four studies assessed between 84 and 63,500 workers, though did not report number of companies included [21, 27, 36, 47]. All performed testing to identify which relationships were statistically significant.

3.1.5.2 Before-after studies

In before-after studies, lagging indicators are measured following the implementation of a leading indicator, such as a new training programme or safety framework, and compared to baseline data (before the leading indicator was initiated). All twelve before-after studies measured this using rates or frequencies [12-14, 16, 17, 19, 20, 22, 24, 42, 48, 49] and two studies also performed statistical correlations [16, 24]. Only two before-after studies reported statistical significance [16, 24].

Ten before-after studies [12-14, 16, 17, 19, 20, 22, 48, 49] were each conducted at one company and reported safety data before and after (or change in safety data after) a new safety programme or process was implemented. For example, Breitsprecher 2014 compared total recordable incident rate (TRincR) and motor vehicle accident rate before and after the implementation of leadership training for senior and middle managers [22]. López 2013 compared the number and rate of automotive accidents occurring in one company before and after a Regional Journey Management Centre (RJMC) was opened [14], while Pereira 2017 reported change from baseline (and Pearson correlations) of TRIR and total incidence rate before and after implementation of a behaviour-based safety programme in the construction industry.

Two studies reported data from more than one company [24, 42]. Choe 2016 performed an interrupted time series using data from the Bureau of Labor Statistics of the US Department of Labor fatalities to look at the fatality rate and days away rate following the revision of the OSHA Steel Erection Standard to evaluate change from baseline as well as correlation [24]. Govender 2022 looked at safety data (loss of life, total recordable case frequency rate, lost time injury frequency rate [LTIFR], all injury frequency rate, high potential incident) at nine diamond mines before and after a safety framework called the 4Cs framework was introduced [42].

3.1.5.3 Case series

In case series, leading and lagging indicators at multiple companies were measured over time to assess the relationship between them. Four case series assessed the relationship using statistical correlations [30, 40, 45, 50], one compared safety outcome score [28] and one compared expected accident rate to actual accident rate [29]. Those that reported the number of companies included in the study included between three [50] and 126 [30] and those that reported the number of workers included reported between 11 [45] and 23,399 [30].

In the studies using correlation, one used Spearman's correlation [50], one used Pearson's correlation [30], one used both [45] and one used stepwise regression analysis and correlation [40]. Grabowski 2010 used data from surveys to retrospectively identify which leading indicators were correlated with safety outcomes at three companies [50]. Laitinen 2013 evaluated the correlation between a company's score on an observation tool and their accident rates [30]. Salas 2016 and Vosoughi 2021 both assessed the correlation between multiple leading and lagging indicators [40, 45].

3.1.5.4 Cross-sectional studies

Seventeen studies measured leading and lagging indicator performance at one timepoint to examine the relationship between the two. Ten provided information on company participants [11, 25, 26, 33, 35, 38, 44, 52, 53, 55] evaluating between three [35] and 3,514 [25] companies. Hinze 2023 reported on 57 projects but did not provide information about the company or companies involved and Merrick 2014 reported on two shipping fleets [39, 54]. The remaining three studies did not provide any company details [31, 32, 43]. Nine studies included between 112 and 3,578 workers [26, 31, 32, 37, 43, 44, 46, 52, 54], while the remaining eight studies did not report this information.

Methods used to assess the relationship between leading and lagging indicators varied and the statistical significance of the findings was well reported with 15 of the 17 studies reporting this information [25, 26, 31-33, 35, 37, 39, 43, 44, 46, 52-55]. The most common statistical analysis method used was a form of regression analysis which was performed in four of the included studies [25, 31, 43, 54]. For example, Moore 2022 used linear regression analysis to assess the relationship between the SH-26 assessment rating and workers' compensation claims.

Four studies used correlations to assess the relationship between leading and lagging indicators [35, 39, 53, 55]. For example, Tang 2018 used Pearson's correlation to examine whether safety factors (e.g. emergency management, personal safety and hazard identification) were associated with near miss and lost time injury rates. Three studies used mixed methods. Dennerlein 2020 used a both univariate correlation and multivariable regression to assess the relationship between assessments of contractor safety (ACES), OSHA recordable injuries and injuries involving days away, restricted or transferred (DART) [52]. Manjourides 2019 used a combination of spearman correlation and zero-inflated Poisson modelling to estimate the relationship between safety management, safety programs, special elements and DART injuries [53]. Robson 2017 used unstandardised correlation coefficients and predictive modelling to investigate the relationship between audit scores and workers' lost time claims [33].

Sá 2023 did not use any statistical any methods to assess the relationship between the leading and lagging indicators. The five remaining studies each used different approaches, including: confirmatory factor analysis to examine the relationship between safety climate, accidents, injuries and near misses [46]; Wilcoxon signed ranks tests, factor analysis and hierarchical cluster analysis to assess the links between safety factors and incidence occurrence [26]; multilevel modelling to examine near misses and incidents [37]; partial least square-based structural equation modelling to determine the significance of leading indicators in being able to predict OHS performance [44]; and classification tree approaches to look at the association between prevention management strategies and accident rates in difference companies [38].

3.1.5.5 Case studies

We have defined case studies as evaluations of a single company measuring leading and lagging indicators over time. All eight studies provided details of company participants. Seven reported on one company [5, 6, 15, 34, 41, 51, 56] and The Campbell Institute 2015 reported on five different case studies [23]. Two studies described the companies they reported on to be large [15, 34] and one as 'mid-sized' but the number of employees were not reported [6]. The remaining five studies did not report on company size.

Methods used to assess the impact of leading indicators on lagging indicators differed, and the statistical significance of findings was rarely reported (one study [51]). The most common approach was linear correlation to estimate the strength and direction of relationships between leading and lagging indicators (n=4 [15, 23, 51, 56]). For example, Rajendran 2013 calculated scores for the degree of implementation of three leading indicators in construction sites (worker safe behaviour observations, pre task planning and site safety audits) and computed correlation coefficients for their relationships with near miss incident rates, first aid injury rates, OSHA recordable injury rates and total incident/injury rates [56].

The remaining four studies each used different approaches, including: univariate regression modelling to estimate the relationship between the number of safety talks/inspections and medical injury rates as measured across 47 construction projects [6]; temporal vector autoregression (VAR) modelling to estimate the temporal relationships between 14 different leading indicators and TRIR [5]; simple frequency comparisons across multiple timepoints to examine trends in safety performance over time (number of safety observations against several lagging indicators, including fatalities, total recordable injury rates, lost time injury rates and first aid cases) [41]; and qualitative comparative analysis to identify connections between safety practice indicators and safety performance [34].

3.1.6 Study validity

Internal and external validity are discussed in the following sections. As discussed in both Section 3.1.6.1 and 3.1.6.2, poor reporting was a common issue and incomplete information on study methods posed challenges to identifying and assessing how studies were conducted. Brandt 2023 [27] is summarised in Appendix H, Figure H.1 as an example of a study with low risk of bias that may be applicable to a wider population.

3.1.6.1 Risk of bias

RoB assessments are summarised below. The detailed RoB assessments for each study are reported in Appendix E, Table E.1.

At least one study design weakness was identified in each study, and so no studies were considered at very low RoB. In total, 15% (seven studies) were considered to be at low RoB (weakness in one domain) [6, 24, 27, 33, 37, 46, 47], 56% (27 studies) at moderate RoB (weakness in two to three domains) [5, 12-14, 16, 21, 22, 25, 26, 29-32, 34-36, 38-40, 43, 45, 50-55] and 29% (14 studies) at high RoB (weaknesses in at least four domains [11, 15, 17-20, 23, 28, 41, 42, 44, 48, 49, 56]. Reviewer judgements for each domain in each included study are provided in Figure 3.1.

The most common limitation was the failure (or the potential failure) to adequately minimise selection bias, for example by using convenience samples [13, 30, 34, 52], removing companies reporting poor safety outcomes from analyses [53], or most frequently by simply not providing the rationale for why study participants were selected (Figure 3.2) [11, 14, 18-20, 23, 28, 32, 35, 37, 39, 42, 46, 48-51, 54, 55].

Few studies provided information on methods to account for confounding factors, for example by adjusting for key characteristics in regression models [27], and the majority of studies failed to consider confounding or to report data on potential confounders (such as country or worker characteristics). The use of incorrect statistical analysis was the least common issue, while potential sources of bias outside of the four assessed domains were rarely identified. Studies often provided insufficient information to determine whether leading indicators or lagging indicators had been measured without bias; this bias was most commonly judged to be at high risk when studies collected self-reported data to measure leading and/or lagging indicators, without using validated tools [11, 21, 34, 39, 40, 50, 53-55].

When examined by study design, cohort designs appeared to more often be at a lower RoB and all other designs were judged to be at moderate to high overall RoB (Table 3.5).

Figure 3.1a: Summary risk of bias judgements for included studies

Study	D1	D2	D3	D4	D5	Overall
Alarcón 2016	+	X	X	+	-	X
Amir-Heidari 2017	X	-	X	X	X	X
Bitar 2018	+	X	X	+	-	-
Brant 2023	-	+	+	+	+	+
Breitsprecher 2014	+	X	-	X	+	-
Campbell Institute 2015	X	-	-	-	-	X
Cao 2019	X	-	-	+	+	-
Chen 2017	X	-	-	+	+	-
Choe 2016	+	+	-	+	+	+
Coetzee 2023	X	X	-	X	+	X
Dadashi Haji 2023	X	X	+	X	X	X
Dennerlein 2020	X	+	X	+	+	-
Doherty 2010	X	X	+	X	+	-
Gale 2011	-	-	+	X	+	-
Govender 2022	-	-	-	X	-	X
Grabowski 2010	X	-	X	+	+	-
Haas 2018	X	-	X	X	-	X
Hinze 2013	X	-	X	+	+	-
Lagerstrom 2019	-	-	+	+	+	-
Laitinen 2010	X	X	+	+	+	-
Laitinen 2013	X	-	+	+	+	-
Lingard 2017	X	-	+	+	+	-
López 2013	X	X	+	X	+	-
Manjourides 2019	X	X	X	+	+	-

Risk of bias

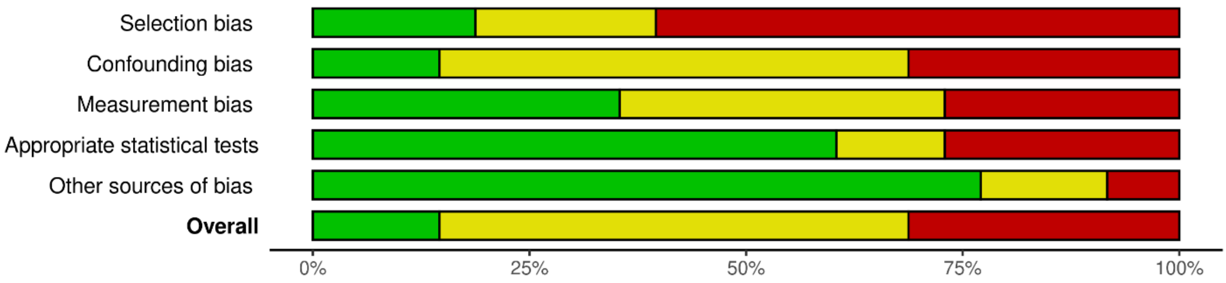
D1: Are you confident that the means of selecting and maintaining the sample minimised bias?
D2: Are you confident that the potential confounders were adequately considered, and then were either well controlled or appropriately discounted as a source of bias?
D3: Are you confident that the measurement methods did not introduce bias to the corresponding findings?
D4: Were appropriate statistical tests applied to the data?
D5: 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?
Red=high RoB; yellow=moderate RoB; green=low RoB.

Figure 3.1b: Summary risk of bias judgements for included studies

	Risk of bias					
	D1	D2	D3	D4	D5	Overall
Manjourides 2019	X	X	X	+	+	-
Merrick 2014	X	-	X	+	+	-
Mohammed 2019	X	X	-	-	X	X
Moore 2022	+	-	-	+	+	-
Mousavi 2020	-	X	-	+	-	X
Pereira 2017	-	-	+	+	+	-
Quaigrain 2023	X	X	-	+	+	-
Rajendran 2013	X	X	+	-	X	X
Robson 2017	+	+	-	+	+	+
Sá 2023	X	-	X	-	+	X
Salas 2016	-	X	X	+	+	-
Schiavi 2013	X	X	-	X	+	X
Sheehan 2016	X	+	+	+	+	+
Stough 2012	X	-	-	X	+	X
Tang 2017	-	-	+	-	+	-
Tang 2018	-	-	+	+	+	-
Tauseef 2012	X	-	-	X	+	X
Thananan 2014	X	-	-	X	+	X
Van Derlyke 2022	+	-	+	+	+	+
Versteeg 2019	+	-	+	+	+	+
Vosoughi 2021	-	-	-	+	+	-
Watcher 2014	+	-	-	+	+	-
Wei 2020	X	-	X	+	+	-
Winge 2019	X	+	X	+	-	-
Zahoor 2017	X	+	+	-	+	+

D1: Are you confident that the means of selecting and maintaining the sample minimised bias?
D2: Are you confident that the potential confounders were adequately considered, and then were either well controlled or appropriately discounted as a source of bias?
D3: Are you confident that the measurement methods did not introduce bias to the corresponding findings?
D4: Were appropriate statistical tests applied to the data?
D5: 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?
Red=high RoB; yellow=moderate RoB; green=low RoB.

Figure 3.2: Summary risk of bias plot by domain, for 48 included studies.



For each of the five domains: Green=% studies judged at low RoB, yellow=% studies judged at unclear RoB, red=% studies judged at high RoB.
For the overall RoB: Green=% studies judged at low RoB, yellow=% studies judged at moderate RoB, red=% studies judged at high RoB.
Plot generated using Robvis shiny web app [58].

Table 3.5: Overall risk of bias for 48 included studies, by study design

Study Design	Overall risk of bias			
	Low n (%)	**Moderate n (%)**	**High n (%)**	**Total n**
Before-after	1 (8)	5 (42)	6 (50)	12
Cohort	2 (40)	2 (40)	1 (20)	5
Cross-sectional	3 (18)	12 (71)	2 (12)	17
Case series	0 (-)	5 (83)	1 (17)	6
Case study	1 (13)	3 (38)	4 (50)	8
Total	**7 (15)**	**27 (56)**	**14 (29)**	**48**

3.1.6.2 External validity

It is difficult to assess the degree to which the results of included studies are generalisable beyond the study setting due to the limited reporting of key information, particularly company and worker characteristics, study setting, study size, and in some cases country of study. This means that, for the greater majority of studies, it is not possible to assess whether they could be considered representative of real-world settings.

Overall, generalisability is likely to be weak. Over a third of included studies (n=18) evaluated data from only one company [5, 6, 12-17, 19, 20, 22, 34, 39, 41, 48, 49, 51, 56]. Most of these studies did not report company size or the number of workers included in the study. Nonetheless, differences between how companies operate may mean that even studies of one company evaluating high worker numbers are not necessarily more widely generalisable.

Another common issue was the evaluation of leading indicators that were specific to the company being evaluated in the study. For example, Doherty 2010 was a study of an oil company in Qatar, RasGas, evaluating the RasGas Elements of Excellence, the company's operations integrity management system [12]. Also, Breitsprecher 2014 evaluated the company's own health, safety and environment (HSE) leadership academies [22]. Since the leading indicators and mechanisms for measuring them were developed within the context of each company's specific setting, these studies are likely to be poorly generalisable to other companies.

Lastly, most studies evaluated multiple leading indicators (n=25) and/or multiple lagging indicators (n=32), or used complex scores developed within-study to measure them (n=7). This, in addition to the substantial variation in all study methods, hinders the ability to identify which factors could be contributing to observed correlations or relationships between leading and lagging indicators, with implications for being able to replicate beneficial findings in future practice.

3.2 Effectiveness of Leading Indicators

In total, 27 studies found that at least some of the leading indicators measured were favourably associated with lagging indicators (in 20 studies it was unclear whether there was an effect and in only one was there not an effect). However, studies were too heterogeneous in the methods used to present a meaningful distribution of leading indicator effects. The following section summarises the evidence base for each type of leading indicator, describing the variation in methods used.

3.2.1 New guideline, tool or process

All 15 studies found new guidelines, tools or processes were associated with improvements in safety outcomes [11-14, 16, 17, 19, 20, 22, 24, 29, 41, 42, 48, 49] but only three reported that the relationship was statistically significant [16, 24, 29] while the remaining 12 did not report statistical significance. The quality of the study methods and reporting, and therefore confidence in their results, varied.

Choe 2016 and Pereira 2017 explored whether observed differences in lagging indicators could be explained by chance using statistical hypothesis testing for significance (reported as p values for those differences in outcomes) [16, 24]. Using a before-after comparison, Choe 2016 reported significant reductions in the days away injury rate (and normalised days away injury rate) following the implementation of the revised OSHA Steel Erection Standard, although the reduction in fatality rate (and normalised fatality rate) was not found to be statistically significant [24]. Pereira 2017 also reported significantly improved total incidence rates following the implementation of a behaviour-based safety programme at building sites [16]. Laitinen 2010 evaluated the trend in accident rates following the implementation of a safety contest that rewarded the best performing companies and reported that safety significantly improved in the years following the implementation of the contest [29].

The remaining 12 studies [11-14, 17, 19, 20, 22, 41, 42, 48, 49] reported improvements in lagging indicators associated with the implementation of the leading indicator (reported in Table 3.3), but none reported whether differences were statistically significant. In addition, only one study considered whether the relationship could be explained by potential confounders [24] while participant and population characteristics were not well reported in any studies.

Studies varied in the lagging indicators they measured, the methods of measuring lagging indicators, and the methods of evaluating the extent to which they were associated with leading indicators. Some studies for example reported lagging indicators that no other study reported, such as driving score [13] or heat injury rate during Ramadan [12]; others used differing methods to measure similar lagging indicators, for example Dadashi Haji 2023 reported site-specific fatality rates before and after the leading indicator [48] while Choe 2016 reported the total number of fatalities before and after [24]. Similarly, Laitinen 2010 reported accident risk per cubic metre of construction [29] while Breitsprecher 2014 reported the TRincR [22] and Coetzee 2023 the all injury frequency rate [41].

Use of a broadly similar analytic approach did not guarantee the comparability of study findings. For example, the two studies that performed correlations varied in specific approach, Choe 2016 conducting a time series analysis to evaluate the change in fatality rate and days away injury rate [24], while Pereira 2017 evaluated incident rates using linear correlations at a single timepoint [16].

3.2.2 Safety climate/culture

Five included studies (three cross-sectional [32, 43, 46] and two cohort studies [21, 27]) evaluated safety climate or safety culture within organisations as a safety leading indicator, three of which were carried out in the construction industry [21, 32, 46]. Between 426 [46] and 63,500 [27] workers were assessed, although characteristics of both companies and workers were poorly reported with only two studies reporting this information [32, 46].

The five studies varied in the methods used and quality of reporting. For example, methods of collecting the safety climate data differed with two studies using the NOSACQ-50 [27, 43] while the remaining three studies [21, 32, 46] used unvalidated questionnaires based on previous research or developed by the researchers. Studies sought to determine whether safety climate was associated with a range of different lagging indicators, including long term sickness absence [27], injury outcomes [21, 32], musculoskeletal symptoms and subsequent missed work [43], unsafe events [21] and number of self-reported accidents/injuries and near misses [46]. Only two studies considered potential confounders [27, 46].

Four studies reported safety climate/culture to be associated with improvements in lagging indicators [21, 27, 32, 43]. One cohort study [21] and one cross-sectional study [32] reported statistical correlations between safety climate/culture and lagging indicators. Chen 2017 found a statistically significant negative correlation between safety climate and physical symptoms [32], while Zahoor 2017 reported a significant negative correlation between safety climate and self-reported accidents/injuries but did not provide p values or information on whether or how these were conducted [46].

Brandt 2023 [27] and Lagerstrom 2019 [43] reported similar findings. In a time-to-first-event analysis, Brandt 2023 found that the risk of long-term sickness absence was significantly higher in groups of workers reporting four to five safety climate problems in the workplace [27]. Using multinomial logistic regression, Lagerstrom 2019 found statistically significant relationships between safety climate and both the frequency of musculoskeletal symptoms and associated missed working days [43].

3.2.3 Audits and inspections

Two studies (a case study [30] and a cross-sectional study [33]) evaluated the effect of an audit or inspection on lagging indicators and both reported statistically significant correlations between the two, although one was judged overall not to demonstrate a relationship between leading and lagging indicators.

The Laitinen 2013 case series found significant negative correlations between the scores from an audit tool called Elmeri+ (collected from mechanical engineering, metal industry and electronics industry companies) and all three accident rates evaluated, although subgroup analyses found this correlation was not observed in every industry or for different audit sub-scores [30].

Robson 2017 evaluated data from firms operating in various industries that had undergone one or two audits as part of the WorkWell audit programme in Canada. Correlations between total audit score and various different lagging indicators were calculated for all included companies as a group, as well as for specific industry subgroups [33]. A high number of correlations were calculated, with the majority suggesting an absence of statistically significant relationships. The small number of significant negative correlations between audit and claims differed between audit 1 and audit 2 showing inconsistency of effect; therefore, this study was categorised as not showing an effect between leading and lagging indicators.

3.2.4 Monitoring of safety

Moore 2022 used questionnaires to collect safety hazards and management practices data from employees from various industries and found a statistically significant correlation between twelve of the hazards and higher claims outcomes [31].

3.2.5 Mixed

All remaining 25 studies evaluated multiple leading indicators, five of which [18, 28, 37, 47, 52] combined multiple leading indicators to give an overall study-specific leading indicator score. None of these used the same score, evaluated similar lagging indicators, or used similar methods. The leading and lagging indicators combined to get the scores are listed in Appendix G, Table G.2.

The remaining 20 studies reported multiple leading indicators separately, though with considerable heterogeneity in the study methods used and the lagging indicators measured.

In fifteen of the remaining studies, there were statistically significant relationships between at least some of the leading and lagging indicators evaluated [6, 25, 34-36, 38-40, 44, 45, 50, 51, 53-55]. In the remaining five, there was a relationship between at least some of the leading and lagging indicators but no statistical tests were conducted to show whether the relationships were statistically significant [5, 15, 23, 26, 56].

4 Discussion

A scoping review was conducted to map the key characteristics of studies that assess the impact of leading indicators on safety lagging indicators in the discipline of OSH. The review was conducted according to the methods for scoping reviews, outlined by the Joanna Briggs Institute [8]. In total, 48 studies met the eligibility criteria for this review.

4.1 Summary of Key Findings

Overall, the studies identified by this review were found to be very heterogeneous in every study characteristic examined. Reporting of study characteristics was also variable, with key information missing from studies, particularly concerning worker characteristics. This section draws together the main findings to address the five key review objectives.

4.1.1 For which leading indicators has effectiveness been evaluated?

Due to variations in terminology, it was challenging to identify whether studies had sought to measure the same leading indicators. Differences in methods of measurement and company-specific safety management systems and practices prevented an assessment of the total number of different leading indicators evaluated by these studies. Using the information reported, we assessed that included studies evaluated a large number of different leading indicators, representing a wide range of safety practices. Studies most commonly evaluated the impact of new guidelines, frameworks, tools or processes, though the specific leading indicator(s) differed in each of those 15 studies.

More than half of all eligible studies (n=25) evaluated multiple leading indicators, either assessing each individually or as groups using summary scores to investigate relationships with lagging indicators. Across all studies, leading indicators were often either specific to the company in which they were being measured or specific to that industry.

The leading indicators evaluated in this evidence base did not all fully meet previously published definitions and were twice as likely to meet the Campbell Institute criteria than the Xu 2021 criteria (30 vs 14 respectively), mainly due to the indicators not being judged to be measuring current performance.

4.1.2 For which industries and in which countries does this evidence come from?

Studies were performed across eight different industries, most commonly construction (n=18) or energy (oil and gas drilling and transportation) (n=13), while a small proportion sought to evaluate leading indicators across multiple industries. Country settings were not always reported. The 39 studies providing this information were conducted in 22 different countries across all five continents, most commonly in North America and Asia. Most studies were conducted in a single country, at least 12 involved multiple countries and at least two multiple continents.

4.1.3 Are different leading indicators evaluated in lower to middle income countries, compared to high income countries?

While country income status as designated by the World Bank could not be assessed for 12 studies due to lack of reporting, most research was conducted in HI settings (27/48), with only four each in UMI countries and LMI countries, and no eligible studies identified from LI countries. When examining type of leading indicator, there did not appear to be any clear association between certain indicator types and countries in each outcome status: new guideline, tool or process indicators were evaluated in HI, UMI and LMI countries, as were mixed leading indicators. The five studies of safety climate were conducted in both HI and LMI settings, while other leading indicator categories included too few studies to identify any potential difference.

4.1.4 How has effectiveness been evaluated?

Eligible studies used one of five observational designs to assess the impact of leading on lagging indicators, most commonly cross-sectional designs that measured the performance of all indicators at one timepoint and statistically assessed the relationship between the two using a method of statistical correlation and/or regression. However, study designs and methods of analysis varied considerably across the 48 studies.

Other designs identified by the review team included cohort studies (n=5), before-after studies (n=12), case series (n=6), and case studies (n=8). The most common methods of analysis were linear correlation (most often calculating Pearson's coefficient) and simple comparison of frequencies, each used by a third of included studies. Other methods included regression analyses, modelling approaches (including structural equation modelling), interrupted time series, classification trees, principal factor analysis and qualitative analysis, most of which comprised more specific methods used by a single study.

Studies assessed the impact of leading indicators against a wide variety of lagging indicators, covering ten categories of safety outcomes, including injuries, accidents, incidents, near misses, lost time, fatalities, compensation claims, sickness/illness, safety and other outcomes. Most studies evaluated the impact of a leading indicator against multiple lagging indicators. Injuries and accidents were the most frequently reported, although methods of defining and measuring these outcomes differed considerably between studies, which could restrict the meaningful synthesis of effects in future systematic reviews.

4.1.5 Is there any evidence for the effectiveness of leading indicators, and what is the nature of that evidence?

All but one included study reported a positive effect of leading indicator on lagging indicator outcomes. However, results were very varied and while most studies reported multiple analyses, none found all analyses to produce evidence of an effect. In addition, claims of impact were not always verified by the use of appropriate statistical methods to determine whether effects were likely due to chance, therefore this review concludes that evidence for some impact was produced by 27 (56%) studies.

Unfortunately, the considerable heterogeneity across all aspects of how these studies were conducted precludes the ability to meaningfully synthesise or summarise individual effect sizes. A systematic review published in 2019 (identified during study selection) performed a meta-analysis of effects across eight studies assessing the effects of leading on lagging indicators in the construction industry [59], of which four studies were included in this scoping review [38-40, 56]. In order to overcome the different measures of effect used by these studies, the authors performed statistical calculations to standardise the effects and so make them comparable. Other than class of leading indicator (passive and active), no other variation in study characteristics was taken into account prior to undertaking the pooling of effects and the authors found significant heterogeneity of effect across both classes of leading indicator, which may cast doubt on the validity of their findings.

A 2020 report from an American research group research suggests it might be challenging to make valid inferences on safety performance using accident frequency rates. A large study involving statistical analysis of a very substantial dataset from the construction industry (10 organisations providing 15 years of data covering 3.26 trillion worker-hours) found that TRIR was not correlated with fatality rates, suggesting that the causes for accident rates might differ from the causes for fatalities. In addition, when building models to try and predict future TRIR, the best models could only predict the observed rates in 2-4% of model iterations, suggesting that TRIR does not have a causal relationship with implementation of safety management systems. More broadly, these findings may imply that some observed associations could be spurious, rather than presenting evidence for a causal relationship between use of leading indicators and some measures of accident frequency [60].

4.1.6 How robust is this evidence base, and how can it be improved?

Overall, the quality of included studies for demonstrating that leading indicators have causal impacts on lagging indicators is weak, for three reasons.

First, the evidence base identified by this review consisted entirely of observational study designs, and only a fraction (10%, the five cohort studies [18, 21, 27, 36, 47]) provided comparisons of lagging indicators measured in groups of workers who had experienced different leading indicators. Yet establishing whether leading indicators have a causal impact on lagging indicators is likely to require prospective, comparative studies (ideally interventional in nature) whereas many differences between groups as possible are controlled, so that the only difference between them is the leading indicator(s) being applied. None of the included studies was of a controlled, comparative design.

Second, the internal validity of included studies was determined to be moderate to low indicating that the results observed in these studies are likely to reflect problematic study design rather than true findings. Fewer than one in five studies was judged to be at low RoB using the Robson 2007 tool. Over half were at moderate RoB and 30% were determined to be at high RoB, meaning that those studies are highly likely to have produced distorted estimates of effects.

Including biased study samples was one of the most common design deficiencies, indicating that studies may have excluded companies or workers less likely to respond to leading indicators. Studies also commonly failed to account for confounding factors thus risking confounding bias [61], namely a distortion of study effects due to confounding by unmeasured or uncontrolled risk factors. Together with frequently scant reporting of worker characteristics, this suggests that confounding is not commonly considered during the design and conduct of studies that evaluate leading indicators. The identification of key confounders, their measurement, and incorporation into the analysis and interpretation of study results, are likely to constitute important improvements to the quality of the evidence base. Similarly, measuring indicators (particularly lagging) using objective tools (such as through electronic records) rather than questionnaires based on self-reported information is likely to improve the internal validity of studies.

The Robson 2007 systematic review reported similar findings in terms of the quality of their 23 included studies: few were comparative (3/23), the overall quality was determined to be weak, and common sources of potential bias were failing to account for confounders and measurement methods. The authors are not aware of any other published reviews that have assessed the internal validity of primary studies undertaken in OSH settings.

Finally, the generalisability of the evidence base was determined to be weak for several reasons, including: limited reporting of key information (particularly company and worker characteristics, study setting, study size and, in some cases, country of study), evaluating a single company, evaluating leading indicators specific to the participating company, and evaluating multiple leading indicators and/or multiple lagging indicators, and often using complex scores developed within-study to measure them. This, in addition to the substantial variation in the lagging indicators evaluated, hinders the ability to identify which factors could be contributing to observed correlations or relationships between leading and lagging indicators, with implications for being able to replicate beneficial findings in future practice.

4.2 Strengths of the Review

This review has used explicit, systematic methods captured in a predefined review protocol that was registered on the OSF. Transparent and replicable search methods to find relevant studies were designed by experienced Information Specialists, and while the review type did not require these searches to be comprehensive, multiple resources were searched and previous reviews were examined for potentially relevant studies.

A key strength was the ability to consult with a group of OSH experts from varying backgrounds during the development of the review protocol to ensure validity of research aims, objectives and review methods.

The results of our search underscored the absence of existing reviews seeking to summarise the characteristics of the existing evidence base, without limiting to particular industries or leading indicators. This review therefore makes a unique contribution to the research literature.

4.3 Limitations of the Review

Search terms were limited to retrieve records which referred to the concept of leading indicators. Studies that only referred to specific indicators would not have been retrieved. A limit was placed on the language of the search for reasons of resource constraints. Also, during screening, a date limit (pre-2010) was placed on the records due to the volume of literature. In addition, a large proportion of documents could not be obtained for assessment of relevance at full text stage (50 of 266). Using the same proportion of relevant studies as those identified from obtained full text reports, we estimate to have missed around 11 eligible studies. While their contents remain unknown, they are unlikely to impact substantially on the key findings in this report.

A related limitation concerns the focus on published studies, which may have missed relevant unpublished data that demonstrate positive impacts, such as internal company benchmarking data. While we did obtain and examine some unpublished reports, those not known or available to our expert group could not be considered or included. The impact of this to the conclusions in this review cannot be quantified. It is also not possible to determine whether organisations that were represented by the included studies would be judged to have a good safety performance, since this review did not set out to examine the interdependencies between programmes and leading indicators.

This review used single reviewer study selection and data extraction, increasing the possibility of having missed studies or extracted information incorrectly. However, both of these critical review stages were performed in duplicate for the first 10% of studies in order to ensure consistency of approach.

A large number of studies (n=25) reviewed multiple leading indicators, some of which (n=5) did not report the performance of individual leading indicators separately but gave a composite score summarising the multiple leading indicators in use. Furthermore, some of the guidelines, tools or processes that were introduced involved a programme with a number of leading indicators. These features limit our ability to understand the performance of individual leading indicators by comparing across studies. However, it may also suggest that the approach of looking for associations between individual leading indicators and safety outcomes may not adequately represent real practice where a range of indicators, appropriate to the company’s health and safety needs, tend to be tracked in tandem.

Finally, the scope of this review focussed on understanding impacts to safety performance, as measured by safety lagging indicators, to the exclusion of impacts to occupational health. However, it is well recognised that employee state of mind and health can adversely impact situational awareness and decision making, which might lead to adverse safety outcomes. It is therefore possible that this review has missed studies demonstrating positive changes to occupational health outcomes that could be valid surrogates for downstream impacts to safety.

4.4 Implications for Practice

This review shows a huge amount of variation in the current evidence base for leading indicators, which makes it difficult to identify recommendations for practice based on this existing literature. The inconsistency in relationship between leading and lagging indicators, which reflects the findings of previous studies, suggests practitioners should exercise caution when seeking to evaluate the impact of leading indicators. Setting out to identify one particular metric or intervention that is a "silver bullet" is unlikely to produce meaningful findings or findings that can be translated across settings.

Perhaps the principal challenge to being able to discern whether different leading indicators have a consistent impact was the inability to compare findings across studies. If comparability across sites, units, companies or industries remains a goal of future research, then more standardised data collection and wider data sharing should be encouraged.

Although the principle of using leading indicators to support organisational safety and health decision-making is still valid, currently their greatest value lays in organisation-specific evaluations that set out to assure the organisation's key risk control systems. The challenge for this more introspective goal will be to identify a set of indicators that demonstrate the ongoing integrity of that safety and management system.

4.5 Future Research

There is a clear need for research to address whether leading indicators can improve occupational health, and to investigate whether this outcome could be a valid surrogate for downstream safety performance.

In order to understand whether specific leading indicators improve safety outcomes across settings, syntheses with statistical pooling of effects would be ideal. These would ideally pool the results of prospective, comparative studies that have controlled for confounders, selection bias and measurement bias.

To achieve this, future studies would also need to:

- Carefully recruit under study conditions, rather than using convenience samples.
- Be designed with consideration for the future generalisability of results.
- Use standardised metrics to measure and report leading and lagging indicators.

Primary studies should also report study methods more completely. For example, understanding worker characteristics would help researchers and other organisations to assess the applicability of results to other settings, while more complete reporting of measurement methods would enable an assessment of the comparability of evidence produced by that study. Clear reporting also enables accurate risk of bias assessments, ensuring that stronger studies are more reliably identified. Since the risk of bias tool used by this review was developed from similar tools to assess the internal validity of epidemiological studies, future research is needed to understand whether the tool captures the most important potential sources of bias in OSH studies of leading indicators.

Given the large number of leading indicators that were identified in the review, it may be beneficial to first identify which leading indicators (or types of leading indicators) are most commonly used in practice, in order to better focus future research. While we developed a definition of leading indicators for use in this review based on the literature, there was such a variation in definitions used across the literature that some leading indicators in studies included in this review did not meet other definitions of leading indicators. Greater clarity on the definition of a leading indicator is needed.

Creating a more comparable evidence base would also enable researchers to identify and investigate reasons for not observing expected associations with lagging indicators. Further research is needed to determine which lagging indicators are the most valid to reach conclusions on safety performance.

It is important to acknowledge that these recommendations are provided from the perspective of healthcare science research, in which an understanding of which evidence is most reliable and useful has been developed specifically for healthcare, over several decades. Indeed, this scoping review applied a research framework developed within evidence-based medicine and aimed to document evidence for the impact of individual leading indicators on lagging indicators. However, the resulting heterogeneity of evidence it identified may also reflect the complexity of how leading indicators are typically used in practice. In contrast to pharmaceutical interventions where a single intervention is administered and the effects measured shortly after (and with all related behaviours controlled for to eliminate confounding), in OSH practice a range of indicators covering a range of safety management functions are selected and tracked together. Subsequent safety and health decision-making tends to be informed by directional changes in those indicators. A key challenge to future research will be to identify methods by which more controlled evaluations of groups of indicators can be performed, and what level of heterogeneity might be acceptable when seeking to compare results across multiple studies.

Some of these elements of safety management systems and decision-making share features with the ‘complex interventions’ used in health and social care, that are composed of multiple components which interact in complex ways to achieve their intended effect [62]. Complex intervention research typically seeks to address broader questions than whether an intervention achieves its outcome, for example understanding how it contributes to system change, theorising how it works or exploring its value relative to the resources required for delivery [63].

Considering the differences in practice and decision-making between healthcare and OSH, it may be valuable to consider developing evidence standards that are specific to the OSH context. Such standards would help to ensure that future research in this area is useful and contributing to improvements in leading indicator practice and implementation. Developing the standards through active engagement with a range of stakeholders will also be important to ensure that the standards produce the most appropriate evidence, and to maximise the potential for it to impact on OSH practice.

4.6 Conclusion

This review has identified a substantial, though disparate, evidence base evaluating the impact of leading indicators on safety lagging indicators. Almost all studies reported a positive impact, though the degree to which these findings are reliable indicators that leading indicators cause changes to lagging outcomes remains unknown for several reasons. The overriding characteristic of the evidence base is the heterogeneity of topics evaluated, and the methods used to evaluate them. This research effort provides an optimistic signal to the discipline that leading indicators are being empirically evaluated in terms of their impact on lagging indicators. Yet the variation in approach, as well as the dynamic nature of risk and culture in different industries and in different countries, may also prevent generalisations to be made regarding the ability of individual (or specific groups of) leading indicators to reliably improve the safety performance across industries. In order to gain an understanding of the general utility of leading indicators and which leading indicators are best to deploy in different setting, future studies should improve their approaches to minimising bias, and identify common tools to measure both leading and lagging indicators to facilitate the comparison of studies. Developing an OSH-specific evidence standards framework is likely to guide and assist this process.

5 References

1. International Labour Organization. Nearly 3 million people die of work-related accidents and diseases. Geneva: ILO; 2023. [cited 15th March 2024]. Available from: https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_902220/lang--en/index.htm.
2. Occupational Safety and Health Administration. Using Leading Indicators to Improve Safety and Health Outcomes. Washington: Occupational Safety and Health Administration (OSHA); 2019. [cited 11 July 2023]. Available from: https://www.osha.gov/sites/default/files/OSHA_Leading_Indicators.pdf.
3. Grabowski M, Ayyalasomayajula P, Merrick J, Harrald JR, Roberts K. Leading indicators of safety in virtual organizations. Safety Science. 2007.45(10):1013-43. doi: https://doi.org/10.1016/j.ssci.2006.09.007
4. Inouye J. An Implementation Guide to Leading Indicators. Campbell Institute; 2019. [cited 17 July 2023]. Available from: https://www.thecampbellinstitute.org/wp-content/uploads/2019/08/Campbell-Institute-An-Implementation-Guide-to-Leading-Indicators.pdf.
5. Lingard H, Hallowell M, Salas R, Pirzadeh P. Leading or lagging? Temporal analysis of safety indicators on a large infrastructure construction project. Safety Science. 2017.91:206-20. doi: https://doi.org/10.1016/j.ssci.2016.08.020
6. Versteeg K, Bigelow P, Dale AM, Chaurasia A. Utilizing construction safety leading and lagging indicators to measure project safety performance: A case study. Safety Science. 2019.120:411-21. doi: https://doi.org/10.1016/j.ssci.2019.06.035
7. Pawlowska Z. Using lagging and leading indicators for the evaluation of occupational safety and health performance in industry. Int J Occup Saf Ergon. 2015.21(3):284-90. doi: 10.1080/10803548.2015.1081769
8. Peters M, Godfrey C, McInerney P, Munn Z, Tricco A, Khalil H. Chapter 11: Scoping reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis; 2020.
9. Robson L, Clarke J, Cullen K, Bielecky A, Severin C, Bigelow P*, et al.* The effectiveness of occupational health and safety management system interventions: a systematic review. Safety Science. 2007.45(3):329 - 59. doi: https://doi.org/10.1016/j.ssci.2006.07.003
10. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD*, et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021.372:n71. doi: 10.1136/bmj.n71
11. Sá JC, Dinis-Carvalho J, Fraga H, Lima V, Silva FJG, Bastos J. The impact of lean on occupational safety in organisations. In: IFIP Advances in Information and Communication Technology, 2023. 184-92
12. Doherty BD, Fragu LP. Sustainable HSE performance: Successful management systems and monitoring tools in the Middle East LNG industry. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2010, 2010. 1445-63
13. Gale BR, Trostel MV, Armitage DL, Mason MA, Bautista LS. Case study: Successful implementation of driving safety and IVMS program. In: Proceedings - SPE Annual Technical Conference and Exhibition, 2011. 4699-704

14. López JC, Zidan A. Step change in driving performance: A case study. In: SPE Latin American and Caribbean Health / Safety / Environment / Social Responsibility Conference 2013: Sustainable Solutions for Challenging HSSE Environments in Latin America and the Caribbean, 2013. 7-16
15. Mohamed E, Jafari P, Kang SC, Pereira E, AbouRizk S. Leading indicators for safety management: Understanding the impact of project performance data on safety performance. In: Proceedings, Annual Conference - Canadian Society for Civil Engineering, 2019.
16. Pereira E, Guo X, Soleimanifar M, Siu MFF, Jeddry V, AbouRizk S. Effectiveness of applying a behavior-based safety program in industrial modular construction. In: 6th CSCE-CRC International Construction Specialty Conference 2017 - Held as Part of the Canadian Society for Civil Engineering Annual Conference and General Meeting 2017, 2017. 545-53
17. Schiavi AR. Using safety audits as a leading indicator. In: ASSE Professional Development Conference and Exposition 2013, 2013.
18. Stough J. From research to practice: A story of an actionable safety leading indicator index. In: SPE Latin American and Caribbean Health / Safety / Environment / Social Responsibility Conference 2013: Sustainable Solutions for Challenging HSSE Environments in Latin America and the Caribbean, 2012. 17-28
19. Tauseef A, Villegas M, Bordage P, Turner L. Behavior management: A successful approach. In: Society of Petroleum Engineers - SPE/APPEA Int. Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2012: Protecting People and the Environment - Evolving Challenges, 2012. 640-46
20. Thananan T. Target zero: The challenge of achieving sustainable safety excellence. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment 2014: The Journey Continues, 2014. 1502-10
21. Wei L, Yang R, Chen Y, Shahi A, Safa M, Hanna A, McCabe B. Comparison of safety cultures and performances between the construction industries in the United States and Canada: A case study of Texas and Ontario. In: Construction Research Congress 2020: Safety, Workforce, and Education - Selected Papers from the Construction Research Congress 2020, 2020. 346-55
22. Breitsprecher K, Ndahbros S, Hinton JJ, Jacques P. Organizational ownership of an interdependent HSE culture yields quantifiable HSE performance. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment 2014: The Journey Continues, 2014. 1834-41
23. Campbell Institute. Practical guide to leading indicators: Metrics, case studies and strategies. 2015. Available from: https://www.thecampbellinstitute.org/wp-content/uploads/2017/05/Campbell-Institute-Practical-Guide-Leading-Indicators-WP.pdf.
24. Choe S, Yun S, Leite F. Analysis of the effectiveness of the OSHA steel erection standard in the construction industry. Safety Science. 2016.89:190-200. doi: https://doi.org/10.1016/j.ssci.2016.06.016
25. Bitar FK, Chadwick-Jones D, Lawrie M, Nazaruk M, Boodhai C. Empirical validation of operating discipline as a leading indicator of safety outputs and plant performance. Safety Science. 2018.104:144-56. doi: https://doi.org/10.1016/j.ssci.2017.12.036
26. Tang DKH, Leiliabadi F, Olugu EU, Md Dawal SZB. Factors affecting safety of processes in the Malaysian oil and gas industry. Safety Science. 2017.92:44-52. doi: https://doi.org/10.1016/j.ssci.2016.09.017

27. Brandt M, Andersen LL, Kines P, Ajslev JZN. Safety climate at work and risk of long-term sickness absence: Prospective cohort with register follow-up among 63,500 workers. Safety Science. 2023.166doi: https://doi.org/10.1016/j.ssci.2023.106217

28. Amir-Heidari P, Maknoon R, Taheri B, Bazyari M. A new framework for HSE performance measurement and monitoring. Safety Science. 2017.100:157-67. doi: 10.1016/j.ssci.2016.11.001

29. Laitinen H, Päivärinta K. A new-generation safety contest in the construction industry - A long-term evaluation of a real-life intervention. Safety Science. 2010.48(5):680-86. doi: https://doi.org/10.1016/j.ssci.2010.01.018

30. Laitinen H, Vuorinen M, Simola A, Yrjänheikki E. Observation-based proactive OHS outcome indicators - Validity of the Elmeri+ method. Safety Science. 2013.54:69-79. doi: https://doi.org/10.1016/j.ssci.2012.11.005

31. Moore LL, Wurzelbacher SJ, Chen IC, Lampl MP, Naber SJ. Reliability and validity of an employer-completed safety hazard and management assessment questionnaire. Journal of Safety Research. 2022.81:283-96. doi: https://doi.org/10.1016/j.jsr.2022.03.005

32. Chen Y, McCabe B, Hyatt D. Impact of individual resilience and safety climate on safety performance and psychological stress of construction workers: A case study of the Ontario construction industry. Journal of Safety Research. 2017.61:167-76. doi: https://doi.org/10.1016/j.jsr.2017.02.014

33. Robson LS, Ibrahim S, Hogg-Johnson S, Steenstra IA, Van Eerd D, Amick BC. Developing leading indicators from OHS management audit data: Determining the measurement properties of audit data from the field. Journal of Safety Research. 2017.61:93-103. doi: https://doi.org/10.1016/j.jsr.2017.02.008

34. Winge S, Albrechtsen E, Arnesen J. A comparative analysis of safety management and safety performance in twelve construction projects. Journal of Safety Research. 2019.71:139-52. doi: https://doi.org/10.1016/j.jsr.2019.09.015

35. Tang DKH, Md Dawal SZ, Olugu EU. Actual safety performance of the Malaysian offshore oil platforms: Correlations between the leading and lagging indicators. Journal of Safety Research. 2018.66:9-19. doi: https://doi.org/10.1016/j.jsr.2018.05.003

36. Wachter JK, Yorio PL. A system of safety management practices and worker engagement for reducing and preventing accidents: An empirical and theoretical investigation. Accid Anal Prev. 2014.68:117-30. doi: https://doi.org/10.1016/j.aap.2013.07.029

37. Sheehan C, Donohue R, Shea T, Cooper B, De Cieri H. Leading and lagging indicators of occupational health and safety: The moderating role of safety leadership. Accid Anal Prev. 2016.92:130-38. doi: https://doi.org/10.1016/j.aap.2016.03.018

38. Alarcón L, Acuña D, Diethelm S, Pellicer E. Strategies for improving safety performance on construction firms. Accid Anal Prev. 2016.94:107-18. doi: https://doi.org/10.1016/j.aap.2016.05.021

39. Hinze J, Hallowell M, Baud K. Construction-safety best practices and relationships to safety performance. Journal of Construction Engineering and Management. 2013.139(10)doi: http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000751

40. Salas R, Hallowell M. Predictive validity of safety leading indicators: Empirical assessment in the oil and gas sector. Journal of Construction Engineering and Management. 2016.142(10)doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001167

41. Coetzee E, Govender U, Ndeunyema P, Genc B, Maré Y, Roux J, *et al.* An integrated safety framework for the diamond mines: A case study from Namibia. Resources Policy. 2023.82doi: https://doi.org/10.1016/j.resourpol.2023.103564

42. Govender U, van Eck G, Genc B. An integrated 4Cs safety framework for the diamond industry of Southern Africa. Resources Policy. 2022.77doi: https://doi.org/10.1016/j.resourpol.2022.102774

43. Lagerstrom E, Magzamen S, Kines P, Brazile W, Rosecrance J. Determinants of safety climate in the professional logging industry. Safety. 2019.5(2)doi: https://doi.org/10.3390/safety5020035

44. Mousavi SS, Khani Jazani R, Cudney EA, Trucco P. Quantifying the relationship between lean maturity and occupational health and safety: Antecedents and leading indicators. International Journal of Lean Six Sigma. 2020.11(1):150-70. doi: https://doi.org/10.1108/IJLSS-04-2018-0043

45. Vosoughi S, Chalak MH, Yarahmadi R, Abolghasemi J, Alimohammadi I, Anbardan AN, Kanrash FA. Prioritization and assessment of safety key performance indicators in an automotive industry. Sigurnost. 2021.63(4):347-61. doi: https://doi.org/10.31306/s.63.4.1

46. Zahoor H, Chan APC, Utama WP, Gao R, Zafar I. Modeling the relationship between safety climate and safety performance in a developing construction industry: A cross-cultural validation study. IJERGQ. 2017.14(4)doi: https://doi.org/10.3390%2Fijerph14040351

47. Van Derlyke P, Marín LS, Zreiqat M. Discrepancies between implementation and perceived effectiveness of leading safety indicators in the US dairy product manufacturing industry. Safety and Health at Work. 2022.13(3):343-49. doi: https://doi.org/10.1016%2Fj.shaw.2022.04.004

48. Dadashi Haji M, Behnam B, Sebt MH, Ardeshir A, Katooziani A. BIM-based safety leading indicators measurement tool for construction sites. International Journal of Civil Engineering. 2023.21(2):265-82. doi: https://doi.org/10.1007/s40999-022-00754-9

49. Haas EJ, Connor BP, Vendetti J, Heiser R. A case study exploring field level risk assessments as a leading safety indicator. Trans Soc Min Metall Explor. 2018.342:22-28. doi: https://doi.org/10.19150/trans.8104

50. Grabowski M, You Z, Song H, Wang H, Merrick JRW. Sailing on friday: Developing the link between safety culture and performance in safety-critical systems. IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans. 2010.40(2):263-84. doi: https://doi.org/10.1109/TSMCA.2009.2035300

51. Cao HC, Goh YM. Analyzing construction safety through time series methods. Front. Eng. Manag. 2019.6(2):262-74. doi: https://doi.org/10.1007/s42524-019-0015-6

52. Dennerlein JT, Weinstein D, Huynh W, Tessler J, Bigger L, Murphy L, Manjourides J. Associations between a safety prequalification survey and worker safety experiences on commercial construction sites. Am J Ind Med. 2020.63(9):766-73. doi: https://doi.org/10.1002/ajim.23143

53. Manjourides J, Dennerlein J. Testing the associations between leading and lagging indicators in a contractor safety pre-qualification database. National Occupational Injury Research Symposium. 2019.84.

54. Merrick JRW, Grabowski M. Decision performance and safety performance: A value-focused thinking study in the oil industry. Decision Analysis. 2014.11(2):105-16. doi: http://dx.doi.org/10.1287/deca.2014.0291

55. Quaigrain RA, Issa MH. Comparative analysis of leading and lagging indicators of construction disability management performance: An exploratory study. International Journal of Construction Management. 2023.23(7):1205-13. doi: https://doi.org/10.1080/15623599.2021.1963921

56. Rajendran S. Enhancing construction worker safety performance using leading indicators. Practice Periodical on Structural Design and Construction. 2013.18(1):45-51. doi: https://doi.org/10.1061/(asce)sc.1943-5576.0000137

57. World Bank. New World Bank country classifications by income level: 2022-2023 2022. Available from: https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2022-2023.

58. McGuinness L, Higgins J. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Research Synthesis Methods. 2020.1-2. doi: https://doi.org/10.1002/jrsm.1411

59. Alruqi WM, Hallowell MR. Critical Success Factors for Construction Safety: Review and Meta-Analysis of Safety Leading Indicators. Journal of Construction Engineering and Management. 2019.145(3)doi: 10.1061/(asce)co.1943-7862.0001626

60. Hallowell M, Quashne M, Salas R, Jones M, MacLean B, Quinn E. The statistical invalidity of TRIR as a measure of safety performance. Construction Safety Research Alliance; 2020.

61. Pearce N, Checkoway H, Kriebel D. Bias in occupational epidemiology studies. Occup Environ Med. 2007.64(8):562-68. doi: 10.1136/oem.2006.026690

62. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, M P. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 8008.337:A1655. doi: https://doi.org/10.1136/bmj.a1655

63. Skivington K, Matthews L, Simpson S A, Craig P, Baird J, M BJ*, et al.* A new framework for developing and evaluating complex interventions: Update of Medical Research Council guidance. BMJ. 2021.374:n2061. doi: https://doi.org/10.1136/bmj.n2061

64. Clarivate. EndNote 20. Clarivate; 2021. Available from: http://endnote.com/.

65. Xu J, Cheung C, Manu P, Ejohwomu O. Safety leading indicators in construction: A systematic review. Safety Science. 2021.139doi: https://doi.org/10.1016/j.ssci.2021.105250

Appendix A: Review Methods

A.1 Eligibility Criteria

To ensure that relevant studies were consistently identified, a clear definition of the eligible study participants, concept and context was developed.

A.1.1 Population

Studies of any company comprising workers of any age were eligible for inclusion in this review. Workers included employees and contractors (including people providing casual labour).

Studies including both workers and non-workers (for example, members of the public) were only eligible if data were reported separately for the workers.

A.1.2 Concept

The concept of interest refers to the focus of the scoping review and includes elements from standard systematic review eligibility criteria such as “interventions” and “outcomes”. Studies of any leading indicator(s) were eligible for inclusion in this review.

During discussions with OSH experts to identify a suitable review question, all stakeholders (YHEC, the Foundation and OSH experts) revisited the definition of what constitutes a leading indicator, resulting in the following definition.

A factor linked to the practice of safety and health within a workplace, which is intended to prevent future accidents or other adverse safety outcomes, or otherwise improve workplace safety. These factors include:

- The active monitoring of existing safety and health practices and performance, including monitoring for known hazards.
- The implementation of a new intervention that may impact the existing safety management system (with the intervention not limited to safety, but also including managerial and organisational interventions).
- Measures used to identify or monitor the effectiveness of a relevant intervention.

These factors do not include:

- Reactive monitoring.
- Establishing the cause of safety outcomes that have already occurred, such as a retrospective investigation seeking to identify the causes of accidents. Studies that perform this retrospective investigation would only be eligible if the act of retrospective investigation were used as an intervention itself; for example, if it is an established part of the safety and health management system that is used to prevent future accidents.

Leading indicators were therefore considered to include (but not be limited to):

- Safety audits.
- Behaviour based safety.
- Visible leadership.
- Safety training.
- Other staff training, such as leadership skills.
- Corrective action measurements.
- Organisational, scheduling and timing related leading indicators.

The literature search was specifically targeted to identify studies of leading indicators in general and did not include specific examples of leading indicators.

Lagging indicators were considered to be any safety outcome, including (but not limited to):

- Injury.
- Illness.
- Fatality.

Studies reporting occupational health (for example, wellbeing) or performance were not eligible. Studies reporting only qualitative outcomes were also not eligible.

A.1.3 Context

Studies conducted in any industry were eligible for inclusion. Studies in settings that were not occupational were not eligible.

Studies of participants in a mixture of occupational and non-occupational settings were eligible for inclusion if data are reported separately for participants in occupational settings.

A.1.4 Study design

Any study design was eligible if eligible outcomes were reported in ten or more workers.

Reviews were not eligible. However, if any relevant systematic reviews published in the last five years were identified in the searches, then the reviewers checked their included studies lists for eligible studies that may have been missed by the database searches. The systematic reviews were not extracted.

A.1.5 Limits

Studies published before 2010 were not eligible for inclusion. Studies published only as conference abstracts were eligible for inclusion.

Non-English language studies were excluded, as were pre-prints, editorials and news items.

A.2 Identification of Relevant Studies

A.2.1 Search strategy

A search strategy for the Scopus database was designed to identify studies on the impact of leading indicators on safety outcomes in the workplace. The search strategy design reflected the pragmatic scoping review context. The search methods were not designed to be exhaustive. They were designed to target a selection of potentially relevant studies, whilst enabling searches to be conducted and results assessed within the context of the project resource and timeline. The final Scopus strategy is presented in Figure A.1.

The strategy comprised three concepts:

- Leading indicators (search lines 1 to 6).
- Safety outcomes (search lines 7 to 10).
- Workplace setting (search line 11).

The concepts were combined as follows: leading indicators AND safety outcomes AND workplace setting.

The strategy was devised using a combination of terms in the Title, Abstract and Keyword fields. The search terms for each concept were identified through discussion within the research team, scanning background literature and browsing database thesauri. Reflecting the pragmatic review context, the terms for each concept were relatively restricted. This approach was discussed and agreed within the research team and with Lloyd's Register Foundation. This approach meant, for example, that the terms for the leading indicators concept were designed to target studies where the database record explicitly refer to generic, non-specific terms for the concept. The strategy was not designed to retrieve studies where the record only referred to specific, named indicators. We used supplementary search approaches to mitigate the possibility of missing important papers (see Table A.1).

Reflecting the eligibility criteria, the strategy was restricted to studies published in English language (search line 14). The strategy was not restricted by date.

The final Scopus strategy was peer-reviewed before execution by a second Information Specialist. Peer review considered the appropriateness of the strategy for the review scope and eligibility criteria, inclusion of key search terms, errors in spelling, syntax and line combinations, and application of exclusions.

Figure A.1: Search strategy for Scopus

#1. TITLE({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) OR KEY({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) 1336
#2. TITLE-ABS-KEY({lagging indicator} OR {proactive indicator} OR {pro-active indicator} OR {pro active indicator} OR {predictive indicator} OR {upstream indicator} OR {heading indicator} OR {positive indicator} OR {process indicator} OR {activities indicator} OR {downstream indicator} OR {historical indicator} OR {trailing indicator} OR {negative indicator} OR {safety indicator} OR {lagging indicators} OR {proactive indicators} OR {pro-active indicators} OR {pro active indicators} OR {predictive indicators} OR {upstream indicators} OR {heading indicators} OR {positive indicators} OR {process indicators} OR {activities indicators} OR {downstream indicators} OR {historical indicators} OR {trailing indicators} OR {negative indicators} OR {safety indicators}) 7543
#3. ABS({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) 2548
#4. TITLE-ABS-KEY({lead measure} OR {lead measures} OR {leading measure} OR {leading measures} OR {proactive measure} OR {pro-active measure} OR {pro active measure} OR {predictive measure} OR {upstream measure} OR {proactive measures} OR {pro-active measures} OR {pro active measures} OR {predictive measures} OR {upstream measures}) 2338
#5. TITLE-ABS-KEY({lead metric} OR {lead metrics} OR {leading metric} OR {leading metrics} OR {proactive metric} OR {pro-active metric} OR {pro active metric} OR {predictive metric} OR {upstream metric} OR {proactive metrics} OR {pro-active metrics} OR {pro active metrics} OR {predictive metrics} OR {upstream metrics}) 234
#6. #2 OR #3 OR #4 OR #5 12412
#7. TITLE-ABS-KEY({OHS} OR {OSH} OR {ESH} OR {OH&S} OR {OS&H} OR {ES&H} OR {EH&S}) 8426
#8. TITLE-ABS-KEY(safety OR safe OR "near miss" OR "near misses" OR benchmarking OR benchmark OR "bench mark" OR "bench marking") 2956706
#9. TITLE-ABS-KEY(hazard* OR accident* OR incident* OR fatalities OR fatality OR illness* OR death OR deaths OR mortality OR injure OR injured OR injuries OR injury OR prevent*) 9241586
#10. #7 or #8 or #9 11373218
#11. TITLE-ABS-KEY(industry OR industrial OR industries OR workplace or workplaces OR work OR worker OR workers OR staff OR employee OR employees OR personnel OR construction OR occupation* OR organisation* OR organization* OR companies OR company OR manufacture OR manufacturing OR construction OR mining OR agricultur* OR transport* OR haulage OR contractor OR contractors OR labour OR labor OR workforce OR {human resource} OR {human resources}) 18093469
#12. #6 AND #10 AND #11 2709
#13. #1 OR #12 3709
#14. Limit to English language 3518

Key to Scopus symbols and commands:

*	Unlimited right-hand truncation symbol
" "	Loose phrase search
{ }	Exact phrase search
TITLE-ABS-KEY	Searches are restricted to the Title, Abstract and Keyword fields.

A.2.2 Resources searched

We conducted the literature search in the databases and information sources shown in Table A.1. The selection of resources reflected the pragmatic review context.

Table A.1: Databases and information sources searched

Resource	Interface / URL
Databases	
Scopus	Elsevier
Web of Science Core Collection: ▪ Science Citation Index Expanded. ▪ Social Sciences Citation Index. ▪ Arts & Humanities Citation Index. ▪ Conference Proceedings Citation Index – Science. ▪ Conference Proceedings Citation Index - Social Science & Humanities. ▪ Emerging Sources Citation Index.	Clarivate
The NIOSH NIOSHTIC-2 publications search	https://www2a.cdc.gov/nioshtic-2/default.asp
Reference list checking	NA
Expert input (EAG)	NA

Key: EAG – expert advisory group; NA – not applicable; NIOSH - National Institute for Occupational Safety and Health.

We requested relevant unpublished reports from members of the EAG.

For reference list checking, we checked the included studies list of any retrieved relevant systematic reviews published in the last five years for eligible studies that may have been missed by the database searches. Only reviews assessing quantitative performance data were eligible for checking.

A.2.3 Running the search strategy and downloading the results

We conducted the searches on 1 August 2023 using each database or resource listed in the protocol, translating the agreed Scopus strategy appropriately. Translation included consideration of differences in database interfaces and functionality, in addition to variation in indexing languages and thesauri. The final translated database strategies were peer-reviewed by a second Information Specialist. Peer review considered the appropriateness of the translation for the database being searched, errors in syntax and line combinations, and application of exclusions.

Appendix B contains the full strategies for all sources searched.

Where possible, we downloaded the results of searches in a tagged format and loaded them into bibliographic management software (EndNote) [64]. The results were deduplicated using several algorithms and the deduplicated references were held in a duplicates EndNote database for checking if required. Results from resources which did not allow export in a format compatible with EndNote were saved in Word or Excel documents as appropriate and manually deduplicated.

A.3 Selection of Relevant Studies

Record assessment was undertaken as follows:

- A single researcher assessed the search results and removed the obviously irrelevant records such as those in ineligible settings.
- Two reviewers independently assessed the titles and abstracts of 10% of the remaining records for relevance against the eligibility criteria, with disagreements adjudicated by a third reviewer. A single reviewer assessed the relevance of the remaining records.
- We obtained the full text of potentially relevant studies. Two reviewers independently assessed 10% of the full texts for relevance against the eligibility criteria. A third reviewer adjudicated any disagreements. A single reviewer screened the remaining full texts.
- We recorded the number of records included and removed at each selection stage in the PRISMA flow diagram. We listed studies excluded after assessment of the full document in an excluded studies table, with the reasons for exclusion.

We obtained electronic or paper copies of potentially relevant full documents meeting the systematic review's eligibility criteria in liaison with the Foundation, the EAG or via local access routes.

A.4 Data Charting

A data extraction template was developed in Excel and piloted in duplicate on 10% of included studies before progressing to full data extraction.

One researcher extracted data from the remaining included studies, and a second researcher checked all data points for another 10% of included studies. In total therefore, the extractions of 20% of the included studies were checked by a second reviewer.

This scoping review extracted the following elements from the eligible studies:

- Study details (bibliographic details).
- Study characteristics:
    - Country:
        - World Bank income level classification [57] of the country (HI, UMI, LMI, LI).
    - Study design and methods.
    - Date of study and data collection.
    - Details about statistical analyses
- Participant details:
    - Including reported characteristics of companies and workers.

- Leading indicators:
    - Author-reported definition and description of leading indicators.
    - Method of measurement.
    - Level of leading indicator evaluated (individual level, project level, company/ institutional level).
    - Author-reported definition in relation to literature definitions:
        - Xu et al 2021 [65]: “measures that indicate the current performance of a safety management system of a project or firm. They can: 1) identify the system’s weaknesses and strengths, 2) identify situations that might cause incidents and injuries, 3) drive proactive actions to prevent an incident or injury before it occurs and achieve continuous improvement".
        - The Campbell Institute [23]: “proactive, preventative and predictive measures that monitor and provide current information about the effective performance, activities, and processes of an environment, health and safety (EHS) management system that drive the identification and elimination or control of risks in the workplace that can cause incidents and injuries”.
- Lagging indicators:
    - Author reported definition and description of lagging indicator.
    - Method of measurement.
    - Unit of measurement.
    - Number of workers analysed.
    - The size of the effect.
    - A measure of precision for each estimate of effect (95% confidence intervals, standard error [SE] or standard deviation [SD]).
    - Data at all reported time points.
- Method of evaluating impact of leading indicator on lagging indicator (e.g., correlation, multivariate modelling, simple comparison, development of a theoretical model with a case study).
- Measures of effectiveness of leading indicators other than lagging indicators.
- Applicability / generalisability of study findings (this was a subjective assessment regarding the extent to which study results were applicable beyond the remit of the study, for example to the whole industry, or to other countries).
- Stakeholders involved.
- Outcomes quantifying the relationship between a leading and lagging indicator.

A.5 Risk of Bias Assessment

One reviewer appraised the internal validity of each included study, and a second reviewer checked 20% of these assessments. This was performed using a simple tool designed by Robson and coauthors for reviews of effectiveness in OSH published as part of a systematic review of OSH management systems [9].

The assessment consists of five questions designed to probe the possibility of:

- Selection bias (failing to guard against the selection or elimination of study cases that are systematically more or less likely to demonstrate a correlation between the evaluated leading and lagging indicators).
- Confounding bias (failing to control for confounding factors that might explain a correlation or effect, outside of the relationship between leading and lagging indicator).
- Outcome measurement bias (measuring outcomes using methods that may lead to either over or underestimate the effect of the leading indicators).
- Reliability of statistical methods (failure to conduct statistical analysis, or the use of inappropriate statistical tests given the data collected, potentially leading to the misinterpretation of results).
- Any other possible study design artefact resulting in a systematic deviation of results from the truth.

The five questions are listed in Appendix D with detailed assessments of each included study. Each study was assigned an overall grading using the following rationale:

- High RoB (serious limitations): the answer to at least four of the five questions was ‘no’ or ‘unclear’.
- Moderate RoB (moderate limitations): the answer to two or three questions was ‘no’ or ‘unclear’.
- Low RoB (minor limitations): the answer to one question was ‘no’ or ‘unclear’.
- Very low RoB (no limitations): all questions received positive ‘yes’ answers indicating an absence of perceived risk of bias.

RoB plots were generated using the Robvis shiny web app [58].

Appendix B: Search Strategies

B.1: Source: Scopus

Interface / URL: Elsevier

Database coverage dates: Information not found. Possibly 1823 to present (https://service.elsevier.com/app/answers/detail/a_id/11274/c/10547/supporthub/scopus/)

Search date: 1 August 2023

Retrieved records: 3518

Search strategy:

Two versions of the strategy are shown below. Firstly, the search strategy is shown line by line as entered into the Scopus interface. This version shows the number of results retrieved for each line and is presented for readability. Secondly, in order to meet PRISMA-S guidance, the strategy is presented as a single line search showing the same strategy "as run" in the database.

#1. TITLE({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) OR KEY({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) 1,336

#2. TITLE-ABS-KEY({lagging indicator} OR {proactive indicator} OR {pro-active indicator} OR {pro active indicator} OR {predictive indicator} OR {upstream indicator} OR {heading indicator} OR {positive indicator} OR {process indicator} OR {activities indicator} OR {downstream indicator} OR {historical indicator} OR {trailing indicator} OR {negative indicator} OR {safety indicator} OR {lagging indicators} OR {proactive indicators} OR {pro-active indicators} OR {pro active indicators} OR {predictive indicators} OR {upstream indicators} OR {heading indicators} OR {positive indicators} OR {process indicators} OR {activities indicators} OR {downstream indicators} OR {historical indicators} OR {trailing indicators} OR {negative indicators} OR {safety indicators}) 7,543

#3. ABS({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) 2,548

#4. TITLE-ABS-KEY({lead measure} OR {lead measures} OR {leading measure} OR {leading measures} OR {proactive measure} OR {pro-active measure} OR {pro active measure} OR {predictive measure} OR {upstream measure} OR {proactive measures} OR {pro-active measures} OR {pro active measures} OR {predictive measures} OR {upstream measures}) 2,338

#5. TITLE-ABS-KEY({lead metric} OR {lead metrics} OR {leading metric} OR {leading metrics} OR {proactive metric} OR {pro-active metric} OR {pro active metric} OR {predictive metric} OR {upstream metric} OR {proactive metrics} OR {pro-active metrics} OR {pro active metrics} OR {predictive metrics} OR {upstream metrics}) 234

#6. #2 OR #3 OR #4 OR #5 12,412

#7. TITLE-ABS-KEY({OHS} OR {OSH} OR {ESH} OR {OH&S} OR {OS&H} OR {ES&H} OR {EH&S}) 8,426

#8. TITLE-ABS-KEY(safety OR safe OR "near miss" OR "near misses" OR benchmarking OR benchmark OR "bench mark" OR "bench marking") 2,956,706

#9. TITLE-ABS-KEY(hazard* OR accident* OR incident* OR fatalities OR fatality OR illness* OR death OR deaths OR mortality OR injure OR injured OR injuries OR injury OR prevent*) 9,241,586

#10. #7 or #8 or #9 11,373,218

#11. TITLE-ABS-KEY(industry OR industrial OR industries OR workplace or workplaces OR work OR worker OR workers OR staff OR employee OR employees OR personnel OR construction OR occupation* OR organisation* OR organization* OR companies OR company OR manufacture OR manufacturing OR construction OR mining OR agricultur* OR transport* OR haulage OR contractor OR contractors OR labour OR labor OR workforce OR {human resource} OR {human resources}) 18,093,471

#12. #6 AND #10 AND #11 2,709

#13. #1 OR #12 3,709

#14. Limit to English language 3,518

Single line strategy copied directly after running in Scopus:

(TITLE ({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators}) OR KEY ({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators})) OR (((TITLE-ABS-KEY ({lagging indicator} OR {proactive indicator} OR {pro-active indicator} OR {pro active indicator} OR {predictive indicator} OR {upstream indicator} OR {heading indicator} OR {positive indicator} OR {process indicator} OR {activities indicator} OR {downstream indicator} OR {historical indicator} OR {trailing indicator} OR {negative indicator} OR {safety indicator} OR {lagging indicators} OR {proactive indicators} OR {pro-active indicators} OR {pro active indicators} OR {predictive indicators} OR {upstream indicators} OR {heading indicators} OR {positive indicators} OR {process indicators} OR {activities indicators} OR {downstream indicators} OR {historical indicators} OR {trailing indicators} OR {negative indicators} OR {safety indicators})) OR (ABS ({lead indicator} OR {leading indicator} OR {lead indicators} OR {leading indicators})) OR (TITLE-ABS-KEY ({lead measure} OR {lead measures} OR {leading measure} OR {leading measures} OR {proactive measure} OR {pro-active measure} OR {pro active measure} OR {predictive measure} OR {upstream measure} OR {proactive measures} OR {pro-active measures} OR {pro active measures} OR {predictive measures} OR {upstream measures})) OR (TITLE-ABS-KEY ({lead metric} OR {lead metrics} OR {leading metric} OR {leading metrics} OR {proactive metric} OR {pro-active metric} OR {pro active metric} OR {predictive metric} OR {upstream metric} OR {proactive metrics} OR {pro-active metrics} OR {pro active metrics} OR {predictive metrics} OR {upstream metrics}))) AND ((TITLE-ABS-KEY ({OHS} OR {OSH} OR {ESH} OR {OH&S} OR {OS&H} OR {ES&H} OR {EH&S})) OR (TITLE-ABS-KEY (safety OR safe OR "near miss" OR "near misses" OR benchmarking OR benchmark OR "bench mark" OR "bench marking")) OR (TITLE-ABS-KEY (hazard* OR accident* OR incident* OR fatalities OR fatality OR illness* OR death OR deaths OR mortality OR injure OR injured OR injuries OR injury OR prevent*))) AND (TITLE-ABS-KEY (industry OR industrial OR industries OR workplace OR workplaces OR work OR worker OR workers OR staff OR employee OR employees OR personnel OR construction OR occupation* OR organisation* OR organization* OR companies OR company OR manufacture OR manufacturing OR construction OR mining OR agricultur* OR transport* OR haulage OR contractor OR contractors OR labour OR labor OR workforce OR {human resource} OR {human resources}))) AND (LIMIT-TO (LANGUAGE , "English"))

B.2: Source: Web of Science Core Collection (see below)

Interface / URL: Clarivate

Database coverage dates: see below

Search date: 1 Aug 2023

Retrieved records: 2512

Search strategy:

The Web of Science Core Collection on the date of the search consisted of the following individual databases:

- Science Citation Index Expanded (SCI-EXPANDED)--1900-present.
- Social Sciences Citation Index (SSCI)--1956-present.
- Arts & Humanities Citation Index (AHCI)--1975-present.
- Conference Proceedings Citation Index – Science (CPCI-S)--1990-present.
- Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH)--1990-present.
- Emerging Sources Citation Index (ESCI)--2015-present.

1 TI=("lead* indicator*") OR AK=("lead* indicator*") OR KP=("lead* indicator*") 1,112

2 TS=("lagging indicator*" OR "proactive indicator*" OR "pro-active indicator*" OR "predictive indicator*" OR "upstream indicator*" OR "heading indicator*" OR "positive indicator*" OR "process indicator*" OR "activities indicator*" OR "downstream indicator*" OR "historical indicator*" OR "trailing indicator*" OR "negative indicator*" OR "safety indicator*") 7,040

3 AB=("lead* indicator*") 1,654

4 TS=("lead* measure*" OR "proactive measure*" OR "pro-active measure*" OR "predictive measure*" OR "upstream measure*") 2,431

5 TS=("lead* metric*" OR "proactive metric*" OR "pro-active metric*" OR "predictive metric*" OR "upstream metric*") 180

6 #2 OR #3 OR #4 OR #5 11,176

7 TS=(OHS OR OSH OR ESH OR "OH&S" OR "OS&H" OR "ES&H") 6,726

8 TS=(safety OR safe OR "near miss" OR "near misses" OR benchmarking OR benchmark* OR "bench mark*" OR "bench marking") 2,013,521

9 TS=(hazard* OR accident* OR incident* OR fatalities OR fatality OR illness* OR death OR deaths OR mortality OR injure OR injured OR injuries OR injury OR prevent*) 6,109,114

10 #7 OR #8 OR #9 7,663,691

11 TS=(industry OR industrial OR industries OR workplace or workplaces OR work OR worker OR workers OR staff OR employee OR employees OR personnel OR construction OR occupation* OR organisation* OR organization* OR companies OR company OR manufacture OR manufacturing OR construction OR mining OR agricultur* OR transport* OR haulage OR contractor OR contractors OR labour OR labor OR workforce OR "human resource" OR "human resources") 11,809,313

12 #6 AND #10 AND #11 1,603

13	#12 OR #1	2,590
14	#12 OR #1 and English (Languages)	2,512

B.3: Source: NIOSHTIC-2 Publications Search

Interface / URL: https://www2a.cdc.gov/nioshtic-2/default.asp

Database coverage dates: From the site: A significant portion of the citations (39,000) date from 1971 to the present. An additional 13,800 resources in NIOSHTIC-2 are publications dating from the 1930's to the present from the NIOSH Mining Safety & Health Research Laboratories (formerly the U. S. Bureau of Mines). https://www2a.cdc.gov/nioshtic-2/n2info.asp

Search date: 1 Aug 2023

Retrieved records: 117

Search strategy:

The following search terms were entered in the Basic search box at the URL above. Given the relatively small result numbers it was decided not to combine these terms with any of the terms for safety outcomes or the workplace setting.

'lead* indicator*' or 'lead* metric*' or 'proactive metric*' or 'pro-active metric*' or 'pro active metric*' or 'predictive metric*' or 'upstream metric*' or 'lagging indicator*' or 'proactive indicator*' or 'pro-active indicator*' or 'pro active indicator*' or 'predictive indicator*' or 'upstream indicator*' or 'heading indicator*' or 'positive indicator*' or 'process indicator*' or 'activities indicator*' or 'downstream indicator*' or 'historical indicator*' or 'trailing indicator*' or 'negative indicator*' or 'safety indicator*' or 'lead* measure*' or 'proactive measure*' or 'pro-active measure*' or 'pro active measure*' or 'predictive measure*' or 'upstream measure*'

=117 results

Appendix C: Identification of Studies

C.1 Literature Search Results

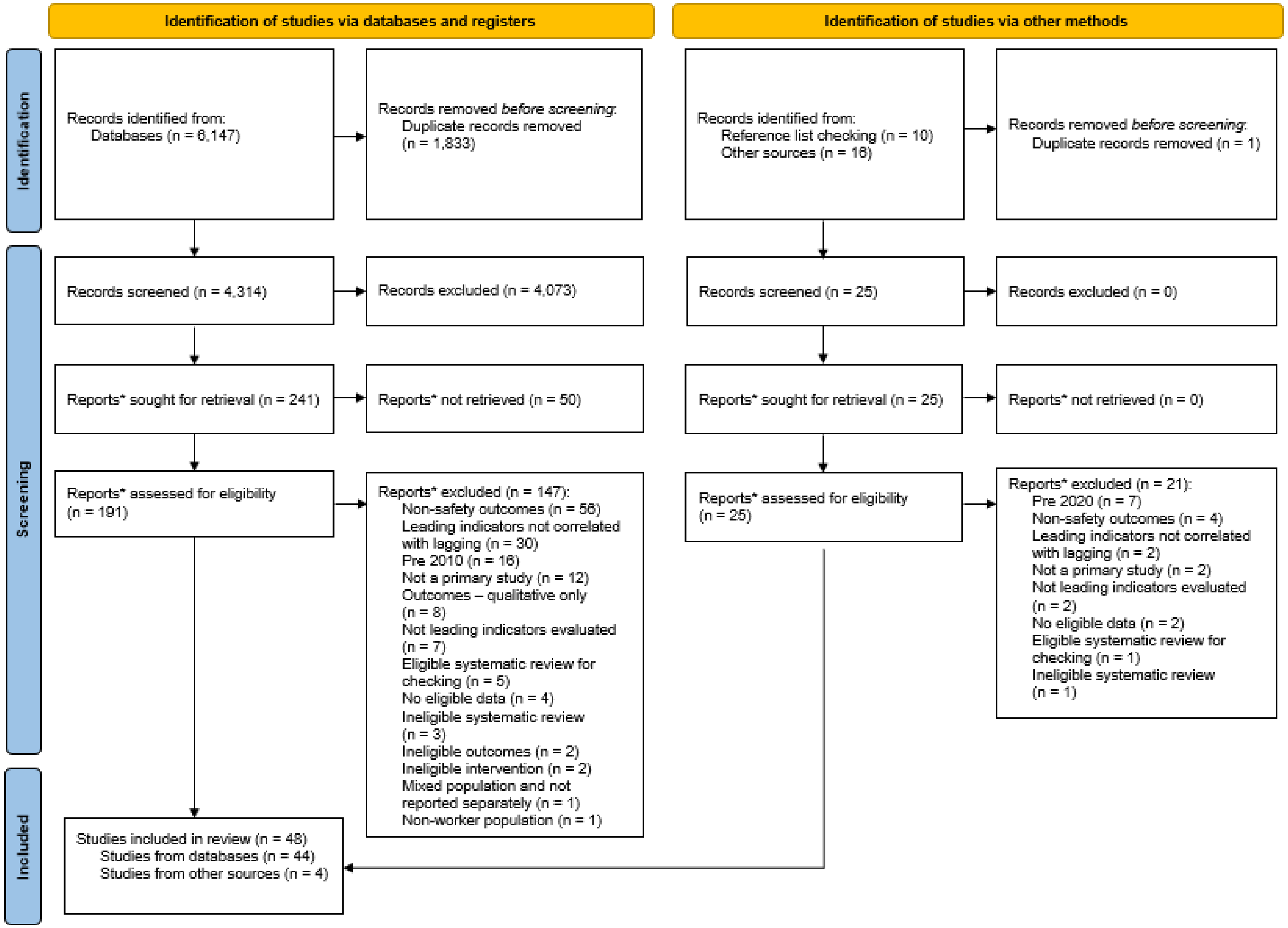
The searches were conducted on 1 August 2023 and identified 6,163 records (Table C.1). Following deduplication, 4,339 records were assessed for relevance.

Table C.1: Literature search results

Resource	Number of records identified
Databases	
Scopus	3,518
Web of Science Core Collection: ▪ Science Citation Index Expanded. ▪ Social Sciences Citation Index. ▪ Arts & Humanities Citation Index. ▪ Conference Proceedings Citation Index - Science. ▪ Conference Proceedings Citation Index - Social Science & Humanities. ▪ Emerging Sources Citation Index.	2,512
The NIOSH NIOSHTIC-2 publications search	117
Total records identified through database searching	6,147
Reference list checking	10
Expert input (EAG)	16
Total additional records identified through other sources	26
Total number of records retrieved	**6,173**
Total number of records after deduplication	**4,339**

Abbreviations: EAG – Expert advisory group, NIOSH - National Institute for Occupational Safety and Health.

Figure C.1: PRISMA flow diagram



*"Note that a "report" could be a journal article, preprint, conference abstract, study register entry, clinical study report, dissertation, unpublished manuscript, government report or any other document providing relevant information": https://www.bmj.com/content/372/bmj.n71.

Adapted from: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: http://www.prisma-statement.org/

C.2 Study Selection

Of the 4,339 records assessed for relevance, 4,073 records were excluded following first pass (n=1,145) and title and abstract review (n=2,928). The full texts for 266 records were sought for retrieval; 50 records were unobtainable, and 216 records were retrieved and assessed for relevance. Of these, 168 were determined to be ineligible, most commonly because they did not assess lagging indicators (n=60) or did report lagging indicators but did not evaluate whether these safety outcomes were associated with leading indicators (n=32). Twenty-three studies would have been eligible but for their date of publication, which preceded 2010. A full list of all 168 excluded studies is provided in Appendix F.

Appendix D: Excluded Studies

Table D.1: Excluded studies (n=168)

Reference	Exclusion reason
Aalberg A, Kvalheim SA, Nilsen IB, Bye RJ. Safety climate and work conditions related to acute spills and hydrocarbon leaks in the offshore oil and gas industry— A repeated cross-sectional study. In: Safety and Reliability - Safe Societies in a Changing World - Proceedings of the 28th International European Safety and Reliability Conference, ESREL 2018 2018; 53-62. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058103348&partnerID=40&md5=536dc3ba64abb009ed396b8da7a7043a	Ineligible outcomes
Abdul Razak N, Ejohwomu O, Fenn P, Okedara K, Dosumu B, Muhammad-Sukki F. Identification of health and safety prequalification criteria for contractor selection in construction projects: A systematic review. Energies. 2021.14(21):7244. doi: https://doi.org/10.3390/en14217244	Limit - eligible SR for checking
Abikenova S, Daumova G, Kurmanbayeva A, Yesbenbetova Z, Kazbekova D. Relationship between occupational risk and personal protective equipment on the example of ferroalloy production. International Journal of Safety and Security Engineering. 2022.12(5):609-14. doi: https://doi.org/10.18280/ijsse.120509	Leading indicators not correlated with lagging
Abubakar M, Zailani BM, Abdullahi M, Auwal AM. Potential of adopting a resilient safety culture toward improving the safety performance of construction organizations in Nigeria. J. Eng. Des. Technol. 2022.20(5):1236-56. doi: https://doi.org/10.1108/JEDT-09-2020-0354	Outcomes - not safety
Acheampong T, Kemp AG. Health, safety and environmental (HSE) regulation and outcomes in the offshore oil and gas industry: Performance review of trends in the United Kingdom Continental Shelf. Safety Science. 2022.148doi: https://doi.org/10.1016/j.ssci.2021.105634	Ineligible intervention
Adinyira E, Manu P, Agyekum K, Mahamadu AM, Olomolaiye PO. Violent behaviour on construction sites: Structural equation modelling of its impact on unsafe behaviour using partial least squares. Eng. Constr. Archit. Manag. 2020.27(10):3363-93.	Ineligible intervention
Aguilar GE, Hewage KN. IT based system for construction safety management and monitoring: C-RTICS2. Automation in Construction. 2013.35:217-28. doi: https://doi.org/10.1016/j.autcon.2013.05.007	Leading indicators not correlated with lagging
Agumba JN, Haupt TC. Identification of health and safety performance improvement indicators for small and medium construction enterprises: A delphi consensus study. Mediterranean Journal of Social Sciences. 2012.3(3):545-57. doi: http://dx.doi.org/10.5901/mjss.2012.v3n3p545	Outcomes - not safety
Agumba JN, Haupt TC. The influence of health and safety practices on health and safety performance outcomes in small and medium enterprise projects in the South African construction industry. J. S. Afr. Inst. Civ. Eng. 2018.60(3):61-72.	Outcomes - qualitative only
Agustina F, Ansori N, Yuliatin. An ergonomic intervention model by sampling inspection and personal protective equipment in SMEs Batik Madura. Adv. Sci. Lett. 2017.23(12):12372-76.	Outcomes - not safety
Ahmed Naji GM, Nizam Isha AS, Al-Mekhlafi ABA, Sharafaddin O, Ajmal M. Implementation of leading and lagging indicators to improve safety performance in the upstream oil and gas industry. Journal of Critical Reviews. 2020.7(14):265-69. doi: http://dx.doi.org/10.31838/jcr.07.14.45	Limit - Not primary study
Akroush NS, El-Adaway IH. Utilizing construction leading safety indicators: Case study of Tennessee. Journal of Management in Engineering. 2017.33(5)doi: https://doi.org/10.1061/(ASCE)ME.1943-5479.0000546	Ineligible outcomes
Aksorn T, Hadikusumo BHW. Measuring effectiveness of safety programmes in the Thai construction industry. Construction Management and Economics. 2008.26(4):409-21. doi: 10.1080/01446190801918722	Date limit - pre-2010
Alexander D, Hallowell M, Gambatese J. Precursors of construction fatalities. II: Predictive modeling and empirical validation. Journal of Construction Engineering and Management. 2017.143(7)doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001297	Outcomes - not safety
Ali MXM, Arifin K, Abas A, Ahmad MA, Khairil M, Cyio MB, et al. Systematic literature review on indicators use in safety management practices among utility industries. IJERGQ. 2022.19(10):6198. doi: https://doi.org/10.3390/ijerph19106198	Limit - eligible SR for checking

Reference	Exclusion reason
Al-Kudmani AS. Building a safety culture - Our experience in Saudi Aramco. In: Society of Petroleum Engineers - 9th International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2008 - "In Search of Sustainable Excellence" 2008; 1772-80. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-52349103901&partnerID=40&md5=f9679bde40e96a14b88292efb492aef7	Leading indicators not correlated with lagging
Almost JM, Vandenkerkhof EG, Strahlendorf P, Caicco Tett L, Noonan J, Hayes T, et al. A study of leading indicators for occupational health and safety management systems in healthcare. BMC Health Serv Res. 2018.18(1):296. doi: https://doi.org/10.1186/s12913-018-3103-0	Outcomes - not safety
Alolah T, Stewart RA, Panuwatwanich K, Mohamed S. Determining the causal relationships among balanced scorecard perspectives on school safety performance: Case of Saudi Arabia. Accid Anal Prev. 2014.68:57-74. doi: https://doi.org/10.1016/j.aap.2014.02.002	Outcomes - qualitative only
Alves DTS, Lima GBA. Establishing an onshore pipeline incident database to support operational risk management in Brazil - Part 2: Bowtie proposition and statistics of failure. Process Safety and Environmental Protection. 2021.155:80-97. doi: https://doi.org/10.1016/j.psep.2021.09.003	Outcomes - not safety
Arias DR. Management team role in safety performance improvement. In: Society of Petroleum Engineers - 9th International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2008 - "In Search of Sustainable Excellence" 2008; 683-91. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-52349124383&doi=10.2118%2f111603-ms&partnerID=40&md5=3f795d848709ace0cda84445cd277270	Date limit - pre-2010
Arntz-Gray J. Plan, do, check, act: The need for independent audit of the internal responsibility system in occupational health and safety. Safety Science. 2016.84:12-23. doi: https://doi.org/10.1016/j.ssci.2015.11.019	Leading indicators not correlated with lagging
Arunachalam R. Behavioral safety at RasGas company limited Doha Qatar, success factors. In: Society of Petroleum Engineers - 13th Abu Dhabi International Petroleum Exhibition and Conference, ADIPEC 2008 2008; 1691-98. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-70349757734&partnerID=40&md5=5845de0ae69893baf6114b5ae4be4be1	Date limit - pre-2010
Asadzadeh SM, Azadeh A, Negahban A, Sotoudeh A. Assessment and improvement of integrated HSE and macroergonomics factors by fuzzy cognitive maps: The case of a large gas refinery. Journal of Loss Prevention in the Process Industries. 2013.26(6):1015-26. doi: https://doi.org/10.1016/j.jlp.2013.03.007	Outcomes - qualitative only
Awolusi I, Marks E, Hainen A, Alzarrad A. Incident analysis and prediction of safety performance on construction sites CivilEng. 2022.3(3):669-86. doi: https://doi.org/10.3390/civileng3030039	Leading indicators not correlated with lagging
Awolusi I, Marks E. Near-miss reporting to enhance safety in the steel industry. Iron and Steel Technology. 2015.12(10):62-68.	Outcomes - not safety
Awolusi IG, Marks ED. Safety activity analysis framework to evaluate safety performance in construction. Journal of Construction Engineering and Management. 2017.143(3)	Leading indicators not correlated with lagging
Backman C, Forster AJ, Vanderloo S. Barriers and success factors to the implementation of a multi-site prospective adverse event surveillance system. Int J Qual Health Care. 2014.26(4):418-25.	Population - not workers
Baek SH, Kwon HM, Byun HS. A study on process safety incident precursors to prevent major process safety incidents in the Yeosu chemical complex. Korean Chemical Engineering Research. 2018.56(2):212-21. doi: https://doi.org/10.9713/kcer.2018.56.2.212	Outcomes - not safety
Bahari SF, Clarke S. Cross-validation of an employee safety climate model in Malaysia. Journal of Safety Research. 2013.45:1-6. doi: https://doi.org/10.1016/j.jsr.2012.12.003	Outcomes - not safety
Baker K, Olson J, Morisseau D. Work practices, fatigue, and nuclear power plant safety performance. Hum Factors. 1994.36(2):244-57. doi: https://doi.org/10.1177/001872089403600206	Outcomes - not safety
Banda OAV, Hanninen M, Lappalainen J, Kujala P, Goerlandt F. A method for extracting key performance indicators from maritime safety management norms. WMU J. Marit. Aff. 2016.15(2):237-65.	Outcomes - not safety
Barbosa C, Azevedo R, Rodrigues MA. Occupational safety and health performance indicators in SMEs: A literature review. Work-a Journal of Prevention Assessment & Rehabilitation. 2019.64(2):217-27.	Limit - Ineligible SR
Bayramova A, Edwards DJ, Roberts C, Rillie I. Constructs of leading indicators: A synthesis of safety literature. Journal of Safety Research. 2023.85:469-84.	Limit - Ineligible SR

Reference	Exclusion reason
Bayramova A, Edwards DJ, Roberts C, Rillie I. Constructs of leading indicators: A synthesis of safety literature. Journal of Safety Research. 2023.85:469-84. doi: https://doi.org/10.1016/j.jsr.2023.04.015	Limit - Ineligible SR
Behie SW, Halim SZ, Efaw B, O'Connor TM, Quddus N. Guidance to improve the effectiveness of process safety management systems in operating facilities. Journal of Loss Prevention in the Process Industries. 2020.68	Leading indicators not correlated with lagging
Bergman ME, Payne SC, Taylor AB, Beus JM. The shelf life of a safety climate assessment: How long until the relationship with safety–critical incidents expires? Journal of Business and Psychology. 2014.29(4):519-40. doi: https://doi.org/10.1007/s10869-013-9337-2	Leading indicators not correlated with lagging
Bhagwat K, Delhi VSK, Nanthagopalan P. Construction safety performance measurement using a leading indicator-based jobsite safety inspection method: Case study of a building construction project. International Journal of Occupational Safety and Ergonomics. 2022.28(4):2645-56. doi: https://doi.org/10.1080/10803548.2021.2012350	Outcomes - not safety
Borsos A, Farah H, Laureshyn A, Hagenzieker M. Are collision and crossing course surrogate safety indicators transferable? A probability based approach using extreme value theory. Accid Anal Prev. 2020.143:105517. doi: https://doi.org/10.1016/j.aap.2020.105517	Population - mixed and NR separately
Bradshaw CP, Milam AJ, Furr-Holden CDM, Lindstrom Johnson S. The school assessment for environmental typology (SAfETy): An observational measure of the school environment. Am J Community Psychol. 2015.56(3-4):280-92. doi: https://doi.org/10.1007/s10464-015-9743-x	Outcomes - not safety
Brahmasrene T, Smith SS. The influence of training, safety audits, and disciplinary action on safety management. Journal of Organizational Culture, Communications and Conflict. 2009.13(1):9-19.	Date limit - pre-2010
Brioso X, Calderon-Hernandez C. Framework for integrating productive, contributory, and noncontributory work with safe and unsafe acts and conditions. IJERGQ. 2023.20(4):3412. doi: https://doi.org/10.3390/ijerph20043412	Leading indicators not correlated with lagging
Broadribb MP, Boyle B, Tanzi SJ. Cheddar or swiss? How strong are your barriers? (One company's experience with process safety metrics). In: Conference Proceedings - 2009 AIChE Spring National Meeting and 5th Global Congress on Process Safety 2008. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-78049286865&partnerID=40&md5=efdc8d4a87ea58fb2c03f457884f0e4f	Limit - Not primary study
Bumbary KM. Using velocity, acceleration, and jerk to manage agile schedule risk. In: Proceedings - 2016 International Conference on Information Systems Engineering, ICISE 2016 2016; 73-80. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84978473022&doi=10.1109%2fICISE.2016.21&partnerID=40&md5=cc8f83876dc7af879d367035f3495c02	Outcomes - not safety
Bush L. Growing season. Putting your reward emphasis on leading indicators will improve safety and stimulate a cultural shift. Occupational health & safety (Waco, Tex.). 2003.72(1):28-30.	Limit - Not primary study
C. PWR-The Center for Construction Research Training. Improving safety climate through a communication and recognition program for construction: A mixed-methods study. In: Key findings from research. Silver Spring, MD: CPWR-The Center for Construction Research and Training. MD; MO: Key findings from research; 2016. p. 1	Outcomes - qualitative only
Cabon P, Deharvengt S, Grau JY, Maille N, Berechet I, Mollard R. Research and guidelines for implementing fatigue risk management systems for the French regional airlines. Accid Anal Prev. 2012.45(Suppl):41-44. doi: https://doi.org/10.1016/j.aap.2011.09.024	Outcomes - not safety
Cadieux J, Roy M, Desmarais L. A preliminary validation of a new measure of occupational health and safety. Journal of Safety Research. 2006.37(4):413-19. doi: https://doi.org/10.1016/j.jsr.2006.04.008	Outcomes - not safety
Camplin JC. Demonstrating safety performance through leading indicators. In: ASSE Professional Development Conference and Exposition 2012 2012. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058498714&partnerID=40&md5=f9fd048a561276fc81b07400c2ac946c	Limit - Not primary study
Carmeli A, Reiter-Palmon R, Ziv E. Inclusive leadership and employee involvement in creative tasks in the workplace: The mediating role of psychological safety. Creativ. Res. J. 2010.22(3):250-60.	Outcomes - not safety
Carson RS. Correlating product and process measures as a model for systems engineering measurement. In: 21st Annual International Symposium of the International Council on Systems Engineering, INCOSE 2011 2011; 3138-51. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84877877954&partnerID=40&md5=ca97475e1b9feac4048e1c9040a5729b	Outcomes - not safety

Reference	Exclusion reason
Carter DT, Prevette SS. Leading with leading indicators. In: Proceedings - 10th International Conference on Environmental Remediation and Radioactive Waste Management, ICEM'05 2005; 1954-60. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-33646554515&partnerID=40&md5=e85100d8b731ac66eff91804c6401475	Date limit - pre-2010
Cermelli D, Pettinato M, Currò F, Fabiano B. Major accident prevention: A construction site approach for pro-active management of unsafe conditions. Chemical Engineering Transactions. 2019.74:1387-92. doi: https://doi.org/10.3303/CET1974232	Outcomes - not safety
Chen FL, Chen PY, Chen CC, Tung TH. Development and validation of an integrated healthy workplace management model in Taiwan. Safety and Health at Work. 2022.13(4):394-400. doi: https://doi.org/10.1016/j.shaw.2022.09.004	Outcomes - not safety
Cheung C, Xu D, Ejohwomu O, P M. Safety leading indicators. Manchester:	Outcomes - not safety
Cheung CM, Xu J, Manu P, Ejohwomu O, Freitas A. Implementing safety leading indicators in construction: Insights on relative importance of indicators. In: CIB W099 & TG59 Annual Conference 2020 Webinar 2020: Online.	Outcomes - not safety
Cheung CM, Zhang RP, Wang R, Hsu SC, Manu P. Group-level safety climate in the construction industry: Influence of organizational, group, and individual factors. Journal of Management in Engineering. 2022.38(1)doi: https://doi.org/10.1061/(ASCE)ME.1943-5479.0000978	Outcomes - not safety
Cheung DC, Xu DJ, Ejohwomu DO, Manu DP. Safety leading indicators research report. Manchester: 2019.	Outcomes - not safety
Choice and use of leading indicators to monitor health and safety performance on construction projects. White paper.	Limit - Not primary study
Cieslewicz W, Araszkiewicz K, Sikora P. Accident rate as a measure of safety assessment in Polish civil engineering. Safety. 2019.5(4)	Leading indicators not correlated with lagging
Cloos GW. More on reference dates and leading indicators. Journal of Business. 1963.36(3):352-64.	Limit - Not primary study
Cloostermans L, Bekkers MB, Uiters E, Proper KI. The effectiveness of interventions for ageing workers on (early) retirement, work ability and productivity: A systematic review. Int Arch Occup Environ Health. 2015.88(5):521-32.	Limit - Ineligible SR
Costin A, Wehle A, Adibfar A. Leading indicators—A conceptual IoT-based framework to produce active leading indicators for construction safety. Safety. 2019.5(4)doi: 10.3390/safety5040086	Outcome - no eligible outcome data reported
Cournoyer ME, Renner CM, Lee MB, Kleinsteuber JF, Trujillo CM, Krieger EW, et al. Lean six sigma tools, part III: Input metrics for a glovebox glove integrity program. Journal of Chemical Health and Safety. 2011.18(1):31-40. doi: https://doi.org/10.1016/j.jchas.2010.06.001	Outcomes - not safety
Craig BN, Das KP, Khago A. Shipboard and shore side perception of safety culture. In: IIE Annual Conference and Expo 2010 Proceedings 2010. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84901045231&partnerID=40&md5=7394b90677148fc1be4ca98449bc6fec	Outcomes - not safety
da Silva AR, Oliveira LMOS, de Moraes LMBS. Process safety integrated map in power BI. In: 2021 Spring Meeting and 17th Global Congress on Process Safety, GCPS 2021 2021. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85135217303&partnerID=40&md5=1111a92e0023620df5a454f8b027d97b	Leading indicators not correlated with lagging
Dadashi Haji M, Behnam B. An automated BIM and system dynamics tool for assessing safety leading indicators in construction projects. International Journal of Building Pathology and Adaptation. 2023.doi: http://dx.doi.org/10.1108/IJBPA-05-2022-0072	Outcomes - not safety
Davis-Street J, Stevens C, Grimsley M, Kendrick J, Boyers K, Erickson H. Cognitive issues for safe operations: Addressing bias and mindfulness to decrease human error. In: Society of Petroleum Engineers - SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability 2020, HSE and Sustainability 2020 2020. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85091406656&partnerID=40&md5=319f2a0fb1b542878046aaa77e639079	Outcomes - not safety
De Almeida AG. Identification of global indicators for risk, health, safety and environment management in production platforms - Case study of Brazilian industry. In: 5th CCPS Latin American Conference on Process Safety 2013, LACPS 2013 2013; 323-36. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84964725443&partnerID=40&md5=d628fff4c5b96a66b4164e2440abc04b	Outcomes - not safety
de La Garza JM, Hancher DE, Decker L. Analysis of safety indicators in construction. Journal of Construction Engineering and Management. 1998.124(4):312-14. doi: 10.1061/(asce)0733-9364(1998)124:4(312)	Date limit - pre-2010
DeHart Ii RE, Brand J. SHE&S excellence and innovation for the Barzan Onshore Project. In: Society of Petroleum Engineers - International Petroleum Technology Conference 2014, IPTC 2014: Unlocking Energy Through Innovation, Technology and Capability 2014; 1632-49.	Leading indicators not correlated with lagging

Reference	Exclusion reason
Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84900304373&partnerID=40&md5=8b514d7bb303cda7c00c889a766470c8	
Den Haan K. Downstream oil industry safety statistics for 2010. CONCAWE Review. 2011.20(2):16-18.	Concept - not leading indicators
Donham K, Schneiders S, Rautiainen R. Certified safe farm: Prospective research and sustainability. (Cooperative-Agreement-Number-U60-CCU-717552). Final Cooperative Agreement Report: 2004. Available from: https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB2004104752.xhtml.	Date limit - pre-2010
Dos Santos Grecco CH, Vidal MCR, Cosenza CAN, Dos Santos IJAL, De Carvalho PVR. Safety culture assessment: A fuzzy model for improving safety performance in a radioactive installation. Progress in Nuclear Energy. 2014.70:71-83. doi: https://www.sciencedirect.com/science/article/pii/S0149197013001480	Outcomes - not safety
dos Santos TR, Padoveze MC, Nichiata LYI, Takahashi RF, Ciosak SI, Gryschek ALFPL. Indicators to assess the quality of programs to prevent occupational risk for tuberculosis: Are they feasible? Rev Latino-Am Enfermagem. 2016.24:e2695. doi: https://doi.org/10.1590/1518-8345.0591.2695	Leading indicators not correlated with lagging
Dyck D, Roithmayr T. Great safety performance: An improvement process using leading indicators. Workplace Health Saf. 2004.52(12):511-20. doi: 10.1177/216507990405201205	Date limit - pre-2010
Falahati M, Karimi A, Mohammadfam I, Mazloumi A, Khanteymoori AR, Yaseri M. Development of safety and health leading performance indicators in the phase of construction of a gas refinery plant using Bayesian network and AHP. Int. J. Adv. Biotechnol. Res. 2017.8(2):1440-53.	Leading indicators not correlated with lagging
Falahatia M, Karimib A, Mohammadfamc I, Mazloumib A, Khanteymoorid AR, Yaserie M. Multi-dimensional model for determining the leading performance indicators of safety management systems. Work. 2020.67:959–69.	Leading indicators not correlated with lagging
Floyd HL. A balanced scorecard of leading and lagging indicators for your electrical safety program: Copyright Material IEEE ESW2021-12. In: IEEE IAS Electrical Safety Workshop 2021. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85113204453&doi=10.1109%2fESW45993.2021.9461566&partnerID=40&md5=617daf89de97c3a5cf5b7bfa5f3818eb	Leading indicators not correlated with lagging
Forest JJ, Kessler K. Correlating process safety leading indicators with performance. Process Safety Progress. 2013.32(2):185-88. doi: https://doi.org/10.1002/prs.11562	Outcomes - not safety
Forteza FJ, Carretero-Gómez JM, Sesé A. Safety in the construction industry: Accidents and precursors. Revista de la Construccion. 2020.19(2):271-81. doi: http://dx.doi.org/10.7764/rdlc.19.2.271.	Limit - Not primary study
Foster PJ, Parand A, Bennett JG. Improving the safety performance of the UK quarrying industry through a behavioural based safety intervention. In: Journal of the Southern African Institute of Mining and Metallurgy 2008; 683-90. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-58149293694&partnerID=40&md5=fd53a29f71752e2ce1940052776515cf	Date limit - pre-2010
Ghosh S, Nourihamedani M, Reyes M, Snyder L. Association between leading indicators of safety performance in construction projects. International Journal of Construction Education and Research. 2023.doi: https://doi.org/10.1080/15578771.2023.2195209	Outcomes - not safety
Goodman PS. Determining the effect of incentive programs on the occurrence of accidents, injuries, and productivity- Final report. (OFR 47-88). Final Contract Report: 1987. 143. Available from: https://ntrl.ntis.gov/NTRL/dashboard/searchResults/titleDetail/PB88234257.xhtml.	Date limit - pre-2010
Grabowski M, Ayyalasomayajula P, Merrick J, Harrald JR, Roberts K. Leading indicators of safety in virtual organizations. Safety Science. 2007.45(10):1013-43. doi: https://doi.org/10.1016/j.ssci.2006.09.007	Outcomes - not safety
Grabowski M, Ayyalasomayajula P, Merrick J, McCafferty D. Accident precursors and safety nets: Leading indicators of tanker operations safety. Maritime Policy and Management. 2007.34(5):405-25. doi: 10.1080/03088830701585084	Outcome - no eligible outcome data reported
Grabowski MR, Ayyalasomayajula P, Haiyuan W, Merrick JR, McCafferty D, Meador ML, et al. Accident precursors and safety nets: Initial results from the leading indicators of safety project. In: Transactions - Society of Naval Architects and Marine Engineers 2007; 288-95. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-49249107695&partnerID=40&md5=6ff49a58cbf84ca1fd9613658ca0d80f	Date limit - pre-2010

Reference	Exclusion reason
Hameed H, Sarfraz MA. Measuring vital signs of process safety culture. In: 36th Center for Chemical Process Safety International Conference, CCPS 2021 - Topical Conference at the 2021 AIChE Spring Meeting and 17th Global Congress on Process Safety 2021; 455-69. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85141681484&partnerID=40&md5=696897bd64241ff663f189b129f697b8	Leading indicators not correlated with lagging
Henselwood F. The use of the pareto shape parameter as a leading indicator of process safety performance. Process Safety Progress. 2009.28(3):221-26. doi: https://doi.org/10.1002/prs.10290	Concept - not leading indicators
Hinze J, J G. Factors that influence safety performance of specialty contractors. Journal of Construction Engineering and Management. 2003.129(2)doi: http://dx.doi.org/10.1061/(ASCE)0733-9364(2003)129:2(159)	Date limit - pre-2010
Hinze J, P R. Safety on large building construction projects. Journal of Construction Engineering and Management. 1988.114(2)doi: https://doi.org/10.1061/(ASCE)0733-9364(1988)114:2(286)	Date limit - pre-2010
Iodoro G I. Health and safety management efforts as correlates of performance in the Nigerian construction industry. Journal of Civil Engineering and Management. 2008.14(4):277-85. doi: https://doi.org/10.3846/1392-3730.2008.14.27	Date limit - pre-2010
Jafari P, Mohamed E, Pereira E, Kang S, AbouRizk S. Leading safety indicators: Application of machine learning for safety performance measurement. In: Proceedings of the 36th International Symposium on Automation and Robotics in Construction, ISARC 2019 2019; 501-06. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85071497699&doi=10.22260%2fisarc2019%2f0067&partnerID=40&md5=465be9dd6cb36d2ece55c46e965d9be7	Leading indicators not correlated with lagging
Jaselski E J, Anderson S D, S RJ. Strategies for achieving excellence in construction safety performance. Journal of Construction Engineering and Management. 1996.122(1)doi: https://doi.org/10.1061/(ASCE)0733-9364(1996)122:1(61)	Date limit - pre-2010
Johnsen SO, Okstad E, Aas AL, Skramstad T. Proactive indicators of risk in remote operations of oil and gas fields. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2010 2010; 804-25. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-77954230104&doi=10.2523%2f126560-ms&partnerID=40&md5=8218f20413d41ccc6a4269f13b837e87	Outcomes - qualitative only
Johnson SE. The predictive validity of safety climate. Journal of Safety Research. 2007.38(5):511-21. doi: 10.1016/j.jsr.2007.07.001	Date limit - pre-2010
Kandola R, Curcuruto M, Griffin M, Morgan JI. The influence of organisational safety climate on group safety outcomes: The mediation role of supervisor safety communication and monitoring. In: Advances in Intelligent Systems and Computing; 2019. p. 35-46.	Outcomes - not safety
Karimi A, Abbasi M, Zokaei M, Falahati M. Development of leading indicators for the assessment of occupational health performance using Reason's Swiss cheese model. Journal of Education and Health Promotion. 2021.10(1)doi: https://doi.org/10.4103/jehp.jehp_1326_20	Leading indicators not correlated with lagging
Keshavarzi A, Rezapour M, Safari S. Analysis of the factors affecting the safety performance in the Iranian power distribution companies - Hybrid approach of DEMATEL and ISM. Iran Occupational Health. 2021.18(1):169-85. doi: http://dx.doi.org/10.52547/ioh.18.1.169	Outcomes - not safety
Khan F, Abunada H, John D, Benmosbah T. Development of risk-based process safety indicators. Process Safety Progress. 2010.29(2):133-43. doi: https://doi.org/10.1002/prs.10354	Outcomes - not safety
Kline J. Can safety be the master measure? In: Conference Proceedings - IEEE-IAS/PCA Cement Industry Technical Conference 2019. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85067975794&doi=10.1109%2fCITCON.2019.8729099&partnerID=40&md5=80a0126dfce20b20c88a3c9b6aa3ee70	Leading indicators not correlated with lagging
Knijff P, Allford L, Schmelzer P. Process safety leading indicators - A perspective from Europe. In: 28th Center for Chemical Process Safety International Conference 2013, CCPS - Topical Conference at the 2013 AIChE Spring Meeting and 9th Global Congress on Process Safety 2013; 54-66. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84883662499&partnerID=40&md5=e82fcf5936616b3c9e37639654217435	Outcomes - not safety
Kongsvik T, Kjøs Johnsen ST, Sklet S. Safety climate and hydrocarbon leaks: An empirical contribution to the leading-lagging indicator discussion. Journal of Loss Prevention in the Process Industries. 2011.24(4):405-11. doi: https://doi.org/10.1016/j.jlp.2011.02.004	Outcomes - not safety
Landon P, Weaver P, Fitch JP. Tracking minor and near-miss events and sharing lessons learned as a way to prevent accidents. Applied Biosafety. 2016.21(2):61-65. doi: https://doi.org/10.1177/1535676016646642	Limit - Not primary study

Reference	Exclusion reason
Law R, Dollard MF, Tuckey MR, Dormann C. Psychosocial safety climate as a lead indicator of workplace bullying and harassment, job resources, psychological health and employee engagement. Accid Anal Prev. 2011.43(5):1782-93. doi: https://doi.org/10.1016/j.aap.2011.04.010	Outcomes - qualitative only
Lewchuk W, Robb AL, Walters V. The effectiveness of bill 70 and joint health and safety committees in reducing injuries in the workplace: The case of Ontario. Can Public Pol. 1996.22(3):225-43. doi: 10.2307/3551503	Date limit - pre-2010
Li RYM, Leung TH. Leading safety indicators and automated tools in the construction industry. In: ISARC 2017 - Proceedings of the 34th International Symposium on Automation and Robotics in Construction 2017; 758-65. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85032344443&doi=10.22260%2fisarc2017%2f0106&partnerID=40&md5=1e0f40a82768f870e7fef8ed5809c72d	Outcomes - not safety
Li X, Ma W. An investigation of safety management in construction workplace in China. In: Lecture Notes in Electrical Engineering 2012; 321-29. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84944061927&doi=10.1007%2f978-3-642-27326-1_42&partnerID=40&md5=fefcc0c04201e0beae651073c083e07f	Outcomes - not safety
Liu KH, Tessler J, Murphy LA, Chang CC, Dennerlein JT. The gap between tools and best practice: An analysis of safety prequalification surveys in the construction industry. New Solutions. 2019.28(4):683-703. doi: https://doi.org/10.1177/1048291118813583	Leading indicators not correlated with lagging
Maduabuchi E. Development of process safety leading indicators for major hazard installation using the causal reasoning approach. In: Institution of Chemical Engineers Symposium Series 2018. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85049568248&partnerID=40&md5=1b67ff3ea17eb8caa95c7ccdda7a1502	Leading indicators not correlated with lagging
Martin A. Health performance in the oil and gas industry - The results. In: Society of Petroleum Engineers - SPE/APPEA Int. Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2012: Protecting People and the Environment - Evolving Challenges 2012; 196-205. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84875133169&partnerID=40&md5=b9dfdf266f20ddf99c87427b1c56de87	Outcomes - not safety
Mathisen GE, Tjora T. Safety voice climate: A psychometric evaluation and validation. Journal of Safety Research. 2023.doi: https://doi.org/10.1016/j.jsr.2023.05.008	Outcomes - not safety
Mentzer R, Mannan SM. Normalization of process safety metrics Mengtian Wang. In: 12AIChE - 2012 AIChE Spring Meeting and 8th Global Congress on Process Safety, Conference Proceedings 2012. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84861423887&partnerID=40&md5=a7912f0e5a40fcda2b0c188ec12df3fd	Concept - not leading indicators
Mohamed S. Empirical investigation of construction safety management activities and performance in Australia. Safety Science. 1999.33(3):129-42. doi: 10.1016/s0925-7535(99)00028-4	Date limit - pre-2010
Mohammadfam I, Kamalinia M, Momeni M, Golmohammadi R, Hamidi Y, Soltanian A. Evaluation of the quality of occupational health and safety management systems based on key performance indicators in certified organizations. Safety and Health at Work. 2017.8(2):156-61.	Outcomes - not safety
Moore P. New safety methods boost jobsite results. ENR (Engineering News-Record). 2013.270(5)	Limit - Not primary study
Mousavi SS, Cudney EA, Trucco P. Towards a framework for steering safety performance: A review of the literature on leading indicators. In: Advances in Intelligent Systems and Computing 2018; 195-204. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85022204450&doi=10.1007%2f978-3-319-60525-8_21&partnerID=40&md5=67b4bee4301632793b65c7b263197007	Limit - Not primary study
Navarro MFL, Gracia Lerín FJ, Tomás I, Peiró Silla JM. Validation of the group nuclear safety climate questionnaire. Journal of Safety Research. 2013.46:21-30. doi: https://doi.org/10.1016/j.jsr.2013.03.005	Outcomes - not safety
Neamat SDS. A comparative study of safety leading and lagging indicators measuring project safety performance. Advances in Science, Technology and Engineering Systems. 2019.4(6):306-12. doi: http://dx.doi.org/10.25046/aj040639	Limit - eligible SR for checking
Ni Y, Sattari F, Lefsrud L, Tufail M. A rising tide raises all boats: Regional promotion of process safety through joint government/industry management. Journal of Loss Prevention in the Process Industries. 2020.68:104331. doi: http://dx.doi.org/10.1016/j.jlp.2020.104331	Leading indicators not correlated with lagging
Niu Y, Fan Y, Li Y. Safety performance measurement in collectivized oil companies in China: Contribution of leading indicators to lagging indicators. Journal of Loss Prevention in the Process Industries. 2023.83doi: 10.1016/j.jlp.2023.105090	Outcome - no eligible outcome data reported

Reference	Exclusion reason
O'Connor P, Cowan S, Alton J. A comparison of leading and lagging indicators of safety in naval aviation. Aviat Space Environ Med. 2010.81(7):677-82. doi: https://doi.org/10.3357/asem.2734.2010	Concept - not leading indicators
Parn EA, Edwards D, Riaz Z, Mehmood F, Lai J. Engineering-out hazards: Digitising the management working safety in confined spaces. Facilities. 2019.37(3-4):196-215. doi: https://doi.org/10.1108/F-03-2018-0039	Outcomes - not safety
Payne SC, Bergman ME, Rodríguez JM, Beus JM, Henning JB. Leading and lagging: Process safety climate-incident relationships at one year. Journal of Loss Prevention in the Process Industries. 2010.23(6):806-12. doi: https://doi.org/10.1016/j.jlp.2010.06.004	Outcomes - not safety
Pereira E, Ahn S, Han S, Abourizk S. Finding causal paths between safety management system factors and accident precursors. Journal of Management in Engineering. 2020.36(2)	Outcomes - not safety
Poh CQX, Ubeynarayana CU, Goh YM. Safety leading indicators for construction sites: A machine learning approach. Automation in Construction. 2018.93:375-86. doi: https://doi.org/10.1016/j.autcon.2018.03.022	Leading indicators not correlated with lagging
Reis AM, Teixeira JVS, Vergara LGL, de Sousa SDT. Worker location-based safety performance indicator. In: Studies in Systems, Decision and Control; 2020. p. 65-73.	Outcomes - not safety
Robson LS, Clarke JA, Cullen K, Bielecky A, Severin C, Bigelow PL, et al. The effectiveness of occupational health and safety management system interventions: A systematic review. Safety Science. 2007.45:329-53.	Limit - eligible SR for checking
Saba F, Mohamed Y. Integrating construction safety and project performance through safety information management in an industrial construction company. In: Proceedings, Annual Conference - Canadian Society for Civil Engineering 2009; 1324-33. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-72849110102&partnerID=40&md5=a900b9751dd618d7485d26a9029c63e2	Date limit - pre-2010
Saksvik PO, Nytro K. Implementation of internal control (IC) of health, environment and safety (HES) in Norwegian enterprises. Safety Science. 1996.23(1):53-61. doi: 10.1016/0925-7535(96)00030-6	Date limit - pre-2010
Santos AJR, Santos SP, Amado CAF, Rebelo EL, Mendes JC. Labor inspectorates' efficiency and effectiveness assessment as a learning path to improve work-related accident prevention. Ann. Oper. Res. 2020.288(2):609-51.	Outcome - no eligible outcome data reported
Sarkar S, Ejaz N, Promod CS, Maiti J. Pattern extraction using proactive and reactive data: A case study of contractors' Safety in a steel plant. Proceedings of Icetit 2019: Emerging Trends in Information Technology. 2020.605:731-42.	Concept - not leading indicators
Shaikh AY, Osei-Kyei R, Hardie M. A critical analysis of safety performance indicators in construction. International Journal of Building Pathology and Adaptation. 2020.39(3):547-80. doi: https://doi.org/10.1108/IJBPA-03-2020-0018	Limit - eligible SR for checking
Shamim MY, Buang A, Shariff AM, Anjum H. Implementation of safety performance framework (SPF) in process industries to avoid disasters. International Journal of Automotive and Mechanical Engineering. 2018.15(1):5022-35. doi: https://doi.org/10.15282/ijame.15.1.2018.10.0389	Outcomes - not safety
Shohet IM, Wei HH, Skibniewski MJ, Tak B, Revivi M. Integrated communication, control, and command of construction safety and quality. Journal of Construction Engineering and Management. 2019.145(9)doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001679	Outcomes - not safety
Short L, Ariawan A. The capability engine, analyzing the past to protect the future. In: Global Congress on Process Safety 2017 - Topical Conference at the 2017 AIChE Spring Meeting and 13th Global Congress on Process Safety 2017; 1171-75. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85027120152&partnerID=40&md5=22c42620374e376b0c555a236cc23579	Outcomes - not safety
Smith D. Health performance indicators. In: International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2002; 829-34. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-1642476679&doi=10.2118%2f73998-ms&partnerID=40&md5=03f92d931d6971aacad950a4d6d94a6d	Leading indicators not correlated with lagging
Soares BB, de Sousa LF, dos Santos MA. Evaluation of occupational safety conditions in a waste plastic recycling plant in Brazil. Indep. J. Manag. Prod. 2021.12(8):2281-96.	Outcomes - not safety
Sparer EH, Herrick RF, Dennerlein JT. Development of a safety communication and recognition program for construction. New Solutions. 2015.25(1):42-58. doi: https://doi.org/10.1177/1048291115569025	Outcomes - not safety
Stough J, Wichienrut W. Using research-vetted, actionable leading indicators to DRIVE safety performance and operational excellence. In: Society of Petroleum Engineers - International Petroleum Technology Conference 2012, IPTC 2012 2012; 2377-86. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84861379091&partnerID=40&md5=6334528bbbd11b588abb7c7b800d1f19	Leading indicators not correlated with lagging

Reference	Exclusion reason
Stough J. Strong reporting culture as stepping stone to continuously drive safety performance improvement. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2011 2011; 432-39. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-79960017510&doi=10.2118%2f140859-ms&partnerID=40&md5=9d88ac13fca7e123a7d3810bc7aa362f	Limit - Not primary study
Stough J. Strong reporting culture as stepping stone to continuously drive safety performance improvement. In: Society of Petroleum Engineers - SPE Middle East Health, Safety, Security, and Environment Conf. and Exhibition 2012, MEHSSE - Sustaining World Energy Through an Integrated HSSE and Business Approach 2012; 300-07. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84865018356&doi=10.2118%2f152566-ms&partnerID=40&md5=b5c75606152984bd36d6d000955fca87	Limit - Not primary study
Stough J. Using research-vetted, actionable leading indicators to DRIVE safety performance and operational excellence. In: Global Congress on Process Safety 2012 - Topical Conference at the 2012 AIChE Spring Meeting and 8th Global Congress on Process Safety 2012; 1232-42. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84875659384&partnerID=40&md5=d6d00cee5704cea8a079f979383521b8	Outcomes - not safety
Straub F. Leading ergonomic indicators: their importance in the American workplace, part 1. Prof Saf. 2018.63(10):60-67.	Outcomes - not safety
Tabish SZS, Jha KN. POINT of VIEW success factors for safety performance in public construction projects. Indian Concrete Journal. 2015.89(2):58-72.	Outcomes - qualitative only
Talebi E, Rogers WP, Morgan T, Drews FA. Modeling mine workforce fatigue: Finding leading indicators of fatigue in operational data sets. Minerals. 2021.11(6):621. doi: https://doi.org/10.3390/min11060621	Concept - not leading indicators
Thorsen HK, Njã O. Monitoring major accident risk in offshore oil and gas activities by leading indicators. In: PSAM 2014 - Probabilistic Safety Assessment and Management 2014. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85066559686&partnerID=40&md5=421d3320b6de80a44d141e4387903446	Outcomes - not safety
Torp S, Riise T, Moen BE. Systematic health, environment and safety activities: Do they influence occupational environment, behaviour and health? Occup Med-Oxford. 2000.50(5):326-33. doi: 10.1093/occmed/50.5.326	Date limit - pre-2010
Umer W, Li H, Szeto GPY, Wong AYL. Proactive safety measures: Quantifying the upright standing stability after sustained rebar tying postures. Journal of Construction Engineering and Management. 2018.144(4)doi: https://doi.org/10.1061/(ASCE)CO.1943-7862.0001458	Outcomes - not safety
Veley CD. Unintended consequences of a promising safety management leading indicator. In: Society of Petroleum Engineers - SPE European HSE Conference and Exhibition 2013: Health, Safety, Environment and Social Responsibility in the Oil and Gas Exploration and Production Industry 2013; 153-58. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84883334625&doi=10.2118%2f164961-ms&partnerID=40&md5=00a74b1898d1b8364ab04800c86aec0d	Outcomes - not safety
Vosoughi S, Chalak MH, Yarahmadi R, Abolaghasemi J, Alimohammadi I, Kanrash FA, et al. Identification, selection and prioritization of key performance indicators for the improvement of occupational health (Case study: An automotive company). J UOEH. 2020.42(1):35-49. doi: https://doi.org/10.7888/juoeh.42.35	Leading indicators not correlated with lagging
Vredenburgh AG. Organizational safety: Which management practices are most effective in reducing employee injury rates? Journal of safety research. 2002.33(2):259-76. doi: 10.1016/s0022-4375(02)00016-6	Date limit - pre-2010
Vredenburgh AG. Risk management: Which management practices are best predictors of employee injury rates? Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting, Vols 1 and 2. 1999.902-06.	Date limit - pre-2010
Walaski P. The role of leading and lagging indicators in OSH performance management. Prof Saf. 2020.65(08):29-35.	Limit - Not primary study
Walters N, Ross B. Predicting and mitigating the risk of catastrophic incidents. In: NPRA Annual Meeting Technical Papers 2011; 146-54. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-82555201480&partnerID=40&md5=f2f675506d55fe4c83601ba963295628	Leading indicators not correlated with lagging
Wijnia YC, Hermkens RJM, Flonk J. The safety indicator: Measuring safety in gas distribution networks. In: Engineering Asset Management Review; 2010. p. 327-44.	Concept - not leading indicators

Reference	Exclusion reason
Wong A. Improving management of critical controls to avoid catastrophic incidents. In: 26th Center for Chemical Process Safety International Conference 2011, CCPS - Topical Conference at the 2011 AIChE Spring Meeting and 7th Global Congress on Process Safety 2011; 21-30. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-80051689299&partnerID=40&md5=988b1e7bcc5eee164dce5f83eb756cff	Leading indicators not correlated with lagging
Wood. Data exploration & modelling report. Discovering safety leading indicators. 2021.	Outcome - no eligible outcome data reported
Wood. Data summary & quality report. Discovering safety leading indicators. 2021.	Outcome - no eligible outcome data reported
Wood. Existing software solutions review & assessment. Discovering safety leading indicators. 2021.	Concept - not leading indicators
Wood. Safety data collection survey. Discovering safety leading indicators. 2021.	Concept - not leading indicators
Wood. Solution requirements specification. Discovering safety leading indicators. 2021.	Leading indicators not correlated with lagging
Wozniak SM, Kakkanattu A. Using connected systems to improve process safety. In: Sustainable Engineering Forum 2018 - Topical Conference at the 2018 AIChE Spring Meeting and 14th Global Congress on Process Safety 2018; 373-84. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85095722457&partnerID=40&md5=4370b31a18c284003df4e2f4c21fa38e	Leading indicators not correlated with lagging
Xu J, Cheung C, Manu P, Ejohwomu O, Too J. Implementing safety leading indicators in construction: Toward a proactive approach to safety management. Safety Science. 2023.157doi: https://doi.org/10.1016/j.ssci.2022.105929	Outcomes - not safety
Xu J, Cheung C, Manu P, Ejohwomu O. Safety leading indicators in construction: A systematic review. Safety Science. 2021.139doi: https://doi.org/10.1016/j.ssci.2021.105250	Limit - eligible SR for checking
Yanar B, Amick BC, Lambraki I, D'Elia T, Severin C, Van Eerd D. How are leaders using benchmarking information in occupational health and safety decision-making? Safety Science. 2019.116:245-53. doi: https://doi.org/10.1016/j.ssci.2019.03.016	Outcomes - qualitative only
Zhang S, Xiaoyan S. Study on nuclear safety management based on multiple nuclear power plants experience feedback management. In: International Conference on Nuclear Engineering, Proceedings, ICONE 2022; 6. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143202834&doi=10.1115%2fICONE29-93104&partnerID=40&md5=e3b877df386e7bf6efa6f23ea2a45e9a	Outcomes - not safety
Zohar D. Thirty years of safety climate research: Reflections and future directions. Accid Anal Prev. 2010.42(5):1517-22. doi: https://doi.org/10.1016/j.aap.2009.12.019	Limit - Not primary study
Zwetsloot G, Leka S, Kines P, Jain A. Vision zero: Developing proactive leading indicators for safety, health and wellbeing at work. Safety Science. 2020.130doi: https://doi.org/10.1016/j.ssci.2020.104890	Leading indicators not correlated with lagging
Unobtainable publications*	
Abdelmalek M, Guedes Soares C. Performance-based leading risk indicators of safety barriers on liquefied natural gas carriers. In: Developments in Maritime Technology and Engineering - Proceedings of the 5th International Conference on Maritime Technology and Engineering, MARTECH 2020 2021; 211-20. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85145494115&doi=10.1201%2f9781003216582-23&partnerID=40&md5=c24fc41fa769f84d84a000124aaf3639	Unobtainable publication
Adzivor EK, Emuze F, Das DK. Indicators for safety culture in SME construction firms: a Delphi study in Ghana. Journal of Financial Management of Property and Construction. 2022.doi: 10.1108/jfmpc-04-2022-0020	Unobtainable publication
Ali U, Sarfraz MA. 50 million safe man-hours without a lost time incident – A success story. In: 2020 AIChE Virtual Spring Meeting and 16th Global Congress on Process Safety 2020. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85106036975&partnerID=40&md5=b23dea79d96c79f8fc468fdde07d7d0c	Unobtainable publication

Reference	Exclusion reason
Alruqi WM, Hallowell MR. Critical Success Factors for Construction Safety: Review and Meta-Analysis of Safety Leading Indicators. Journal of Construction Engineering and Management. 2019.145(3)doi: 10.1061/(asce)co.1943-7862.0001626	Unobtainable publication
Al-Shabbani Z, Ammar A, Dadi G. Preventative Safety Metrics with Highway Maintenance Crews. In: Construction Research Congress 2022: Health and Safety, Workforce, and Education - Selected Papers from Construction Research Congress 2022 2022; 510-19. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85128877110&doi=10.1061%2f9780784483985.052&partnerID=40&md5=011cd03161402a90d7053c0414ab9f08	Unobtainable publication
Al-Shuwaikhat HI, Al-Mubarak YA, Al-Khalaf AA, Khan RA, Al-Shuwaier SA, Gazal MA, et al. An effective health, safety and environmental management system for production engineering organization. In: Society of Petroleum Engineers - SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition 2011 2018; 275-82. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058776153&partnerID=40&md5=ef0bc630e4c9177be5ed550ba574e70c	Unobtainable publication
Assis Edelson LB, Modica Jose E. The process of improving the safety culture and behavior of contractors workers based on transpetro policy. In: Rio Pipeline Conference and Exposition, Technical Papers 2015. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85044644343&partnerID=40&md5=100ca258f197f22251337780a0958a86	Unobtainable publication
Austin D, Walters N. Predicting and mitigating the risk of catastrophic incidents. In: NPRA National Safety Conference Papers 2011; 390-411. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84455183855&partnerID=40&md5=dc1f8204961dd788197ccccb20cc68f0	Unobtainable publication
Bada AJ, Abia D, Igwe A, Ugbebor J. No lost time injury: A success story of a multi-disciplinary approach in Amenam Kpono field. In: Society of Petroleum Engineers - Nigeria Annual International Conference and Exhibition 2010, NAICE 2010; 411-18. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-78650498814&doi=10.2118%2f136994-ms&partnerID=40&md5=7099b639c485ac3e2bd7cbf0c2dfcd17	Unobtainable publication
Barry G. Using lagging data to develop leading indicators to improve the management of frontline operational risk. In: Society of Petroleum Engineers - SPE/APPEA Int. Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2012: Protecting People and the Environment - Evolving Challenges 2012; 2364-79. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84875258480&partnerID=40&md5=5f59aa1dd17d21462362f88ce14029a1	Unobtainable publication
Batson RG, Moynihan GP, Zhou L. Control charts for system safety indicators. Asq's 54th Annual Quality Congress Proceedings. 2000.177-89.	Unobtainable publication
Bibby D. Using lagging and leading indicators and the importance of benchmarking contractor performance. In: Proceedings - SPE Annual Technical Conference and Exhibition 2016. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84994067226&doi=10.2118%2f181262-ms&partnerID=40&md5=5c1d7da14b2352beb41b21482838eafa	Unobtainable publication
Bouacha F. HSE performance measurement and proactive reporting system within in Amenas Operations. In: Society of Petroleum Engineers - 9th International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2008 - "In Search of Sustainable Excellence" 2008; 1524-30. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-52349085406&doi=10.2118%2f111803-ms&partnerID=40&md5=11b256ca10d8c9d9c03f0a1e73fea38c	Unobtainable publication
Britten T. Safety from an executive’s point of view: Turning complaints into efficiencies. In: ASSE Professional Development Conference and Exposition 2010 2010. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85060156016&partnerID=40&md5=bb760baaf3cc87682bdc33ebd1cffb00	Unobtainable publication
Cann N. Safe systems development. In: 41st Annual Human Factors and Ergonomics Society of Australia Conference 2005: Technology Improving Performance, HFESA 2005 2005; 24-35. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84879978522&partnerID=40&md5=811606edf3dd863a97cc9c248327e09b	Unobtainable publication
Chang AS, Tsai YW. Developing performance framework and measures for construction projects. In: EASEC-11 - Eleventh East Asia-Pacific Conference on Structural Engineering and Construction 2008. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84870974999&partnerID=40&md5=5ef3ab397e1dd5a95c586a42f3b7578f	Unobtainable publication

Reference	Exclusion reason
Chen Y, Si H. Empirical Study of Occupational Safety Health Management System. Proceedings of the 5th International Academic Conference on Environmental and Occupational Medicine. 2010.125-28.	Unobtainable publication
Dale AM, Barrera M, Evanoff BA. Incorporating Ergonomics into a Construction Safety Management System. In: Lecture Notes in Networks and Systems 2021; 303-08. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85111164244&doi=10.1007%2f978-3-030-74608-7_38&partnerID=40&md5=19066d5b1229450258a16207310d4ae7	Unobtainable publication
Darnbrough J, Elliott M, Durkin S. Fast-tracking safety. Australian Mining. 2008.100(9):14.	Unobtainable publication
Davis-Street J, Kendrick J, Castillejo L, Grimsley M. Stress - Impacts on health and human performance. In: Society of Petroleum Engineers - SPE International Conference and Exhibition on Health, Safety, Security, Environment, and Social Responsibility 2016. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088356131&doi=10.2118%2f179488-ms&partnerID=40&md5=8101bc3fc0e335528f9db42421da794a	Unobtainable publication
Devine J, Borener S. How do safety culture assessments relate to objective operational safety performance? In: 10th International Conference on Probabilistic Safety Assessment and Management 2010, PSAM 2010 2010; 1976-86. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84873575212&partnerID=40&md5=b69d130643fa7ba7cae6dc5f4dd77585	Unobtainable publication
Drennan F. Work site stretching programs: Five key processes for continuous improvement. In: ASSE Professional Development Conference and Exposition 2013 2013. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058505680&partnerID=40&md5=c6c8edce40e1d740988af43f4fc467dc	Unobtainable publication
Dujmovich TJ. Quantitative performance measure of HSE management system effectiveness. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production 1998, HSE 1998 1998. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85059066588&partnerID=40&md5=917ad97490616fe9d9822d4a49cd67f9	Unobtainable publication
Egbevurie B, Mustapha A, Chine O, Amadi A. Promoting safe work environment and good HSE Culture by reporting leading indicators: Case study in oil and gas industry. In: Society of Petroleum Engineers - SPE African Health, Safety, Security and Environment and Social Responsibility Conference and Exhibition 2016 2016; 1-5. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85040367493&partnerID=40&md5=daf249c9ed134adccb7ac988244e0e1a	Unobtainable publication
Es'haghi M, Nikravesh A, Allah-Bakshi H, Yarahmadi H, Poursheikhali E. Identifying the indicators influencing zero accident vision through social network analysis: case study in a mine. International Journal of Mining and Mineral Engineering. 2023.13(3):185-204. doi: 10.1504/ijmme.2022.129526	Unobtainable publication
Flin R, Mearns K, Gordon R, Fleming M. Measuring safety climate on UK offshore oil and gas installations. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production 1998, HSE 1998 1998. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85088773906&doi=10.2523%2f46741-ms&partnerID=40&md5=bb2dfce7d020f08755055b5342c2772f	Unobtainable publication
Flynn SA. HSE leadership - One company's process safety journey. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2010 2010; 873-79. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-77954227940&doi=10.2523%2f126584-ms&partnerID=40&md5=0e6844f239e1683ca5d254451a5e560a	Unobtainable publication
Forst NV. MOC leading indicator for improvements in training and PS performance. In: 5th CCPS Latin American Conference on Process Safety 2013, LACPS 2013 2013; 770-87. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84964699639&partnerID=40&md5=7c75838ceea4e5cb0df641815a5bdd12	Unobtainable publication
Gidley JK, Campbell DL. An offshore drilling company's approach to process safety management. In: SPE/IADC Drilling Conference, Proceedings 2017; 536-45. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85050960860&doi=10.2118%2f184637-ms&partnerID=40&md5=e1bbbe755a1d378253d9397f9876c30b	Unobtainable publication
Haase KM. Measuring process safety incidents. In: NPRA International Petrochemical Conference Papers 2008; 164-76. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-53549119974&partnerID=40&md5=3235f92a301c2008bd581a43678b6e63	Unobtainable publication

Reference	Exclusion reason
Houri AA, Sadeghi K. Safety management of offshore structures: overview. Infrastruct. Asset Manag. 2022.10(2):67-82.	Unobtainable publication
Irshad N, Sarfraz MA. Reducing process safety incidents by revamping leading safety indicators at fatima fertilizers limited Pakistan. In: 2020 AIChE Virtual Spring Meeting and 16th Global Congress on Process Safety 2020. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85106039112&partnerID=40&md5=45ce51b54a9e09c64743235ef4af34d9	Unobtainable publication
Job A, Silva IS. Safety performance and its measurement: An empirical study concerning leading and lagging indicators. In: Occupational Safety and Hygiene VI - Selected contributions from the International Symposium Occupational Safety and Hygiene, SHO 2018 2018; 431-36. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85061309579&doi=10.1201%2f9781351008884-76&partnerID=40&md5=ce913a508e8dcb5deb2d18b8044fb186	Unobtainable publication
Karowich P, Mallett C, Woods C, Cowie H. Creating a culture of safe driving behaviors. In: 8th SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2006 2006; 112-22. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-33745368160&doi=10.2118%2f98376-ms&partnerID=40&md5=7b186d62a1f7cc426206c48dbff7f894	Unobtainable publication
Kenneth Koves G, Barnes V, Maret D, Morrow S. Hard data on soft subjects: The relationship between safety culture and nuclear plant performance. In: 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012 2012; 6417-27. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84873174259&partnerID=40&md5=fc19cf91e3b142d883d99cef93dbbc7f	Unobtainable publication
Maxwell SA, Harley A. The implementation and automation of effective key performance indicators to manage process safety and asset integrity risk: A power company case study. In: Society of Petroleum Engineers - SPE E and P Health, Safety, Security and Environmental Conference - Americas 2015 2015; 64-76. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84942596457&doi=10.2118%2f173494-ms&partnerID=40&md5=996f21f832d2a982073e71112635a5ff	Unobtainable publication
McCormick C. Cultural shift. Pulp and Paper Canada. 2013.114(2):9-11.	Unobtainable publication
Mosher GA. Safety, human interactions, and decision-makings processes. In: Storage of Cereal Grains and their Products; 2022. p. 407-20.	Unobtainable publication
Nonno L, Hinton J. Driving HSE culture change - A case history. In: World Petroleum Congress Proceedings 2017. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85043293473&partnerID=40&md5=561c3719a9b6e38d95363dfe1646d0aa	Unobtainable publication
Paltrinieri N, Tugnoli A, Øien K, Cozzani V. Synergy between DyPASI and well-known safety indicator methodologies in the prevention of atypical accident scenarios. In: 11th International Probabilistic Safety Assessment and Management Conference and the Annual European Safety and Reliability Conference 2012, PSAM11 ESREL 2012 2012; 4209-18. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84873185319&partnerID=40&md5=1d7c6907f74d746b123f13b26ea12f1d	Unobtainable publication
Redinger C. Benchmarking in International Safety and Health. In: Global Occupational Safety and Health Management Handbook; 2019. p. 95-112.	Unobtainable publication
Robert S, Fiffick W, King K, Guillory R. A proactive safety culture leading to reduced TRIR. In: SPE/IADC Drilling Conference, Proceedings 2017; 884-98. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85040458593&doi=10.2118%2f184681-ms&partnerID=40&md5=8f35f765a6c97a2ac9bf76859352dc52	Unobtainable publication
Scott H, Demirtas U. Driving organizational behavior through leading indicators: Active passenger campaign. In: Society of Petroleum Engineers - SPE International Conference and Exhibition on Health, Safety, Environment, and Sustainability 2020, HSE and Sustainability 2020 2020. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85091445044&partnerID=40&md5=d69845a9b330705d4ed25087f9b7d873	Unobtainable publication
Stough J, Robinson N. From research to practice: A story of an actionable safety leading indicator index. In: SPE/IADC Drilling Conference, Proceedings 2013; 1153-544. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84880975822&doi=10.2118%2f163552-ms&partnerID=40&md5=ead7f38812d79fbf616062904509dfa6	Unobtainable publication

Reference	Exclusion reason
Stough J. From research to practice: A story of mathematically predictive actionable process Safety leading indicator index. In: AIChE Annual Meeting, Conference Proceedings 2013. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84881326675&partnerID=40&md5=b811557b43f67038e3a634841170488a	Unobtainable publication
Stough J. Growing a culture of process risk reduction with statistically validated actionable leading indicators. In: 5th CCPS Latin American Conference on Process Safety 2013, LACPS 2013 2013; 272-82. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84964797338&partnerID=40&md5=2c37cbe1c87cec4abf563569b3dd8b26	Unobtainable publication
Thompson K, Heavisides J, Bearfield G, Griffin D. Trialling the use of safety performance indicators within Great Britain's railway industry. In: Advances in Safety, Reliability and Risk Management - Proceedings of the European Safety and Reliability Conference, ESREL 2011 2012; 1515-22. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84856719884&partnerID=40&md5=3b2e0dd97c935607ff11f2a8edca7536	Unobtainable publication
Tugnoli A, Cozzani V. Optimization of safety barriers by staged LOPA. In: Offshore Mediterranean Conference and Exhibition 2009, OMC 2009 2009. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85051553683&partnerID=40&md5=19d92fe930ffe9e814d78ac31557162c	Unobtainable publication
Whitney P, Stanley A. Behavior-based safety implementation in Thailand: Factors critical to success. In: 2005 SPE Asia Pacific Health, Safety and Environment Conference and Exhibition - Proceedings 2005; 303-13. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-33646578145&doi=10.2118%2f96600-ms&partnerID=40&md5=21386db45fb2dbee573a1fb61e7331f6	Unobtainable publication
Zhu W, Curtis R, Curry J, McSweeney K, Craig B, Wari E. IDENTIFYING SAFETY LEADING INDICATORS FOR THE OFFSHORE INDUSTRY. In: SNAME 28th Offshore Symposium, TOS 2023 2023. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85158982371&doi=10.5957%2fTOS-2023-019&partnerID=40&md5=c887faa330125bc0f707a90491dab361	Unobtainable publication

Abbreviations: SR – Systematic review.
Unobtainable publications were those that we were unable to access through the University of York library or the British Library.

Appendix E: Full Risk of Bias Assessment

Table E.1: Full Risk of Bias Assessment

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Alarcón 2016 [38]	Cross-sectional	Judgement	Yes	No	No	Yes	Unclear	Moderate
		Description	All construction companies reported in the database were included in the analysis.	Potential confounders were not reported. Very little information about the companies and no information about the employees reported.	Data were reported by the companies and there was variation in the way that they categorised the practices, although an attempt was made by the authors to homogenise the data. NR whether the authors were blinded to the accident rates when they classified the practices.	Multiple methods for comparing frequencies applied.	Very little further information is reported in the paper, so it is not clear.	Although a large and representative sample were used, these is not enough information to know if there was confounding or differences in the way that the data were collected/ reported There was also potential bias from lack of blinding.
		Judgement	No	Unclear	No	No	No	High

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Amir-Heidari 2017 [28]	Case series	Description	Authors did not report how the three companies used for the case study were selected.	NR	Companies scored themselves which could've introduced bias, especially if the indicators were scored by senior or management staff (authors do not report which employees were asked to score the indicators).	No statistical analysis was conducted.	Only 4 of the 8 indicators included in the lagging indicator composite score were eligible but could not be separated.	NR how the sample was collected and no discussion of potential confounders. Scores were self-reported, and no statistical analysis was conducted.
		Judgement	Yes	No	No	Yes	Unclear	Moderate
Bitar 2018 [25]	Cross-sectional	Description	Although it's NR why this particular company was chosen, the survey was sent to all workers.	Potential confounders were not reported.	It was a company specific survey, and the authors do not report whether it had been validated previously.	A multiple regression analysis was conducted to assess the correlation between leading and lagging indicators.	There was a large sample size across 8 different countries, but there may be unreported confounding factors as the authors do not report basic demographics of the participants.	There was a large sample size and everyone in the company was eligible, but it was unclear why this company was chosen, whether the survey had been validated and what the key population characteristics were.
		Judgement	Unclear	Yes	Yes	Yes	Yes	Low

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Brandt 2023 [27]	Cohort study	Description	Data used was gathered as part of another study and details of their eligibility criteria were provided but not recruitment (sample were described as “probability samples” but no further details given).	Weighted Cox-regression was used to control for age, gender, survey year, education, lifestyle, psychosocial work factors, occupational group, and depressive symptoms.	The survey used has been shown to be reliable and valid.	Survey version of the Cox proportional hazard model to determine the hazard ratio of long-term sickness absence based on number of safety climate problems.	No concerns identified.	Limited data on sampling and recruitment methods.
		Judgement	Yes	No	Unclear	No	Yes	Moderate
Breitsprecher 2014 [22]	Before-after	Description	Outcomes were reported across the whole region rather than a sample of employees.	No confounders were reported or discussed. Authors did not report the number of employees included, the number of sites or which countries. The also did not report key characteristics of the population or discuss any other potential explanations for the change in outcomes over time.	Methods for measurement of outcomes or definition of outcomes NR.	No statistical tests were conducted.	No other concerns identified.	Participants across the whole region included, however the methods were unclear, potential confounders were not discussed and no statistical tests were conducted.
		Judgement	No	Unclear	Unclear	Unclear	Unclear	High

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Campbell Institute 2015 [23]	Case study	Description	The case studies used are companies associated with the Campbell institute and there's no justification for why they were specifically chosen.	No information provided.	Limited information about the content of the interviews or discussions and how the questions were decided on.	Report publishing short summaries of 5 individual case studies, of which relevant data are provided for 2: correlation coefficient (Cummins) and qualitative comparison of frequencies (Honeywell) reported, no indication of statistical significance.	No information on industry or numbers of workers was provided which makes the results hard to interpret in terms of generalisability. Also, not much evidence provided at all to back up the statements in the case studies.	Not much data provided at all, very surface Level case studies of population with associations to the Campbell Institute.
		Judgement	No	Unclear	Unclear	Yes	Yes	Moderate
Cao 2019 [51]	Case study	Description	No information provided about why this company was chosen or who the population were.	No information provided.	The data used was part of already existing data sets gathered by the company. There was no information about what exactly was gathered or	Correlations between variables and modelling (VAR, VEC, artificial neural networks and Box–Jenkins	No concerns identified.	Very little information reported about the company, population or data collection.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
					how it was gathered.	AR Integrated MA) for forecasting.		
Chen 2017 [32]	Cross-sectional	Judgement	No	Unclear	Unclear	Yes	Yes	Moderate
		Description	NR how participants were chosen.	No information provided.	The questionnaire was based on previous research but had been adapted by the authors. NR whether the adapted version had been validated. Also, 54 respondents were excluded due to missing data.	Correlation, structural equation modelling.	No other concerns identified.	Unclear how the participants were selected, whether confounders were considered or whether the outcomes were measured in a validated way.
Choe 2016 [24]	Before-after	Judgement	Yes	Yes	Unclear	Yes	Yes	Low
		Description	Data taken from a national database.	Normalised results given as well.	It is not reported in detail how the fatality data was collected for the database, although examples (death certificates and OSHA reports) are given. Days away injury came from a survey which estimates workplace injuries and illnesses	Interrupted time series analysis to evaluate the mean and trend of outcomes before and after the revision.	No concerns.	Limited information about methods of data collection, but the sample was from a national database and statistical analysis was conducted which took confounders into account.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
					based on safety records of private industry employers. Depending on who completes the survey (e.g. managers, workers) it may be biased or not include injuries that went unreported.			
Coetzee 2023 [41]	Case study	Judgement	No	No	Unclear	No	Yes	High
		Description	Although all of the workers were eligible, the study was only conducted at one company, and it is not reported why this company was selected.	No confounders were discussed, and no baseline characteristics or demographics of the workers were reported.	Safety performance data were collected from company documents, however it's not clear how all of the outcomes are defined or what the timeline of implementing the framework was.	No statistical tests were conducted.	No other concerns identified.	It is unclear why the company included was chosen, outcomes are not clearly defined, confounders are not discussed, and no statistical tests were conducted.
Dadashi Haji 2023 [48]	Before-after	Judgement	No	No	Yes	No	No	High
		Description	Only one company was included, and it is not reported why that company was chosen.	The authors mentioned that the stations were similar in size, construction method and the condition of the	Accidents, injuries and fatalities were just counted and divided by the number of projects to find the rate per	Statistical tests were not conducted on the outcomes of interest for this review.	The study has a small sample size (35 sites but all from 1 project in 1 company).	Although outcomes were clearly described and compared across similar

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
				construction site but do not discuss any other confounders or give any details of the population characteristics or if they were similar in the projects from before and after the population.	station. Authors do not report how they define each outcome, but presumably it is the same before and after.			projects, it is not clear why this company was chosen for study and what the population characteristics are, and no statistical tests were conducted.
Dennerlein 2020 [52]	Cross-sectional	Judgement	No	Yes	No	Yes	Yes	Moderate
		Description	Convenience sample of active commercial construction worksites in the Boston-metro area between January 2017 and August 2018.	Linear regression models for site level were adjusted for confounders: square footage of project, project cost [in dollars], project completion percentage, FTEs worked and presence of a full-time safety manager. Company-level associations were adjusted for the presence of female workers on-site and previous participation in a prequalification survey (other than ACES).	Leading indicator was measured by survey using Likert scale responses to 63 questions and statements on company and site safety, by employees and subcontractors who are likely to have been responsible for site safety. Response bias cannot be ruled out. Insufficient information on how injury rates were recorded. Not validated survey.	Univariate correlation (Pearson's correlation coefficient), multivariable regression (linear) and multivariable regression (log-linear).	No other concerns identified.	Although appropriate statistical tests were conducted, there were some concerns as the participants are a convenience sample and the data collection method was susceptible to bias.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Doherty 2010 [12]	Before-after	Judgement	No	No	Yes	No	Yes	Moderate
		Description	NR how or why this company was the focus of the study.	Confounders are not discussed, and population characteristics are not reported.	Company used their own data.	No statistical tests of before/after frequency comparisons were conducted.	No other concerns identified.	It is unclear how or why this company was chosen for study. Confounding factors and other population characteristics are not discussed or reported and no statistical analysis was conducted.
Gale 2011 [13]	Before-after	Judgement	Unclear	Unclear	Yes	No	Yes	Moderate
		Description	Convenience sample – The company approached the research team to carry out the study with them.	No information provided.	A technology device was used to monitor driving behaviour.	No statistical tests of before/after frequency comparisons were conducted.	No concerns identified.	Convenience sample used and no statistical analysis carried out.
Govender 2022 [42]	Before-after	Judgement	Unclear	Unclear	Unclear	No	Unclear	High
		Description	No information provided about why these 9 mines were chosen.	Non reported or discussed.	Insufficient information provided about how the framework was measured.	No statistical analysis performed – trends looked at over time.	The methods state there are 9 mines from three countries (South Africa, Botswana and Namibia) yet the results say all	Small sample with limited information. Not much info provided on how the leading indicator was

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
							9 mines are from South Africa.	measured, some discrepancies in the results.
		Judgement	No	Unclear	No	Yes	Yes	Moderate
Grabowski 2010 [50]	Case series	Description	No explanation for why these three specific organisations were chosen.	No information provided.	Self-reported data.	Correlations between leading and lagging indicators.	No concerns identified.	Quite a small sample of three companies and data was obtained through surveys so response bias cannot be ruled out.
		Judgement	No	Unclear	No	No	Unclear	High
Haas 2018 [49]	Before-after	Description	No explanation for why this specific mine was chosen.	No information provided.	The data was obtained through a variety of different methods (interviews, observations and through existing documentary materials). Can't be certain that these methods were validated or unbiased.	No statistical analyses were performed. The trend in incident rates is shown over time but not statistically analysed.	No other concerns identified.	Small sample size, data collected through qualitative unvalidated measured, and no statistical analysis was conducted.
		Judgement	No	Unclear	No	Yes	Yes	Moderate

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Hinze 2013 [39]	Cross-sectional	Description	Included organisations were represented by the research team. No details of companies provided.	No information provided.	Survey used to gather data which was obtained through reviewing surveys used in previous studies. No information provided on if this was validated.	Kendall's correlation coefficient	No concerns identified.	Significant concerns around bias introduced from the sampling method. Also limited information about the participants and no information about the population. No information about potential confounders of the whether the questionnaire was validated.
		Judgement	Unclear	Unclear	Yes	Yes	Yes	Moderate
Lagerstrom 2019 [43]	Cross-sectional	Description	Participants were recruited from safety workshops, but it is unclear whether the workshops were mandatory or whether the participants had volunteered for them.	Participant characteristics (logging system, supervisory status, accreditation, education level, years of experience, presence of musculoskeletal symptoms) were converted into categorical variables	A modified version of a validated survey was used.	Multinomial logistic regression.	No other concerns identified.	It is unclear whether the participants were a representative sample and whether confounding factors were appropriately controlled for.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
				and safety climate scores were compared between groups for each characteristic, but it is unclear whether the other characteristics were controlled for in each comparison.				
		Judgement	No	No	Yes	Yes	Yes	Moderate
Laitinen 2010 [29]	Case series	Description	The sample contained companies who were voluntarily participating in the contest. Although this did include the majority of construction companies in the region, we do not know anything about the other construction companies in the region or whether this is generalisable to other regions.	Some discussion in the discussion section that the results could have occurred without the contest, but no confounders were factored into the analysis.	Inter-observer tests were carried out. The accident figures were reported by the Federation of Accident Insurance Institutions.	Accident risk was predicted based on past trends and compared to actual accident risk.	No other concerns identified.	Unclear whether the sample were representative, and confounders were not factored into the analysis.
		Judgement	No	Unclear	Yes	Yes	Yes	Moderate

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Laitinen 2013 [30]	Case series	Description	Convenience sample. Participants were those who had voluntarily participated in the programme.	Some outcomes were limited to blue collar workers as this was found to be significantly different to the number of accidents among white collar workers, but no other demographic factors or confounders were discussed.	The Elmeri+ score was developed based on previous research.	Pearson's correlation coefficient.	No other concerns identified.	The outcome measurement was based on previous research and statistical analysis was conducted; however, the population is a convenience sample and confounders are not really considered.
		Judgement	No	Unclear	Yes	Yes	Yes	Moderate
Lingard 2017 [5]	Case study	Description	Although all data from the project was used in the study there is no reasoning why this particular project was chosen.	No information provided	The data was collected prior to the research being carried out as part of routine safety data collection – There's no reason for it to be bias.	Justification and reasoning are given for their choices, p values provided to show significance of results.	No concerns identified	Data was gathered prior to research being conducted however there was no justification provided for the sample used.
		Judgement	No	No	Yes	No	Yes	Moderate
López 2013 [14]	Before-after	Description	Only one company was included, and it is not reported why that company was chosen.	None reported or discussed.	Company data is used, although the authors do not specify what all of	No statistical analysis was conducted.	No other concerns identified.	Authors do not report why this company was chosen for study or discuss

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
					the abbreviations stand for.			potential confounders and no statistical analysis was conducted.
		Judgement	No	No	No	Yes	Yes	Moderate
Manjourides 2019 [53]	Cross-sectional	Description	Clear eligibility criteria, data from 4 companies with lagging indicator values considered to be 'extreme outliers' (very high values) removed from analysis with unclear implications for bias.	Confounders not discussed and the authors state that demographic data was collected but don't report it.	Data were self-reported by the company which could introduce bias.	Zero-Inflated Poisson models.	No other concerns identified.	Self-reported data used, demographics not reported, and 'outliers' removed from analysis with unclear effects on results.
		Judgement	No	Unclear	No	Yes	Yes	Moderate
Merrick 2014 [54]	Cross-sectional	Description	No explanation for why these two specific organisations were chosen aside from the fact that they were industry partners.	Confounders not discussed.	The data was obtained through surveys and so response bias cannot be ruled out. The questionnaires were developed by the authors and not previously validated.	Stepwise multiple regression analysis.	No concerns identified.	Small sample size, data collected through surveys.
		Judgement	No	No	Unclear	Unclear	No	High

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Mohammed 2019 [15]	Case study	Description	No information provided about why this company was chosen or who the participants population were.	Non reported or discussed.	The data used was part of already existing database gathered every 2 weeks throughout the projects. There was no information about how exactly this was gathered e.g. surveys/interviews/ observation etc.	Correlation and feature selection techniques were used to identify the variables with the greatest impact on safety performance. However – no p values provided so unclear as to the significance of the results.	The paper reports the following "Additionally, the data set used in the current study was limited by its small size and by the large number (up to 50%) of missing data points." The missing data points were for the 'Experience direct hours' variables (5) and so this was removed." However out of the 36 project variables, only 23 are reported. It's not clear why this is the case.	Very limited sample information, missing data discussed but not all explained. No p values.
		Judgement	Yes	Unclear	Unclear	Yes	Yes	Moderate
Moore 2022 [31]	Cross-sectional	Description	Existing data set used. Consists of data gathered routinely.	No control measures mentioned but they did check for internal validity.	The data was obtained through surveys and so response bias cannot be ruled out. However internal	Linear regression analyses caried out to look at relationship between leading and	No concerns identified.	Decent sample size and analysis carried out on routinely collected data. Checks for internal validity carried out.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
					consistency was checked for.	lagging indicators.		However, surveys were used so can't rule out response bias.
		Judgement	Unclear	No	Unclear	Yes	Unclear	High
Mousavi 2020 [44]	Cross-sectional	Description	Multiple sources used: a contact list provided by Productivity Inc (used in previous similar research), social media and personal contacts. Not guaranteed to be fully unbiased.	None reported or discussed.	A specific questionnaire was designed for the study. Unclear if this was tested before implementation or validated in any way.	Partial least square-based structural equation modelling carried out to investigate the relationship between OHS performance and Lean implementation.	A lot of the analysis that was carried out is hard to interpret and not always reported accurately. E.g. insignificant P values marked as being significant.	Small sample size considering the wide range of industries and counties. Data gathered through a custom questionnaire and some inaccuracy in the reporting of results.
		Judgement	Unclear	Unclear	Yes	Yes	Yes	Moderate
Pereira 2017 [16]	Before-after	Description	There is no explanation for why this particular company was chosen aside from the fact that they had implemented a program they were interested in investigating.	No information provided.	The data that was used in the study was existing data gathered on a daily basis throughout the construction project.	Correlations t-tests use to assess the relationship between incident rates and the behaviour-based safety program.	No concerns identified.	Sample size details are limited which makes it hard to draw conclusions.
		Judgement	No	No	Unclear	Yes	Yes	Moderate

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Quaigrain 2023 [55]	Cross-sectional	Description	Authors did not report how the companies were selected.	Confounders not discussed and data about the population was not reported.	Authors do not report whether the CDM3 had been previously validated, but it likely used self-reported data.	Spearman's non-parametric correlation, two-tailed, was used to assess the relationships.	No other concerns identified.	Confounders, population characteristics and method of selecting participants not reported. Unclear whether data collection tool has been validated.
		Judgement	No	No	Yes	Unclear	No	High
Rajendran 2013 [56]	Case study	Description	This was a case study, and no information was given about why this project at this company was chosen. Authors also did not report how many workers were included in the analysis (although lagging indicators were measured per 100 workers) meaning that it is unclear whether the sample size was large enough for the results to be	No confounders were discussed, and no baseline characteristics or demographics of the workers were reported.	Measurement of leading and lagging indicators was clearly described.	No p values or other measure of statistical significance reported.	Only two of the four lagging indicators were reported for two of the three leading indicators.	Some concerns over lack of information about why the company was chosen for study, and no information on confounders or population characteristics.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
			representative. The authors also did not report any other information about the workers such as demographics or baseline characteristics, meaning we cannot know whether the results are applicable to other companies.					
		Judgement	Yes	Yes	Unclear	Yes	Yes	Low
Robson 2017 [33]	Cross-sectional	Description	Data from all companies meeting the eligibility criteria were analysed.	A large number of covariates were considered in the model.	The same audit tool was used at all firms, but the authors raised a concern of auditor bias. Claims data (lagging indicator) was collected from the Workplace Safety and Insurance Board of Ontario which is the sole provider of workers' compensation for occupational illness and injury in Ontario, Canada.	Data was placed in models which gave correlation coefficients.	No other concerns identified.	Concern flagged by authors about auditor boas in the collection of leading indicator data. No other concerns.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Sá 2023 [11]	Cross-sectional	Judgement	No	Unclear	No	Unclear	Yes	High
		Description	It is not reported how the 1,600 organisations that the questionnaire was sent to were chosen.	A Kruskal Wallis test was used to test for some differences between the groups, but it's not clear which groups. Also, other confounders, such as differences with population characteristics or demographics were not considered. The study also only included organisations that did implement the Lean tools and did not compare the included organisations with those that did not implement Lean tools.	The data was self-reported by the organisations.	A Kruskal Wallis test was applied but it is not clear whether the authors looked at the correlation between the leading and lagging indicators (as opposed to several potential confounders) or which groups the tests were performed on.	No other concerns identified.	Due to the unclear reporting, it is difficult to tell whether this study is useful.
Salas 2016 [40]	Case series	Judgement	Unclear	No	No	Yes	Yes	Moderate
		Description	The oil and gas industry are required to collect and report data which is uploaded to a database. NR if all the contractors with data in this database were used	The paper reports that the variability in the client-contractor relationships could impact the generation of safety data and therefore confound the results.	Data was gathered as part of a routine process however this data is self-reported (contractors report the information themselves to their clients). It's	Regression analysis to look at the predictability of leading indicators.	No concerns identified.	Exact sampling process is unclear, some concerns around how the data was gathered.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
			in the study of if it was a sample of these.		possible that this data wasn't accurate.			
Schiavi 2013 [17]	Before-after	Judgement	No	No	Unclear	No	Yes	High
		Description	Only one company was included, and it is not reported why that company was chosen.	Confounders were not discussed. Key characteristics of the company or population were not reported. The trend in accidents and injuries prior to the change in safety process was not reported.	Measurement methods not reported.	No statistical tests were conducted.	No other concerns identified.	Very little information reported about the company, population or data collection. No statistical analysis conducted.
Sheehan 2016 [37]	Cross-sectional	Judgement	No	Yes	Yes	Yes	Yes	Low
		Description	Sampling process NR.	Several confounders were controlled for in the models.	The data was self-reported, but the data collection methods were validated and reliable tools.	Multi-regression analysis.	No other concerns identified.	Unclear how the sample were identified, but otherwise no concerns.
Stough 2012 [18]	Cohort study	Judgement	No	Unclear	Unclear	No	Yes	High
		Description	Very limited sample information provided.	Non reported or discussed.	Not much information provided about how the leading indicators were measured other than it was existing dataset.	No statistical analysis performed.	No other concerns identified.	Limited information provided on sampling methods and sample. Not much usable data either.
		Judgement	Unclear	Unclear	Yes	Unclear	Yes	Moderate

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
Tang 2017 [26]	Cross-sectional	Description	Questionnaire sent out through a variety of methods to people who met their criteria. Not much detail provided about the sample itself. The study reports that the sample size is reflective of the area of interest but not ideal statistically.	Non reported or discussed.	Questionnaire used was piloted to validate and improve the questionnaire. Reliability of survey items measured using Cronbach's alpha which indicated good internal consistency.	No report of statistical testing for linear regression. Wilcoxon signed ranks tests with explanations for selection of statistical method provided.	No concerns identified.	Reflective sample size for the area but still on the small size. Validated questionnaire used to gather data and appropriate statistical analysis.
		Judgement	Unclear	Unclear	Yes	Yes	Yes	Moderate
Tang 2018 [35]	Cross-sectional	Description	No information provided about why these three companies were chosen or why these 10 platforms were chosen. Also, NR which platforms were associated with which company.	Non reported or discussed.	Questionnaire used was based off leading indicators identified in previous research. The questionnaire was piloted to validated and improved upon.	Correlation analysis to investigate the safety performance on 10 oil and gas platforms.	No concerns identified.	Small sample with limited information but otherwise a good paper.
		Judgement	No	Unclear	Unclear	No	Yes	High
Tauseef 2012 [19]	Before-after	Description	No information provided about why this company was chosen or who the participants population were.	Non reported or discussed.	It's not fully clear how the observation intervention program was measured. There is	No statistical significance testing of frequency changes.	No concerns identified.	Scarce information provided about the sample or how exactly the leading

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
					mention of reports being made but no information on what exactly this involves.			indicator data was measured. The study was also carried out after the program was partly implemented and not much information is provided about this - unclear how It impacted the results.
Thananan 2014 [20]	Before-after	Judgement	No	Unclear	Unclear	No	Yes	High
		Description	No information provided about the sample size provided.	Non reported or discussed.	Surveys carried out but detail not provided on what they contained or if they were validated.	No statistical analysis applied - simple before and after comparisons.	No concerns identified.	Scarce information provided about the sample or how exactly the leading indicator was measured.
Van Derlyke 2022 [47]	Cohort study	Judgement	Yes	Unclear	Yes	Yes	Yes	Low
		Description	Well explained and flow of participants shown.	Non reported or discussed.	Survey piloted before use.	Whitney U, Chi-squared.	No concerns. Identified.	Small sample size but otherwise good.
Versteeg 2019 [6]	Case study	Judgement	yes	Unclear	Yes	Yes	Yes	Low
		Description	Rational for sample explained and clear eligibility criteria.	Data collected retrospectively.	Routinely collected data.	Poisson modelling.	No concerns identified.	Some lack of information on how the routinely

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
								collected data was gathered.
Vosoughi 2021 [45]	Case series	Judgement	Unclear	Unclear	Unclear	Yes	Yes	Moderate
		Description	Sample was non-randomised and purposive. However, the size was deemed appropriate based on past research.	Non reported or discussed.	Not much information provided on the content or the interviews or the 'documents' that were used.	Correlations reported with p values.	No concerns identified.	Non-randomised sample and limited information provided on data collection.
Wachter 2014 [36]	Cohort study	Judgement	Yes	Unclear	Unclear	Yes	Yes	Moderate
		Description	Target sample identified within a database and survey sent to all those who were eligible.	Non reported or discussed.	Data gathered through a survey which was designed by the authors and based off past research. The reliability of each subscale was tested and those that were unreliable were removed from the survey. However - still self-reported data so possibility of bias cannot be ruled out.	Correlations and regression analysis were carried out.	No other concerns identified.	Large sample size with good data collection method. Some possible bias in the self-reported nature of the data and possible confounding variables.
Wei 2020 [21]		Judgement	No	Unclear	No	Yes	Yes	Moderate

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
	Cohort study	Description	It was not reported by the authors why or how questionnaires were sent to the individuals that they were.	Demographic data for the participants was reported but it is unclear whether the differences between the two groups was factored into the analysis.	The data was self-reported and although this may have introduced inaccuracies in terms of safety performance.	Spearman's correlation coefficient.	No other concerns identified.	Unclear whether confounders were considered in the analysis or how sampling was conducted. Also, data were self-reported.
		Judgement	No	Yes	No	Yes	Unclear	Moderate
Winge 2019 [34]	Case study	Description	Convenience sample of those who fit the criteria and are associated with the collaboration company.	They ensured the projects were all of a similar type e.g. size, building type and contractual arrangements so that comparisons could be fairly made.	interviews were piloted but still self-reported method of data collection so bias cannot be ruled out. However, this was only a part of the data collection methods.	Qualitative comparative analysis.	There were also no p values or any true indicator of significance - authors used the terms ' sufficient and necessary.	Small convenience sample with a self-report element to data collection. Some concerns around data analysis.
		Judgement	No	Yes	Yes	Unclear	Yes	Low
Zahoor 2017 [46]	Cross-sectional	Description	Sampling method was not reported.	Covariates were factored into the model.	Accident rates were self-reported, but the questionnaire was appropriately adapted and validated before use.	Following exploratory factor analysis, a measurement model was hypothesised and then tested and validated. It is reported that	No other concerns identified.	The only concern is that the authors did not report how the participants were recruited.

Study	Study Design	Description/ Judgement	Are you confident that the means of selecting and maintaining the sample minimized bias?	Are you confident that the potential confounders were adequately considered, and then either well controlled or appropriately discounted as a source of bias?	Are you confident that the measurement methods did not introduce bias to the corresponding findings?	Were appropriate statistical tests applied to the data?	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?	Risk of bias
						standardised path coefficient is "significant" but significance is not defined and no p values are reported.		

Abbreviations: ACES – Assessments of contractor safety, AR – Autoregressive, CDM3 – Construction disability management maturity model, FTE – Full time equivalent, LTSA – Long-term sickness absence, MA – Meta analysis, NR – Not reported, OSHA – Occupational safety and health administration, VAR – Vector autoregression, VEC – Vector error correction.

Appendix F: Included Studies

Table F.1: Included studies (n=48)

Reference
Alarcón L, Acuña D, Diethelm S, Pellicer E. Strategies for improving safety performance on construction firms. Accident; Analysis and Prevention. 2016.94:107-18. doi: https://doi.org/10.1016/j.aap.2016.05.021
Amir-Heidari P, Maknoon R, Taheri B, Bazyari M. A new framework for HSE performance measurement and monitoring. Safety Science. 2017.100:157-67. doi: 10.1016/j.ssci.2016.11.001
Bitar FK, Chadwick-Jones D, Lawrie M, Nazaruk M, Boodhai C. Empirical validation of operating discipline as a leading indicator of safety outputs and plant performance. Safety Science. 2018.104:144-56. doi: 10.1016/j.ssci.2017.12.036
Brandt M, Andersen LL, Kines P, Ajslev JZN. Safety climate at work and risk of long-term sickness absence: Prospective cohort with register follow-up among 63,500 workers. Safety Science. 2023.166doi: 10.1016/j.ssci.2023.106217
Breitsprecher K, Ndahbros S, Hinton JJ, Jacques P. Organizational ownership of an interdependent HSE culture yields quantifiable HSE performance. Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment 2014: The Journey Continues. 2014.3:1834-41. doi: 10.2118/168556-ms
Cambell Institute. Practical guide to leading indicators: Metrics, case studies and strategies. 2015.
Cao HC, Goh YM. Analyzing construction safety through time series methods. Front. Eng. Manag. 2019.6(2):262-74.
Chen Y, McCabe B, Hyatt D. Impact of individual resilience and safety climate on safety performance and psychological stress of construction workers: A case study of the Ontario construction industry. Journal of Safety Research. 2017.61:167-76. doi: 10.1016/j.jsr.2017.02.014
Choe S, Yun S, Leite F. Analysis of the effectiveness of the OSHA steel erection standard in the construction industry. Safety Science. 2016.89:190-200. doi: 10.1016/j.ssci.2016.06.016
Coetzee E, Govender U, Ndeunyema P, Genc B, Maré Y, Roux J, et al. An integrated safety framework for the diamond mines: A case study from Namibia. Resources Policy. 2023.82doi: 10.1016/j.resourpol.2023.103564
Dadashi Haji M, Behnam B, Sebt MH, Ardeshir A, Katooziani A. BIM-based safety leading indicators measurement tool for construction sites. International Journal of Civil Engineering. 2023.21(2):265-82. doi: 10.1007/s40999-022-00754-9
Dennerlein JT, Weinstein D, Huynh W, Tessler J, Bigger L, Murphy L, et al. Associations between a safety prequalification survey and worker safety experiences on commercial construction sites. Am J Ind Med. 2020.63(9):766-73. doi: https://doi.org/10.1002/ajim.23143
Doherty BD, Fragu LP. Sustainable HSE performance: Successful management systems and monitoring tools in the Middle East LNG industry. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2010 2010; 1445-63. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-77954222208&doi=10.2523%2f126823-ms&partnerID=40&md5=97447423e90f1c5ef8902bec5bc03942
Gale BR, Trostel MV, Armitage DL, Mason MA, Bautista LS. Case study: Successful implementation of driving safety and IVMS program. In: Proceedings - SPE Annual Technical Conference and Exhibition 2011; 4699-704. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84856696835&doi=10.2118%2f147521-ms&partnerID=40&md5=c31c637c6ccd53b5ac0196a93915abe9
Govender U, van Eck G, Genc B. An integrated 4Cs safety framework for the diamond industry of Southern Africa. Resources Policy. 2022.77doi: 10.1016/j.resourpol.2022.102774
Grabowski M, You Z, Song H, Wang H, Merrick JRW. Sailing on friday: Developing the link between safety culture and performance in safety-critical systems. IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans. 2010.40(2):263-84. doi: 10.1109/tsmca.2009.2035300
Haas EJ, Connor BP, Vendetti J, Heiser R. A case study exploring field level risk assessments as a leading safety indicator. Trans Soc Min Metall Explor. 2018.342:22-28. doi: https://doi.org/10.19150/trans.8104
Hinze J, Hallowell M, K B. Construction-safety best practices and relationships to safety performance. Journal of Construction Engineering and Management. 2013.139(10)doi: http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0000751
Lagerstrom E, Magzamen S, Kines P, Brazile W, Rosecrance J. Determinants of safety climate in the professional logging industry. Safety. 2019.5(2)doi: 10.3390/safety5020035
Laitinen H, Päivärinta K. A new-generation safety contest in the construction industry - A long-term evaluation of a real-life intervention. Safety Science. 2010.48(5):680-86. doi: 10.1016/j.ssci.2010.01.018
Laitinen H, Vuorinen M, Simola A, Yrjänheikki E. Observation-based proactive OHS outcome indicators - Validity of the Elmeri+ method. Safety Science. 2013.54:69-79. doi: 10.1016/j.ssci.2012.11.005
Lingard H, Hallowell M, Salas R, Pirzadeh P. Leading or lagging? Temporal analysis of safety indicators on a large infrastructure construction project. Safety Science. 2017.91:206-20. doi: 10.1016/j.ssci.2016.08.020
López JC, Zidan A. Step change in driving performance: A case study. In: SPE Latin American and Caribbean Health / Safety / Environment / Social Responsibility Conference 2013: Sustainable Solutions for Challenging

Reference
HSSE Environments in Latin America and the Caribbean 2013; 7-16. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84889800624&partnerID=40&md5=eae9b24d3d13b962603d4aaa82818ea8
Manjourides J, Dennerlein J. Testing the associations between leading and lagging indicators in a contractor safety pre-qualification database. National Occupational Injury Research Symposium. 2019.84.
Merrick JRW, Grabowski M. Decision performance and safety performance: A value-focused thinking study in the oil industry. Decision Analysis. 2014.11(2):105-16. doi: 10.1287/deca.2014.0291
Mohamed E, Jafari P, Kang SC, Pereira E, AbouRizk S. Leading indicators for safety management: Understanding the impact of project performance data on safety performance. In: Proceedings, Annual Conference - Canadian Society for Civil Engineering 2019. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85080721364&partnerID=40&md5=dfa1ec14a3bbea6fc1b3c02450d2fdda
Moore LL, Wurzelbacher SJ, Chen IC, Lampl MP, Naber SJ. Reliability and validity of an employer-completed safety hazard and management assessment questionnaire. Journal of Safety Research. 2022.81:283-96. doi: 10.1016/j.jsr.2022.03.005
Mousavi SS, Khani Jazani R, Cudney EA, Trucco P. Quantifying the relationship between lean maturity and occupational health and safety: Antecedents and leading indicators. International Journal of Lean Six Sigma. 2020.11(1):150-70. doi: 10.1108/ijlss-04-2018-0043
Pereira E, Guo X, Soleimanifar M, Siu MFF, Jeddry V, AbouRizk S. Effectiveness of applying a behavior-based safety program in industrial modular construction. In: 6th CSCE-CRC International Construction Specialty Conference 2017 - Held as Part of the Canadian Society for Civil Engineering Annual Conference and General Meeting 2017 2017; 545-53. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065087408&partnerID=40&md5=bb14108d019fbb687d627a2e77016b26
Quaigrain RA, Issa MH. Comparative analysis of leading and lagging indicators of construction disability management performance: An exploratory study. International Journal of Construction Management. 2023.23(7):1205-13. doi: 10.1080/15623599.2021.1963921
Rajendran S. Enhancing construction worker safety performance using leading indicators. Practice Periodical on Structural Design and Construction. 2013.18(1):45-51. doi: 10.1061/(asce)sc.1943-5576.0000137
Robson LS, Ibrahim S, Hogg-Johnson S, Steenstra IA, Van Eerd D, Amick BC. Developing leading indicators from OHS management audit data: Determining the measurement properties of audit data from the field. Journal of Safety Research. 2017.61:93-103. doi: 10.1016/j.jsr.2017.02.008
Sá JC, Dinis-Carvalho J, Fraga H, Lima V, Silva FJG, Bastos J. The impact of lean on occupational safety in organisations. In: IFIP Advances in Information and Communication Technology, 2023. 184-92
Salas R, Hallowell M. Predictive validity of safety leading indicators: Empirical assessment in the oil and gas sector. Journal of Construction Engineering and Management. 2016.142(10)doi: 10.1061/(asce)co.1943-7862.0001167
Schiavi AR. Using safety audits as a leading indicator. In: ASSE Professional Development Conference and Exposition 2013 2013. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85058534061&partnerID=40&md5=ee20744c2835df000fb8d580e0af55f9
Sheehan C, Donohue R, Shea T, Cooper B, De Cieri H. Leading and lagging indicators of occupational health and safety: The moderating role of safety leadership. Accid Anal Prev. 2016.92:130-38. doi: 10.1016/j.aap.2016.03.018
Stough J. From research to practice: A story of an actionable safety leading indicator index. In: SPE Latin American and Caribbean Health / Safety / Environment / Social Responsibility Conference 2013: Sustainable Solutions for Challenging HSSE Environments in Latin America and the Caribbean 2013; 17-28. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84889772506&partnerID=40&md5=72c03e5237b62a050bcfaf2c1bf4b2a6
Tang DKH, Leiliabadi F, Olugu EU, Md Dawal SZB. Factors affecting safety of processes in the Malaysian oil and gas industry. Safety Science. 2017.92:44-52. doi: 10.1016/j.ssci.2016.09.017
Tang DKH, Md Dawal SZ, Olugu EU. Actual safety performance of the Malaysian offshore oil platforms: Correlations between the leading and lagging indicators. Journal of Safety Research. 2018.66:9-19. doi: 10.1016/j.jsr.2018.05.003
Tauseef A, Villegas M, Bordage P, Turner L. Behavior management: A successful approach. In: Society of Petroleum Engineers - SPE/APPEA Int. Conference on Health, Safety and Environment in Oil and Gas Exploration and Production 2012: Protecting People and the Environment - Evolving Challenges 2012; 640-46. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84875131920&doi=10.2118%2f156673-ms&partnerID=40&md5=2e19fe161f2495309ff0d4ef49680bd7
Thananan T. Target zero: The challenge of achieving sustainable safety excellence. In: Society of Petroleum Engineers - SPE International Conference on Health, Safety and Environment 2014: The Journey Continues 2014; 1502-10. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84905860517&partnerID=40&md5=dd74b15370a33270a05b4aafc6036da8
Van Derlyke P, Marín LS, Zreiqat M. Discrepancies between implementation and perceived effectiveness of leading safety indicators in the US dairy product manufacturing industry. Safety and Health at Work. 2022.13(3):343-49. doi: 10.1016/j.shaw.2022.04.004

Reference
Versteeg K, Bigelow P, Dale AM, Chaurasia A. Utilizing construction safety leading and lagging indicators to measure project safety performance: A case study. Safety Science. 2019.120:411-21. doi: 10.1016/j.ssci.2019.06.035
Vosoughi S, Chalak MH, Yarahmadi R, Abolghasemi J, Alimohammadi I, Anbardan AN, et al. Prioritization and assessment of safety key performance indicators in an automotive industry. Sigurnost. 2021.63(4):347-61. doi: 10.31306/s.63.4.1
Wachter JK, Yorio PL. A system of safety management practices and worker engagement for reducing and preventing accidents: An empirical and theoretical investigation. Accid Anal Prev. 2014.68:117-30. doi: 10.1016/j.aap.2013.07.029
Wei L, Yang R, Chen Y, Shahi A, Safa M, Hanna A, et al. Comparison of safety cultures and performances between the construction industries in the United States and Canada: A case study of Texas and Ontario. In: Construction Research Congress 2020: Safety, Workforce, and Education - Selected Papers from the Construction Research Congress 2020 2020; 346-55. Available from: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85096919148&doi=10.1061%2f9780784482872.038&partnerID=40&md5=b59832b629892be6996b7197b1dd15fb
Winge S, Albrechtsen E, Arnesen J. A comparative analysis of safety management and safety performance in twelve construction projects. Journal of Safety Research. 2019.71:139-52. doi: 10.1016/j.jsr.2019.09.015
Zahoor H, Chan APC, Utama WP, Gao R, Zafar I. Modeling the relationship between safety climate and safety performance in a developing construction industry: A cross-cultural validation study. IJERGQ. 2017.14(4)doi: https://doi.org/10.3390%2Fijerph14040351

Appendix G: Detailed Results Tables

Table G.1: Study methods

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
Alarcón 2016 [38]	Cross-sectional	NR	To detect the best performing combinations of practices or strategies for different sizes of companies.	Multiple (comparison of frequencies, tornado diagram, classification tree method)	NR	NR
Amir-Heidari 2017 [28]	Case series	Framework development and case study	To review the HSE key performance indicators and frameworks for measurement of HSE performance and present a comprehensive classification of key performance indicators based on three factors (time, scope, and type).	Comparison of frequencies	NR but data reported for the past 5 years	NR
Bitar 2018 [25]	Cross-sectional	NR	To explore the relationship between operating discipline and a range of safety outputs including personal safety, process safety and plant reliability.	Multiple regression	September to November 2015	NR
Brandt 2023 [27]	Cohort study	National cohort surveys	To investigate the scientific gap concerning the importance of safety climate at work for the risk of LTSA in the general working population as well as in occupations characterized by shorter education and/or physically demanding work.	Comparative risk (hazard ratios)	Data used was gathered between 2012 and 2018	The Danish Working Environment Research Fund for this project (Arbejdsmiljøforskn ingsfonden, grant number 20195100758).
Breitsprecher 2014 [22]	Before-after study	NR	To describe a case study of the Africa region of an oil company following a two-year period of implementing the HSE leadership programme.	Comparison of frequencies	NR (leadership academies took place between 2010 and 2011, outcomes are reported for 2010, 2011 and 2012)	NR
Campbell Institute 2015 [23]	Case study	NR	To identify and define key leading indicators and to investigate how they are put into practice.	Linear correlation (NR) and qualitative comparison of frequencies	Data gathered between October 2013 and April 2014	NR

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
Cao 2019 [51]	Case study	NR	To use time series methods to identify projects with a high risk of accidents based on leading indicators.	Linear correlation (Pearson's)	Data collected between 2008 and 2015	NR
Chen 2017 [32]	Cross-sectional	NR	To examine the impact of safety climate and individual resilience on physical safety outcomes and job stress of construction workers in the Canadian construction industry.	Modelling (structural equation modelling)	Data collected between July 2015 and July 2016	Ontario Ministry of Labour ROP (Grant number 13-R-047).
Choe 2016 [24]	Before-after study (interrupted time series)	Interrupted time series	To evaluate the effectiveness of the revised OSHA standard 29 Code of Federal Regulations 1926 steel erection subpart in preventing overall occupational injuries in structural iron and steel workers in the US construction industry.	Interrupted time series	NR	NR
Coetzee 2023 [41]	Case study	Case study	To examine the extent to which the application of the Integrated Diamond Safety Maturity Framework has assisted to enhance the safety of frontline workers as well as support growth in safety maturity.	Comparison of frequencies	2022	NR
Dadashi Haji 2023 [48]	Before-after study	Case study	To develop a safety leading indicator measurement tool considering the safety leading indicators.	Comparison of rates	2020	Open Access funding enabled and organized by CAUL and its Member Institutions.
Dennerlein 2020 [52]	Cross-sectional	Cross-sectional	To describe the development of a comprehensive organisational survey for construction contractors that captures the leading indicators associated with best practices of injury and illness prevention programs that can be used as part of a prequalification procedure. To examine the association between scores from the organisational survey and worker safety climate and worker injury rates by construction worksites and by a subcontractor.	Linear correlation (Pearson's) and multivariate regression (linear and log-linear)	Data collection January 2017 to August 2018	Supported by the Center for Construction Research and Training; NIOSH.
Doherty 2010 [12]	Before-after study	NR	To describe the HSE management systems and associated performance and monitoring tool adopted by one of RasGas.	Comparison of frequencies (qualitative)	NR when the data were compiled and analysed, but the data reported are from 1999 to 2010	NR

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
Gale 2011 [13]	Before-after study	NR	To assess the impact of a driving program on driving accidents.	Comparison of frequencies (qualitative)	2009 to 2011	NR
Govender 2022 [42]	Before-after study	NR	To validate the 4Cs safety framework for the diamond industry.	Comparison of frequencies (qualitative)	Leading indicator data gathered between January and May 2021. 12 months of Lagging indicator data gathered - no dates provided	Part of a postdoctoral research study.
Grabowski 2010 [50]	Case series	NR	To identify leading indicators through looking at the link between safety culture and performance and to see if they correlate to safety performance.	Linear correlation (Spearman's)	Data collection between 2005 and 2006	Supported by the American Bureau or Shipping and the Overseas Shipholding Group Inc (Not specified if this is financial support).
Haas 2018 [49]	Before-after study	Case study	To illustrate the methods needed to successfully implement a proactive risk management system through looking at a case study.	Comparison of frequencies (qualitative)	December 2015	NR
Hinze 2013 [39]	Cross sectional	NR	To evaluate the influence of safety practices on safety performance in the constructing industry.	Linear correlation (Kendall's)	NR	Sponsored by the Constructing Industry Institute.
Lagerstrom 2019 [43]	Cross-sectional	Survey	To determine the level of safety climate among the population of professional loggers in Montana and to identify the determinants of safety climate based on surveys of worker demographics, workplace factors, and musculoskeletal symptoms.	Multinomial logistic regression	NR	Supported in part by the NIOSH Mountain and Plains Education and Research Centre and NIOSH funded High Plains Intermountain Centre for Agricultural Health and Safety.
Laitinen 2010 [29]	Case series	Real-life long-term evaluation	To evaluate the effects of the contest and analyse the observed change in the working environment and working habits, and to analyse the accident figures for the Uusimaa region and other parts of the country.	Weighted scores	Data from 1990 to 2006 analysed	NR
Laitinen 2013 [30]	Case series	Cross-sectional	To study the validity of the Elmeri+ observation method in predicting the accident risk of a workplace.	Linear correlation (Pearson's)	2002 to 2004	NR

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
Lingard 2017 [5]	Case study	Retrospective analysis	To uncover causal relationships amongst different leading and lagging indications, quantify time dependencies among indications and measure magnitude and direction of indications over time.	Modelling (temporal using VAR)	Jan 2010 to Jan 2015	The analysis was supported by Australian Linkage Project Grant LP120200440.
López 2013 [14]	Before-after study	NR	To report the results of the RJMC after more than a year.	Comparison of frequencies	2010 to 2011	NR
Manjourides 2019 [53]	Cross-sectional	Cross-sectional	To provide the first test of the associations of safety management systems and other leading indicators with lagging indicators as collected by a prequalification company with over 2,000 construction companies.	Linear correlation (Spearman) and univariate regression (Poisson)	2015 (all data downloaded on 1st October 2015)	Funded in part by a grant from the Centre for Construction Research and Training and the NIOSH.
Merrick 2014 [54]	Cross-sectional	NR	To assess the relationship between assertions from different decision frames (organisational, vessel and crew member) and safety performance.	Stepwise regression	2005 to 2006	NR
Mohammed 2019 [15]	Case study	NR	To investigate and test the feasibility of using existing, yet under-utilized, project-related data together with safety-related data to assess proactive safety performance more accurately in industrial construction projects.	Linear correlation (Pearson's)	2016 to 2017	Collaborative Research and Development Grant (CRDPJ 492657) from the NSERC.
Moore 2022 [31]	Cross-sectional	NR	To evaluate the reliability and predictive validity of a safety management questionnaire in predicting workers compensation claims.	Linear regression	2012 to 2015 data set	This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Mousavi 2020 [44]	Cross-sectional	Deductive quantitative study	To develop a comprehensive model disentangling the relationship between lean and OHS, and to highlight the importance of using leading indicators in understanding and measuring the impact of lean on OHS performance.	Modelling (structural equation modelling)	Data gathered between April and July 2017	NR
Pereira, 2017 [16]	Before-after study	NR	To propose a data-driven framework that can assess the effectiveness of a behaviour-based safety program to improve safety performance.	Linear correlation (Pearson's) and paired sample T-tests	Program implemented in September 2014. Incident data gathered between January 2012 to October 2016 (incident rates between October	NSERC Industrial Research Chair Program (IRCPJ #195558-10).

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
					2013 and May 2015 were removed from data set due to limited projects in this time period.	
Quaigrain 2023 [55]	Cross-sectional	NR	To develop leading and lagging indicators to evaluate the disability management performance of construction organisations.	Linear correlation (Spearman's)	2012 to 2015	Supported by a grant from the Research and Workplace Innovation Program of the Workers Compensation Board of Manitoba.
Rajendran 2013 [56]	Case study	Case study	To evaluate selected leading indicators under real project conditions and develop recommendations for their use by construction contractors.	Linear correlation (Pearson's)	2011	NR
Robson 2017 [33]	Cross-sectional	NR	To answer research questions regarding (a) the ability of audit-based scores to predict workers' compensation claims outcomes, (b) structural characteristics of the data in relation to the organization of the audit instrument, and (c) internal consistency of items within audit elements.	Linear correlation	January 2007 and December 2010	Supported by a grant from the WSIB of Ontario to the lead author, as well as core support to the Institute for Work & Health from WSIB and the Ontario Ministry of Labour.
Sá 2023 [11]	Cross-sectional	NR	To determine whether, in organizations where Lean tools were implemented, if there was an improvement in occupational safety conditions, namely in the reduction of accident rates, and to verify which Lean tools contributed the most to that improvement.	Comparison of frequencies	7 August 2019 to 6 October 2019	"The work of the author Vanda Lima is supported by national funds, through the FCT under the project UIDB/04728/2020" - no other funding information reported.
Salas 2016 [40]	Case series	NR	To identify combinations of leading indicators that reliably predict lagging indicators of safety performance.	Multiple (multivariate regression and principal factor analysis)	2012 to 2014	NR
Schiavi 2013 [17]	Before-after study	NR	To describe the factors that led to the successful reduction of accidents and injuries at the company.	Comparison of frequencies	2002 to 2010	NR
Sheehan 2016 [37]	Cross-sectional	NR	To consider the association between leading and lagging indicators of OHS and to investigate the moderating effect of	Modelling (multi-level)	September 2013 to November 2014	NR

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
			safety leadership on the association between leading and lagging indicators.			
Stough, 2012 [18]	Cohort study	NR	To reveal measurable differences in the organisational factors which show how organisations treat risk reduction events.	Comparison of frequencies	2008 (reported that data has been analysed since 2008. however, the database is a 'multiyear' database so unclear when the analysed data was collected)	NR
Tang 2017 [26]	Cross-sectional	NR	Aimed at leveraging the experience of the health, safety and environmental personnel in identifying safety indicators which are most pertinent to the Malaysian offshore oil and gas sector.	Multiple (Comparison of frequencies and linear regression)	NR	NR
Tang 2018 [35]	Cross-sectional	NR	To identify key safety factors in evaluating and predicting the safety performance of oil and gas platforms.	Linear correlation (Pearson)	2016	The study was not subject to the funding of any grant.
Tauseef 2012 [19]	Before-after study	NR	To assess the impact of a behaviour-based safety program (observation intervention program) on reducing unsafe events.	Comparison of frequencies	2010 to 2012	NR
Thananan 2014 [20]	Before-after study	NR	To measure the level of SSHE culture, based on the International Association of Oil and Gas Producers, Safety Culture Maturity Model. The result provides a framework to assist in the selection and implementation of appropriate behavioural interventions in order to reduce accidents.	Comparison of frequencies	2011 to 2012	NR
Van Derlyke 2022 [47]	Cohort study	NR	To examine leading indicators in the dairy product manufacturing industry and evaluate their perceived effectiveness in reducing work related injuries.	Comparison of frequencies	Outcome data from 2013-2018. NR when questionnaire was carried out which gathered this data.	This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Versteeg 2019 [6]	Case study	Case study	To examine the impact of leading indicators (site inspections and toolbox talks) lead to a lower frequency of injuries.	Univariate regression (Poisson)	Data from 2012-2016	This research did not receive any funding from agencies in the public,

Study	Study design	Author reported study design	Study objective	Statistical method	Date of study	Details of funding
						commercial, or not-for-profit sectors.
Vosoughi 2021 [45]	Case series	Descriptive-analytic applied research	To identify, prioritise and assess key indicators for implementing an effective performance evaluation system in an automotive industry.	Linear correlation (Pearson's and Spearman's)	2009 to 2018	Research Deputy of Iran University of Medical Sciences for research grant scheme (No: IR.IUMS.FMD.REC 1396.9511139002).
Wachter 2014 [36]	Cohort study	NR	To develop ideas around a system of safety management practices and test their relationship with safety statistics e.g. accident rates.	Multiple (Linear correlation and regression)	2011 to 2012	Funded by a grant from the Alcoa Foundation.
Wei 2020 [21]	Cohort study	NR	To compare several dimensions of safety culture and performance and to uncover their shared and distinct characteristics.	Linear regression (Spearman's)	2015 to 2016	Financial support from the Government of Ontario’s Ministry of Labour ROP grant number 16-R-038.
Winge 2019 [34]	Case study	NR	To identify how safety management and contextual factors influence safety performance.	Qualitative comparative analysis	NR	This study is a part of a research project about construction safety funded by the Research Council of Norway and the Norwegian Labour Inspection Authority. Study carried out in cooperation with Statsbygg, a Norwegian government client organisation.
Zahoor 2017 [46]	Cross-sectional	Cross-cultural validation study	To validate a safety performance measurement model in the cross-cultural setting of a developing country and highlight the variations in investigating the relationship between safety climate factors and safety performance indicators.	Modelling (structural equation modelling)	March to June 2015	Fully funded by the International Postgraduate Scholarship of The Hong Kong Polytechnic University (Grant number FTE-IPS-RTLF-88011).

Abbreviations: CAUL – Council of Australian University Librarians, FCT – Portuguese foundation for science and technology, HSE – Health, safety and environment, NIOSH – National Institute for Occupational Safety and Health, NR – Not reported, NSERC - Natural Sciences and Engineering Research Council of Canada, OSH – Occupational safety and health, OSHA – Occupational safety and health administration, RJMC - Regional journey management centre, ROP - Research opportunities program, SSHE - Safety, security, health and environment, US – United States, VAR – Vector autoregression, WSIB - Work safety and insurance board.

Table G.2: Descriptions of leading and lagging indicators

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
New guideline, tool or process									
Breitsprecher 2014 [22]	Before-after study	Oil	HSE Leadership Academies	From the beginning of 2011 to the end of 2012, the 23 HSE leadership academies were conducted in the company's Africa region for senior and middle management.	TRinC and motor vehicle accident rate.	NR - units for both measures are missing (e.g. incident rate = incidents per 1,000 hours worked).	Lagging indicators compared before and after the leadership training academies.	Number of academies implemented and number of attendees (only number not proportion of senior and middle managers who attended reported).	Number of safety observation cards and number of Stop Work actions. Both leading indicators but measured here as outcomes, and could be considered intermediate outcomes, e.g. managers were more likely to make observations and give Stop Work actions which led to improved safety outcome.
Choe 2016 [24]	Before-after study (interru-pted time series)	Construction	Revision of the Steel Erection Standard	In 2002, OSHA revised the steel erection standard, which is designed to protect employees from steel erection related hazards.	Fatality rate, days away injury rate, normalised fatality rate and normalised days away injury rate.	Fatality rate: number of fatalities per 100,000 full-time equivalent construction structural iron and steel workers Days away	Interrupted time series comparing linear regression line of best fit for safety outcome frequencies before the revised steel erection standard to lines	Extracted from Bureau of Labor Statistics of the US Department of Labor.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
						injury rate: number of days away injuries per 10,000 full-time equivalent construction structural iron and steel workers Normalised fatality rate: relative fatality rate of construction structural iron and steel workers adjusted by a fatality rate of total construction workers Normalised days away injury rate: days away injury rate of construction structural iron and steel workers compared to a days away injury rate of construction workers in general.	of best fit for the following 5 years. Data from the US department of labour were used in the analysis.		

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
Coetzee 2023 [41]	Case study	Diamond mining	Implementation of the integrated 4C framework	A novel, innovative approach to safety maturity that embraces four Cs, viz. Culture, Competence, Cultivate and Connectedness (4Cs) in a single framework, underpinned by corporate values and focusing on people, systems, and processes. The journey commenced in 2017 with a baseline safety maturity review.	Loss of life, all injury frequency rate, total recordable case frequency rate, lost time injury frequency rate, MPI and HPI, MPH and HPH, lost time injuries medical treatment cases, first aid cases, days lost. MPI, HPI, MPH and HPH are not defined.	NR	Year-to-year trend in safety performance indicators following the implementation of the framework.	Initial implementa-tion plus monitoring of "safety maturity" (intermediate indicator).	SafeSentry observations (a front-line employee-driven risk discovery system, focusing on the continuous identification of randomly shifting risks), safety maturity, percentage of enabling verses disabling factors towards each of the 4Cs.
Dadashi Haji 2023 [48]	Before-after study	Construction	A tool integrating BIM and knowledge base which visualizes the impacts of the safety leading indicators on a daily basis, with an add-on to aid in integration and to facilitate the process of information exchange between the main components of the proposed system.	In the first step, a safety leading indicator knowledge base is built to encompass three parts: identification of active safety leading indicators, recognition of attributes related to leading indicators, and forming relationships between them. The safety leading indicators are dangerous	Accidents, injuries and fatalities.	The authors do not report the study definitions of the three outcomes, but they report the number of accidents, injuries and fatalities from before and after the tool was introduced and divide it by the number of stations built to give the number of	Comparison between projects completed before the framework was implemented and projects afterwards.	The use of the tool during the construction of 23 of the 35 stations.	A panel of experts reviewed the tool and rated different statements about the accuracy and effectiveness of the tool on a Likert scale between 1 and 5.

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				equipment, equipment proximity, worker movement, slips and leading edge, and the attributes are working at height, collapse, moving objects and vehicles, electricity, manual handling, harmful materials and noise. Then, the project's 3D model is generated inside the BIM environment, e.g., Autodesk Revit. Afterward, project activities are identified. Furthermore, the model in Revit is exported to Navisworks. In the next step, activities, and their features and attributes are defined. Attributes indicate whether this activity is vulnerable to a certain event like a fire or heavyweight, or not. Moreover, the add-on extracts model elements		accidents, injuries and fatalities per station before and after the implement-ation of the tool.			

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				and assigns their relative activities. In this step, the activities which affect each element through the period of the project, are assigned. In the fourth step, identified safety leading indicators, and their features and attributes are defined. Safety leading indicators' attributes indicate whether a safety leading indicator is capable of causing a certain event or not. In the final step, the defined safety leading indicators are allocated to the activities. To achieve this goal the add-on checks whether the safety leading indicators and the activities have matching attributes, or not. In addition, the impact of safety leading indicators on each element is investigated in this step. As a result, a heat map					

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				is generated which visualizes the effects of the safety leading indicators on the project status in time and location. Through this heat map, critical times and locations are recognized. Thus, project staff can use this information to select strategies to prevent accidents.					
Doherty 2010 [12]	Before-after study	Oil	RasGas Elements of Excellence	The company's operations integrity management system introduced in 2001 which integrates various HSE management systems within one framework. Examples of the HSE tools are: emissions reduction steering committee, heat stress prevention programme, interactive centre for safety training, behaviour-based safety programme, health risk and impact, e-incident	TRIR and also heat injury rate during Ramadan.	TRIR: NR. Heat Injury Rate during Ramadan: Heat-stress-related medical treatments per 200,000 work hours.	Before and after comparison.	Timeline of implementa-tion	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				reporting and tracking system, leak detection and repair programme and corporate waste management programme. The heat stress prevention programme included real-time monitoring, heat stress charts, heat index communication, heat stress working controls, heat stress awareness and fitness to work assessments.					
Gale 2011 [13]	Before-after study	Oil and gas	DrivingChange program	A system to identify leading indicators or risk and provide feedback to individual to influence their behaviours.	Motor vehicle incidents and field worker safety.	NR	Before and after the implementa-tion of DrivingChange.	Vehicle speed, acceleration, deceleration, idle time, miles driven and engine operating hours and engine operating parameters.	NR
Govender 2022 [42]	Before-after study	Mining	Integrated 4C safety framework	4C element (operational action) (shift towards improvement) Culture of leadership in people (care for	Loss of life, TRCFR, lost time injury frequency rate, lost time injury severity rate, all injury frequency rate,	Loss of life, TRCFR, lost time injury frequency rate, lost time injury severity rate, all injury frequency rate,	Before and after comparisons of the lagging indicators.	Not much information provided the framework was measured.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				all) (people trust and respect each other) Competence of people (can do work safely) (people can identify hazards and work collaboratively) Connectedness of people and systems (couple safety in every task((integration and leveraging of people, risk and operational efficiencies as part of core planning) Cultivate value in and by the people (commit to improve) (continuous creation of value in the company and people through adopting a futuristic view with mitigation of risk and discovery opportunity).	high potential incident.	high potential incident.			
Haas 2018 [49]	Before-after study	Mining	Field-level risk assessment program	Risk assessment matrix completed by workers to help identify and evaluate risks. Associated with heightened	Incidents	Non-fatal days lost injuries	No statistical tests conducted so the relationship between leading and lagging indicator scores	Workers completed a risk assessment matrix to help identify and evaluate risks.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				engagement, leadership and behaviour change associated with risks.			can only be assessed by visual examination of the different scores.		
Laitinen 2010 [29]	Case series	Construction	Implementation of a contest based on the standardised TR-observation method.	The contest took place from 1997 to 2000, was not held in 2001 but was then restarted on a permanent basis. The winners are determined based on the average TR-index at the company's sites (60% of score), safety plans and annually changing criteria at individual site level (30% of score), accident incidence rate for the company, based on data reported by the company (10% of score). Although the accident rate is a lagging indicator, it is the implementation of the contest which is a leading indicator in this study. The index scores are also a leading indicator	Accidents per cubic meter of construction.	Predicted accident risk per cubic meter of construction compared to actual accident risk per cubic meter of construction in the time period following the implementa-tion of the contest.	Number of accidents per cubic meter after implementation of contest compared to previous years.	The number of accident rates before the contest, predicted number of accidents and actual number of accidents since the implementa-tion of the contest.	None. As reported in leading indicator column, the TR-index is also reported but is only evaluated against the implementa-tion of the contest (another leading indicator).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				and are reported here but are evaluated against the implementation of the contest (another leading indicator) rather the accident rates (the lagging indicator).					
López 2013 [14]	Before-after study	Oil and gas	Implementation of a RJMC	The RJMC is a centralised centre with specialist hardware, application software, tele-communications technology and multilingual staff who manage land transport activities in multiple countries with a diversity of risk environments. They work with local management to provide practical risk assessments before a journey starts, track journeys in real time, apply on-time positive and formative interventions, and supply managers with comprehensive	Automotive Accidents CMS (industry recognised) AARM-CMS (industry recognised) Total rollover (company and contractors) Automotive Accidents CMSL (industry recognised) AARM-CMSL (industry recognised).	"Automotive Accidents CMS (industry recognised)" authors do not define report what CMS (presumably Catastrophic, Major, Serious) stands for or what is meant by industry recognised. "AARM-CMS (industry recognised)" authors do not report what this stands for, presumably Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious "Total rollover (company and contractors)"	Accident rates before and after the RJMC was implemented.	Comparison of company data in 2010 to 2011.	Change in other leading indicators (over speeding, fatigue, % of trips with trained drivers, unauthorised drivers' trips, unauthorised drivers' night driving, comparison of numbers of positive and formative observations).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				reports that they can use to better understand the behaviour of their drivers, and provide recognition and coaching to them in response to their performance.		not defined but presumably refers to automotive accidents in which the car rolled over "Automotive Accidents CMSL (industry recognised)" authors do not define report what CMS (presumably Catastrophic, Major, Serious, Light) stands for or what is meant by industry recognised. "AARM-CMSL (industry recognised)" authors do not report what this stands for, presumably Automotive Accident Rate (in Miles) and Catastrophic, Major, Serious, Light.			
Pereira 2017 [16]	Before-after study	Construction	Behaviour-based safety program.	Observation reports filled out by workers prior to performing construction tasks	TRincR and total incidence rate.	TRincR and TIR.	Comparisons of TRincR and TIR before and after implementation of the	Workers completed an observation report consisting of	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				to eliminate potential hazards. These are categorised into 4 report types: behaviour observation, at-risk, near-miss and improvement opportunities.			behaviour-based safety programme using T values and correlation coefficients with p values and effect sizes.	53 safety requirements.	
Sá 2023 [11]	Cross-section-al	Various	Implementation of Lean tools	Lean tools are not defined within the paper.	Accident rates	Incidence index, frequency index and severity index (none of these are defined by the authors).	Whether accident rates increased, decreased or remained the same following the implementation of Lean tools.	Self-reported through questionnaire.	NR
Schiavi 2013 [17]	Before-after study	Newspaper publishing	"Resurrected safety process"	The company partnered with Dupont Safety Services who provided the benchmarking tools that identified the key elements where attention was needed. They looked at twelve safety areas as compared to world class organizations. Each element was ranked in five groups from fundamental to world class. The analyses	Percentage reduction in total work-related accidents, lost time accidents and Musculo-skeletal diseases.	NR	Before and after the changes were made to the safety process.	Comparison of accidents and injuries before and after the changes to the process were made.	Worker's compensation direct costs (only reported in 2010, not in change over time).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				reviewed: visible management commitment, working safety policy, integrated organization for safety, line organization responsibility and accountability, aggressive safety goals, high standards of performance, supportive safety personnel, progressive motivation, comprehensive accident investigation, effective two way communications continuous safety training, and safety auditing. Boston Globe decided to concentrate on four of the elements: safety audits, accident investigation and management accountability and communications New processes were not implemented without the support of the					

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				unions. Safety audits were tied to 25% of raises and bonuses for supervisors. At least four safety audits per month are conducted by supervisors and foremen using a very simple one-page safety audit form on the company electronic database.					
Tauseef 2012 [19]	Before-after study	Oil	Observation intervention program.	A behaviour-based safety program consisting of three main elements: accountability, risk management and communication. Observations are made of safe and unsafe behaviours to stop unsafe acts and encourage correct procedure.	Total number of recordable injuries per million man-hours and automotive accident rate.	TRIFR: total number of recordable injuries per million man-hours. Automotive accident rate: number of catastrophic, major and serious automotive incidents per million miles driven.	Before and after comparisons of lagging indicators.	Through observation reports.	None
Thananan 2014 [20]	Before-after study	Gas	Step Change roadmap.	Follows the 4S approach (Study, Story, Segment, Strategic). Consists of the following elements: leadership and commitment,	Lost time injury frequency, total recordable injury rate and process safety events.	Lost time injury frequency, total recordable injury rate and process safety events.	SSHE performance reductions after the Step Change program.	Survey to measure the level of SSHE culture.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				policy and strategic objectives, organisation/ recourses and documentation, evaluation and risk management, planning and operational control, implementation and monitoring, audit and review.					
Safety climate/culture									
Brandt 2023 [27]	Cohort study	Mixed	Safety climate	A short version of NOSACQ-50 that consisted of five selected single items with vital aspects of the safety climate concept. The five single item questions are indicative of the safety climate, addressing the main themes concerned in the literature: managerial (questions 1 to 3) and employee commitment, participation and engagement (questions 4 to 5).	LTSA.	Registered sickness absence in the Danish Register for Evaluation of Marginalisation for a period of at least 6 consecutive weeks for a period of up to 2 years, starting the week after replying to the survey.	Comparison between different working populations and occupational groups using hazard ratios.	A shortened version of the NOSACQ-50 questionnaire.	NR
Chen 2017 [32]	Cross-section-al	Construction	Safety climate and individual resilience.	Safety Climate: The shared perception of	Physical safety outcomes and	Physical safety outcomes: physical	Correlations between all variables were	Survey on demographics, attitude	Psychological stress: for example, loss

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				people toward safety in their work environment. Individual resilience: The capacity of individuals to cope with significant change, adversity or risk.	psychological stress.	symptoms (cuts, burns, sprains etc) and unsafe events (slips, overexertion, exposure to chemicals) Reported based on experiencing the issue at least once in the previous three months.	carried out. Structural model built to examine impact of safety climate and individual resilience on physical safety outcomes and psychological distress.	statements and incident reporting.	of sleep due to work worries, feeling under strain, unable to concentrate on work tasks.
Lagerstrom 2019 [43]	Cross-section-al	Logging	Safety climate	Safety climate score from the NOSACQ-50 administered to loggers attending a safety workshop.	Presence of Musculo-skeletal symptoms or missed work due to Musculo-skeletal symptoms.	Self-reported Musculo-skeletal symptoms.	Safety climate dimension scores reported by presence or absence of Musculo-skeletal symptoms and multinomial logistic regression to determine odds ratio of having Musculo-skeletal symptoms depending on having a 'fairly good' or 'low' score on each dimension of the safety climate assessment.	NOSACQ-50 survey.	Scores were also correlated against several demographic factors in table 4.
Wei 2020 [21]	Cohort study	Construction	Safety culture factors, as well as number of projects in previous 3	The second section of the questionnaire was 40 questions with a Likert-scale (1-	Physical injuries and unsafe events.	Respondents also reported the number of physical injuries,	Correlation between leading and lagging indicators. Comparison	Likert scale on a questionnaire.	Psychological stresses measured in the same way as the lagging

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			years and work hours per week.	5, strong disagree to strongly agree) assessing nine safety culture factors: safety consciousness (workers are aware of relevant safety issues), safety program (safety programs are useful and clear), management commitment (management is committed to safety), supervisor safety (supervisors have good safety attitudes and behaviours), coworker safety (coworkers always act safely), job involvement (the job is an important part of my life), fatalism (individuals have little control over their personal safety), role overload (excessive work is assigned to workers) and interpersonal conflict at work (workers		unsafe events and psychological stresses they experienced in the last three months (0=never, 1=once, 2=two to three times, 4=four to five times, and 5=more than five times).	between Ontario and Texas.		indicators. The authors also looked at the relationship between demographic results and safety performance. The authors also evaluated the differences between the two regions for all of the indicators.

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				experience frequent conflict at work). Mann-Whitney test found a statistically significant difference (≤0.05) between score in Texas and Ontario in supervisor safety, fatalism and role overload.					
Zahoor 2017 [46]	Cross-section-al	Construction	Safety climate. Also 2 of the 3 elements of safety performance although their association with accidents is not a primary objective of the study.	Respondents gave a level of agreement on at 1-5 Likert scale for 24 safety climate statements from a validated questionnaire. The 24 statements covered four factors: management commitment and employees' involvement in health and safety, safety enforcement and promotion, applicability of safety rules and safe work practices and safety consciousness and responsibility. Also, the two	Safety performance' included three broad indicators of which one was eligible for this review: number of self-reported accidents/injuries and near-misses in past 12 months.	As the reliable accident statistics were not available, self-reported accident statistics were collected. The questionnaire included the four questions on accidents in the last 12 months: how many times have you exposed to a near-miss incident of any kind at work? How many times have you suffered from an accident/injury of any kind at work, but did NOT require	Correlation.	Likert scale on a questionnaire.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				elements of 'safety performance' other than accidents were safety compliance and safety participation measured in the same was as accidents (lagging indicator).		absence from work? How many times have you suffered from an accident/injury, which required absence from work NOT exceeding three consecutive days? How many times have you suffered from an accident/injury, which required absence from work exceeding three consecutive days? They were measured on a 5-point Likert scale (ranging from no accidents to over five accidents).			
Audits and inspections									
Laitinen 2013 [30]	Case series	Mechanical engineering, metal industry and electronics industry	Elmeri+ score	The Elmeri method includes 14 observable items. It was developed as an easy and simple	Accidents/10^6 working hours, blue collar accidents/10^6 working hours	Accidents/10^6 working hours: Number of accidents causing at least 1 day	Correlations.	An observer would go to the worksite and complete the	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				standard occupational health and safety inspection method that can be used both by safety inspectors and company personnel. The observer divides the workplace into small workstations and observes all or a random sample of them. Each item is scored as correct or incorrect and the final score is the percentage of correct items. The sub-indexes are: 1) order of floors, platforms, walkways and driveways, 2) Worker safety behaviour, 3) Air quality & use of chemicals, 4) Order of worktables, shelves and racks for tools and materials, 5) Waste container, 6) lighting, 7) noise, 8) Design of workstation and work posture, 9) Condition of	and lost hours/blue collar worker.	absence from work/10^6 working hours Blue collar accidents/10^6 working hours and lost hours/blue collar worker: Data were gathered for both blue collar and white collar workers and white collar workers were excluded from these outcomes.		observation form.	

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				machines & machine guards, 10) Structure of floors, walkways, fall protection, 11) Thermal conditions, 12) Control devices of machines, 13) Emergency exits, 14) Musculo-skeletal load.					
Robson 2017 [33]	Cross-section-al	Various (manufacturing, retail and wholesale trade, other services, construction, government and related services).	Audit score	Firms with 20 or more employees and higher than expected claim costs and rates were audited as part of the WorkWell programme in Ontario. If a firm scored 75% or lower, a second audit was scheduled for several months later.	Lost time claims and non-lost time claims.	Lost-time claim: when a worker suffers a work-related injury/disease that leads to being off work past the day of the incident, a loss of earnings, or a permanent disability/impairment. No-lost-time: when no time is lost from work, other than on the day of an incident, but health care is required.	Correlation coefficients with a p value <0.05.	Observation using a standardised audit instrument. Overall audit results were expressed as a percentage of possible points.	NR
Monitoring of safety performance									
Moore 2022 [31]	Cross-section-al	Multi-industry (agriculture, forestry and fishing;	Safety hazards and management practices	SH-26 by the Ohio Bureau of Worker's Compensation	Worker's compensation claims	Claim rates	Correlations between the SH-26 assessment items and regression	Data gathered through the SH-26 questionnaire which	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
		construction; healthcare and social assistance; mining; services; transporta-tion, warehousing and utilities; and wholesale and retail trade).					coefficients between SH-26 item ratings and worker compensation claims.	assessed OHS elements in the workplace.	
Mixed									
Alarcón 2016 [38]	Cross-section-al	Construction	Different combinations of prevention management practices. 221 practices were identified and the authors categorised them into 7 categories: Accidents & Incidents Investigation, Safety Planning & Resources, Management Commitment, Workers' Safety Training, Management Safety Training, Audits & Certifications, Safety	1 Accidents & Incidents Investigation: activities related to the capture of information of accidents and incidents. 2 Safety Planning & Resources: activities carried out by safety staff (such as the preparation of safety plans) as well as activities related with safety equipment that workers should use. 3 Management Commitment: activities that demonstrate the willingness and commitment to	Accident rate	Number of accidents / average labour force	Visual and descriptive analyses and classification tree. Comparison between companies with different practices.	Practices identified from construction companies in the database which the authors sorted into the 7 categories for analysis.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			Incentives & Rewards.	safety from management, which otherwise would not be carried out. 4 Workers' Safety Training: activities such as courses, workshops, seminars, and all kind of safety training for workers. 5 Management Safety Training: similar to the previous group, but focused on the company management. 6 Audits & Certifications: regular activities performed by the Safety Mutual. 7 Safety Incentives & Rewards: all kinds of recognition for good safety records.					
Amir-Heidari 2017 [28]	Case series	Oil and gas drilling	A composite score of 12 leading indicators.	The 12 indicators are: costs assigned to HSE for preventing accidents (per each worker per year, and with respect to total expense), percentage of	A similar composite scale of 8 indicators described by the authors as lagging indicators, of which 4 meet this reviews	The 4 indicators are: rate of occupational accidents which lead to fatality (number per 200,000 man-hours), rate of	No statistical tests conducted so the relationship between leading and lagging indicator scores can only be assessed by visual	A written guideline was prepared and provided for the companies, and they were asked to report the defined	The other 4 indicators included in the lagging indicator composite score are: frequency of process accidents

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				workers who are not qualified or competent for assigned duties, percentage of completion of activities related to risk assessment and control in the planned times, number of reported near misses (per year), percentage of existing of exact, applicable and updated work procedures for doing the works in a standard way, frequency of audits and inspections related to HSE management system, percentage of completion of activities related to technical inspection and quality assurance in determined times, percentage of costs/funds assigned to environmental risk management (per year with respect to the total	definition of a lagging indicator.	recordable occupational accidents (number per 200,000 man-hours), rate of occupational illness/health problem reports (number per 200,000 man-hours), and amount of legal fines/costs related to HSE (per each 100 m of drilling). No further description of any of the indicators is given.	examination of the different scores reported in Table 6.	indicators using the scoring guide tables, for five years. Each indicator was weighted differently in the composite score calculation. Indicators were weighted and prioritised using Analytic Hierarchy Process.	which lead to a hazardous material release, amount of waste discharged to environment (per each 100 m of drilling), amount of oil spill/discharge to environment (per each 100 m of drilling), amount of greenhouse gas release to atmosphere (per each 100 m of drilling).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				expense), amount of costs/funds assigned to personnel health risk management (per each worker per year, and with respect to total expense), percentage of works which are performed under supervision of experienced and skilled supervisors, percentage of hot and cold works which are performed under permit to work system and number of red risks (in risk profile). No further description of any of the indicators is given.					
Bitar 2018 [25]	Cross-section-al	Oil and gas	Operating discipline communication, operating discipline implementation, leadership expectations and trust index.	Operating discipline is defined by the authors as "a deeply rooted dedication and commitment by each member of the organisation to carry out each task the right way, each time". The organisation	Recorded injury frequency and near miss frequency (both Control of Work related and not).	Recordable injury frequency: number of recordable injuries (a work-related injury that requires medical treatment beyond first aid, as well as	Correlation.	A survey sent to all staff on operation sites.	Environmental and efficiency outcomes were also measured (loss of primary containment frequency, plant reliability, operating efficiency and

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				introduced a framework of operating discipline called 3Fs (Follow the rules, Finish what you start, Follow up). Operating discipline communication is a measure of effectiveness of communication and engagement of the 3Fs and operating discipline implementation is a measure of the implementation of the 3Fs. The survey questions measuring leadership expectation consisted of six questions relating to care for safety among team members, being able to speak out without fear of reprisal, being able to stop unsafe work, and managers' behaviours relating to safety and risk, and operating discipline. The		one that causes, loss of consciousness , restricted work or transfer to another job, days away from work or death, not including fatalities) to employees and contractors for every 200,000 hours worked Near miss frequency: number of near misses (an incident that did not, but had the potential to, affect the health, safety or security of people, or assets, or the environment. Potential severity would be at the minimum a First aid case or environmental damage to a non-sensitive environment),			plant efficiency).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				survey questions measuring trust index consisted of four questions relating to perceptions of trust in the company and its management, pride in working for the company, and leaders listening to all perspectives.		for every 200,000 hours worked Control of work near-miss: near misses with a potential severity at or below a Day Away From Work Case or an impact to the environment, categorised as control of work related. Non-control of work near miss: near misses with a potential severity below the same threshold, categorised as not related to control of work.			
Campbell Institute 2015 [23]	Case study	NR	Training hours, safety observations, incident investigation, site audits, leadership engagement.	Proactive, preventative and predictive measures that monitor and provide current information about the effective performance, activates and proc esses of an EHS	Incident rates, accident rates, hazard rates, near miss reporting and stop-work authority.	NR	Correlations and before and after comparisons.	Through guided group discussions and interviews with companies which provide case studies for analysis.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				systems that drive the identification and elimination or control of risks in the workplace that can lead to incidents and industries. They are designed to gibe advance warning of potential problems so preventative actions can be taken.					
Cao 2019 [51]	Case study	Construction	Project delay, project man hours, PPE, overhead protection, excavation work, machine safety guarding, safe means of access, operating crane/lifting, scaffold, tower/mobile scaffold, mech elevated work platform, falling hazard/opening, electrical hazard, first aid facilities, emergency preparedness, handling and storage of hazardous	Factors that may affect construction safety.	Accident rates.	Number of accidents in each project per month.	Correlation between accident rates and different leading indicators.	NR - just stated that project related data was collected.	The indicators listed in table 5 (other than accidents) that do not meet the project definition of a leading indicator (project progress).

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			materials, safe work procedures, power tool safety, earth control measures, noise/ vector and others.						
Dennerlein 2020 [52]	Cross-section-al	Construction	ACES	ACES designed by the study authors to assess construction contractors' safety performance. The survey focuses on these leading indicators: "upstream root causes of workplace incidents and worker injury and illness" including "organizational policies, programs, and practices that monitor, control, and/or eliminate hazards in the workplace".	OSHA reportable injuries, injuries resulting in DART.	Collected from primary contact person for each site (the site's safety manager or project manager) 12 to 18 months after the worker survey and the ACES survey data collection to assure that the building project was completed.	For the correlation analyses the authors calculated Pearson correlation coefficients for worksite average ACES scores and subcontractor ACES scores with worksite and subcontractor level safety climate scores, injury rates, and DART rates. For the regression analysis, regression models to estimate associations between ACES scores and safety climate (linear multivariable	ACES survey focussing on leading indicators was developed by the authors to assess safety performance. The questions each had a 5-point Likert response and scores were averaged across workers for each worksite and each subcontractor.	Worker safety climate, collected via worker surveys. This review would consider this a leading indicator, but in this study the leading indicator of interest, the ACES survey, was assessed against safety climate as well as injury rates.

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
							regression) and injury and DART rates (log-linear multivariable models) at the site level were fitted.		
Grabowski 2010 [50]	Case series	Marine transporta-tion	Safety factors	Organisational (hiring quality people, safety orientation, formal learning system, promotion of safety), individual (empowerment, anonymous reporting, individual feedback).	Safety performance	Number of accidents, number of incidents, number of near losses, number of port state deficiency (vessel factor), number of lost time injuries >3, number of conditions of class (vessel factor).	Correlation between safety performance of three shipping industry partners.	Surveys.	Vessel safety factors (Leading=com munication, problem identification, vessel feedback and vessel responsibility) (Lagging=con ditions of class, port state deficiency).
Hinze 2013 [39]	Cross-section-al	Construction	Implementation of various safety practices (for example, health and safety manual, safety prequalification, subcontractors’ participation in general contractors' orientation and training, subcontractors’ safety standard compared to general contractor's,	Safety practices or programs.	Recordable injury rate.	Recordable injuries sustained per 200,000 hours of worker exposure.	Correlation coefficients (with p values) of the relationship between safety practices and RIR.	Questionnaire delivered to project representa-tives.	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			safety leadership training for foremen,).						
Lingard 2017 [5]	Case study	Construction	Management activity (toolbox meetings, pre-brief meetings/ pre-start meetings, safety observations, site surveillance inspections carried out, penalties/infringements, occupational health and safety audits, non-compliances, hazards reported, hazards closed out, statutory authority inspections carried out, alcohol tests, drug tests, safe work method statements/ JSA documents review and amended, site inductions.	Predicted indicators of safety in the constructing project.	TRIFR.	The number of workplace injuries normalised per million hours worked per year.	Modelling.	Data obtained from routine reporting throughout the project - not stated exactly what was reported or for what purpose.	NR
Manjourides 2019 [53]	Cross-section-al	Construction	Safety Management System, Safety Program Elements	Safety Management System: 17 survey questions related to safety culture,	Recordable injury cases and injuries involving days away,	Both calculated per 100 full-time equivalent person-hours	Correlation and multivariate modelling.	Scores calculated from relevant survey questions in	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			Hazards, Safety Program Elements Programs, Special Elements Non-drug and alcohol, Special Elements Drug & Alcohol Screening and OSHA Citations.	communication, and incident investigations, such as “Does your company have a full-time safety manager on staff?,” and “Does your company have a defined budget for safety?” Safety Program Elements Hazards: 16 survey questions pertaining to potential hazards that workers may be exposed to on a worksite, e.g. using ladders, working in trenches, or using power tools Safety Program Elements Programs: 17 questions regarding programs addressing specific safety practices related to worksite hazards, such as eye protection or fall protection programs Special Elements Non-drug and	restricted, or transferred.	billed from data reported on Construct-Secure by a safety manager.		the contractor safety assessment program of self-reported and validated company-level safety data. Registered and active companies who subscribe to the Construct-Secure service have a safety manager complete an annual survey capturing measures of safety management systems, special programs, and special elements as well as lagging indicators including OSHA recordable injuries and Experience Modifying Rate.	

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				alcohol: 4 questions covering a contractor's return-to-work programs, OSHA partnerships, and participation in OSHA's Safety and Health Achievement and Recognition Program and Voluntary Protection Programs Special Elements Drug & Alcohol Screening: 4 questions pertaining to drug and alcohol testing policies OSHA Citations: number of citation over the last 3 years self-reported by companies and categorised as 0, 1, 2 or >3.					
Merrick 2014 [54]	Cross-section-al	Oil/energy transporta-tion	Organisational and crew member decision frames.	Organisational (objectives: hiring quality people, safety orientation, formal learning system, promotion of safety), vessel (objectives: prioritisation of safety, crew	Safety performance	Number of accidents	Stepwise regression analysis with adjusted R^2, p values and F ratios. Estimates, T ratios and p values for each objective	Questionnaire developed and sent to workers to score different statements about performance in different	NR

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				responsibility, problem identification, communication, crew feedback) and individual (objectives: empowerment, individual responsibility, anonymous reporting, individual feedback). Each objective was measured by asking respondents to rate several assertions (e.g. "My colleagues consider safety issues seriously while performing their job duties") from strongly agree to strongly disagree.			assertion (and some combinations of assertions) are also provided.	decision frames.	
Mohammed 2019 [15]	Case study	Construction	Project performance safety and scheduling data.	Project data on years of experience, foreman hours, shift hours, exposure hours, incident and work order human performance index.	Accidents	Accident occurrence	Correlation coefficients and 'importance' ratings through Boruta analysis.	Measured through project performance data.	Non- safety project performance data (cost, quality, schedule) was also correlated against accident rates but these variables are not safety leading

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
									indicators. All listed in table 1.
Mousavi 2020 [44]	Cross-section-al	Multi-industry (health and safety, engineering, operation / production, human resources, consulting, education, other).	Lean maturity	Lean maturity: fidelity, extensiveness, experience.	OHS performance.	Recordable injuries, worker's compensation cost, accident rates, total lost working days.	R^2 evaluations, multi-group analysis, moderation analysis and mediation analysis Comparisons between company size and sector are made.	Survey on lean approaches implemented and OHS performance.	Working environment, task characteristic, workforce characteristic, organisational factors.
Quaigrain 2023 [55]	Cross-section-al	Construction	CDM3	The CDM3. Used to assess the maturity of multiple practices of disability management by rating the practices.	Recordable Injury Rate, Lost Time Case Rate and Severity Rate.	NR	Correlation coefficients with a p value <0.05 are reported (not significant results are not reported).	Assesses the maturity of multiple practices of disability management by rating the practices on a Likert scale from 1 to 5.	Disability management metrics. 12 metrics, a mixture of leading and lagging indicators.
Rajendran 2013 [56]	Case study	Construction	Worker safe behaviour observation, pretask plan and site safety audits.	Worker safe behaviour observation: An observation (of workers behaviour) was performed by 7 professionals associated with the case study project. The behaviours were rated as safe or unsafe. The professionals then scored the	Near miss incident rates, first aid injury rates, OSHA recordable injury rates, and total incident/injury rates.	Near miss incident rates: Near miss incidents are defined as those incidents considered a close call, an incident that resulted in almost an injury or property damage First aid injury rate: It	Strength and direction of Pearson's correlation.	Each was scored between 0 and 100 by the project team/safety professionals.	There were no other indicators measured but Pretask Plan Review and Worker Safe Behavior Observation Score had a strong correlation.

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				observation from 0 to 100 where the score is the ratio of workers with safe behaviour to the total number of workers observed multiplied by 100. An example of an unsafe behaviour would be a worker exposed to a 12.2 m fall without proper fall protection. Pretask plan: Project safety professionals reviewed the pretask plans and rated them as either adequate (score=100) or inadequate (score=0). An example of an inadequate pretask plan would be steps to complete the task not being identified Site safety audits: Conducted by the project team, audit results were scored from 0 to 100 to reflect the safety of the construction site.		includes all injuries that resulted from an exposure or event in the workplace and that required some first aid treatment OSHA recordable injury rates: It includes all OSHA recordable incidents, which are defined as those incidents that resulted from an exposure or event in the workplace and that required some type of medical treatment beyond first aid Total incident/injury rates: It includes all the project incidents (OSHA recordable injuries, first aid case injuries, and			

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				Each safety site audit was started with a score of 100. One point was deducted for a minor safety violation and two points were deducted for major safety violations.		near miss incidents), which are defined as those incidents that resulted from an exposure or event in the workplace (total incident/injury rates=near miss incident rates + first aid injuries + TRIR).			
Salas 2016 [40]	Case series	Oil and gas	Contractor safety data	Near miss reporting, safety observation, stop work authority, client audit, JSA development, job safety audit engagement, contractor project management engagement and walkthroughs, contractor safety rep engagement and walkthroughs, client safety walk throughs, contractor safety audits, subcontractor safety audits, project orientation sessions, client participation in	TRincR and Severity rate	TRincR: multiplying the number of recordable injuries in a year by 200,000 (100 employees working 2,000 h per year) and dividing this value by the total work hours worked in the year Severity rate: ratio of lost days experienced as compared to the number of incidents experienced.	R^2 values, p values and SE.	Measured using contractor safety data.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				project orientation sessions, corrective action items, risk driven action items, project risk assessment, risk driven project risk management plan, daily tasks orientation, client participation in daily tasks orientation, contractor post job evaluation score, contractor evaluation score, experience modifiers rate.					
Sheehan 2016 [37]	Cross-section-al	Various (Arts and Recreation Services; Construction; Electricity, Gas, Water and Waste Services; Health and Community Services; Mining; and Transport, Postal and Warehousing)	Aggregate OPM-MU score and safety leadership aggregate score	The OPM-MU is an 8-item scale that has been reported to be a reliable and valid measure of leading indicators of occupational health and safety. Employees were asked to report on their perceptions of the workplace they worked in most often, rather than the organisation overall, using a 5-point scale (ranging from 1=strongly	Reported occupational health and safety incidents, unreported occupational health and safety incidents, and near misses	Self-reported by the survey respondents who were asked to report the number of occupational health and safety incidents they had been personally involved in at work over the past 12 months. The measures were sourced from Probst et al., 2013	Inter-correlations and multi-level regression analysis.	OPM-MU	No other indicators reported, but interactions between leading indicator score and safety leadership score were not extracted because these are both considered leading indicators

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				disagree to 5=strongly agree), according to the extent to which they agreed or disagreed with the eight statements (e.g., Formal OHS audits at regular intervals are a normal part of our workplace). Safety leadership (the operations manager safety leadership scale developed by Wu et al. 2010 is a 12-item scale that measure aspects of safety leadership, e.g., I objectively analyse the causes of injuries, I visit the workplace to assess safety, and I encourage employees to be safe in their working behaviour. Middle-level and line managers in each workplace were asked to report on their perceptions of their own safety leadership, using a 5-point scale					

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				(ranging from 1=strongly disagree to 5=strongly agree), with respect to the 12 statements)					
Stough 2012 [18]	Cohort study	Energy	Operating assets and proactive activities	Operating assets: Facilities, management systems, human factors Proactive activities: Audits, inspections, management system assessments and observations	Total recordable injury rates and severity-weights total recordable injury rates, incidents, near misses, investigations and corrective actions	NR	Comparison of safety outcomes between companies with different safety characteristics implemented.	Data obtained from companies which gather safety data as a way to manage quality, health, safety and environment performance.	None
Tang 2017 [26]	Cross-section-al	Oil and gas	Safety performance	Safety factors: inspection and maintenance, management and work management on safety, number of incidents and near misses, personal safety, constrictor's safety, management of plant changes, plant operation and operating procedures, competence, plant design, instrumentation and alarms, hazard identification and risk assessment,	Incident occurrence	Incidents are culmination of risks that often go unheeded due to defective monitoring mechanism	Z values, R^2 values, p values.	Questionnaire in which respondents rated the perceived importance of the provided safety indicators.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				documentation, start-ups and shutdown, emergency management					
Tang 2018 [35]	Cross-section-al	Oil and gas	Safety performance	Safety factors: inspection and maintenance, emergency management, management and work management, number of incidents and near misses, personal safety, constrictor's safety, management of change, operation and operating procedures, competence, hazard identification and risk assessment, plant design, instrumentation and alarm, start-ups and shutdown	Fatality, fatal accident rates, TRincR, LTIR and reported near-misses	Fatality is defined as death, either immediate or within one year of the date of injury, of an employee or a contractor's employee due to work, while fatal accidents are accidents resulting in fatality. Total recordable incidents encompass all fatalities, lost time injuries, illnesses, and medical treatment cases occurring at work but do not include first-aid injury. Lost time injury results in inability of an employee to continue work, hence a loss of productive	Correlation coefficients and p values. Safety scores obtained from the proposed framework are compared against the findings of facility status reports of two oil platforms in to examine how close the findings from both instruments are in terms of the platform safety performance.	Measured through a questionnaire gathering data on safety performance.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
						work time. Near-misses, on the other hand, are unintended occurrences that could potentially harm human, the environment, and properties. Incident and injury rates are counted as occurrences per million man-hours worked. In the case of TRincR for instance, it is counted as number of total recordable cases per million man-hours worked			
Van Derlyke 2022 [47]	Cohort study	Dairy product manufacturing	Safety audits, preventative maintenance, safety training attendance, safety observations, safety inspections, near-miss reporting, stop work authority, JHA/JSA, safety	Statements that can describe and monitor safety conditions	OIR and DART rates	Measurements of successes	Kendall's W, p values. Comparisons between companies who did implement all 8 identified leading indicators and those that did not.	Surveys to workers to assess how they are implemented.	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
			meeting attendance, corrective action completion rate, worker perception survey, and attendance tracking						
Versteeg 2019 [6]	Case study	Construction	Toolbox talks, number of site inspections and number of near misses	Toolbox talk: Short onsite training session that occurs on regular basis to educate workers Site inspection: Walk arounds completed by superintendent or safety representative to look for hazards using a standardised checklist. Near misses: An indicator of potential risk that relies on judgment and reporting	Number of medical injuries, number of first aid injuries	A medical injury is a workplace injury that requires medical intervention, while a first aid injury is a workplace injury that can be treated on site using first aid	Standard errors, confidence intervals, p values.	Through routinely collected project admin data.	None
Vosoughi 2021 [45]	Case series	Automotive	Safety indicators and educational indicators.	Safety indicators: Number of injury per employee, number of near misses, percentage of corrected non-compliance, number of risk assessments conducted,	Total number of work-related LTI, frequency severity index, percentage of total number of work-related LTI reduction compared to previous year	No definition provided.	Percentages, correlation coefficients.	During interviews, experts were asked to evaluate how 'necessary, relevant, clear and simple' they believed the indicators to be.	Frequency of similar events in the same process, average cost per injury, percentage of budget allocated for risk management

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
				number of manoeuvres carried out, safety climate. Educational indicators: man hour training new employees, percentage of trained supervisors within a year, percentage of trained managers in a year, percentage of workers who have been trained in emergency procedures, number of safety educational courses provided, percent of job training programs implemented, number of e-learning programs, percentage increase in training hours					
Wachter 2014 [36]	Cohort study	Multi-industry (named: agriculture, construction, transportation and distribution, education,	Safety management practices	Employee involvement/ influence, pre and post task safety population, safe work procedures, hiring for safety, cooperation facilitation, safety	Accident rates	total recordable cases, DART cases	Correlation coefficients, regression coefficients, p values.	Survey asking respondents to rate the degree to which each safety management practice was implemented	None

Study	Study design	Industry	Leading indicator	Description of leading indicator	Lagging indicator	Description of lagging indicator	Method of evaluating impact	Measurement of leading indicator	Other measures of effectiveness
		government, healthcare, light manufactur-ing, mining, research and develop-ment, service)		training, communication and information sharing, accident investigation, detection and monitoring, safe task assignment				in their organisation on a Likert scale of 1 (strongly disagree) to 7 (strongly agree).	
Winge 2019 [34]	Case study	Construction	Safety practices	Construction complexity, organisational complexity, time, economy, contract management, OHS planning, roles and responsibilities, project management, management commitment to OHS, safety climate, learning, performance population, operative risk management, site management, staff management, hardware management	Safety performance	LTIR, TRIR, Medical treatment injuries, registered unwanted occurrences and site deviations	Coincidence matrix, consistency values. Comparative analysis between 12 projects.	Assessed through safety and health plans, inspection and audit reports, logs of OHS-related information and interviews with OHS coordinators and project managers.	None

Abbreviations: ACES – Assessment of contractor safety, CDM3 – Construction disability management maturity model, DART – Injuries involving days away, restricted or transferred, HSE – Health, safety and environment, JHA – Job hazard analysis, JSA – Job safety analysis, LTI – Lost time incidents, LTIF – Lost time injury frequency LTIR – Lost time incident rate, NR – Not reported, OIR – Occupational incident rate, OPM-UM – Organisational performance metric – Monash University, OSHA – Occupational safety and health administration, PPE – Personal protective equipment, RIR – recordable injury rate, RJMC - Regional journal management centre, TR – Talonrakennus (Residential construction), TIR – Total incident rate, TRCFR – Total recordable case frequency rate, TRIR – Total recordable injury rate, TRincR – Total recordable incident rate.

Table G.3: Worker characteristics

Study	Workers	Eligibility criteria	Key characteristics
Alarcón 2016 [38]	NR	All construction companies in the database were included in the preliminary analysis but only those with 30 or more records were included in subsequent databases	NR
Amir-Heidari 2017 [28]	NR	NR	NR
Bitar 2018 [25]	3,514 (5,533 responses of which 3,514 were used in the analysis). Survey responses came from employees in offshore and onshore operations as well as offices, but only those working at operation sites were included in the analysis	For the analysis: to be a worker at an operation site	n from each country: Angola: 242 Azerbaijan: 980 Georgia: 214 Indonesia: 847 Norway: 145 Trinidad: 131 UK: 397 USA: 558
Brandt 2023 [27]	63,500 workers (managers [n=3,138], professionals [n=19,759], technicians and associate professionals [n=8,152], clerical support workers [n=5,271], service and sales workers [n=9,353], skilled agricultural/ forestry/ fishery workers [n=356], craft and related trades workers [n=5,275], plant and machine operators and assemblers [n=4,040], elementary occupations [n=4,715], military work [n=421])	People who were currently wage earners (not self-employed), aged 18 to 64, employed for minimum of 35 hours a month, income of at least 3,000 DKK per month in the past three months, wage earners without LTSA during the 52 weeks prior to surveys being carried out	Age (mean, SD): 46.2, 10.8. 48.5% male Higher education : 43.3% Vocational education or less: 56.8% BMI (kg/m2) (mean, SD): 25.8, 4.4 Physical activity during leisure (hours per week) (mean, SD): 5.2, 3.3 Smoking: Daily: 15.1%; once in a while: 5.2%; ex-smoker: 29.1%; never: 50.6% Major depression inventory (0 to 50) (mean, SD): 8, 7.3
Breitsprecher 2014 [22]	NR (338 senior and middle managers attended the sessions, rates are reported for the whole region but number of workers in the region NR)	Only senior and middle managers were included in the HSE leadership academies. All employees in the region were included in the outcomes data	NR
Campbell Institute 2015 [23]	NR	NR	NR
Cao 2019 [51]	NR	Projects with complete and sufficient data.	NR
Chen 2017 [32]	783 construction workers (837 surveys completed but 54 excluded from the analysis due to >10% missing data)	NR	98% male. Mean (SD, range) age=37.11 (12, 16 to 67). Average weekly working hours=44.24, 37% worked over 44 (overtime). High safety training percentage (97.7%). 38.1% had experience of being a safety committee member. 60.7% were in a union. Job roles

Study	Workers	Eligibility criteria	Key characteristics
			were supervisor (31.3%), journeyman (50.5%), apprentice (18.2%)
Choe 2016 [24]	Structural iron and steel workers and structural metal workers.	NR	NR
Coetzee 2023 [41]	The whole workforce (>900 employees).	NR	NR
Dadashi Haji 2023 [48]	NR	NR	NR
Dennerlein 2020 [52]	1,426 workers completed the surveys. Only apprentices, journeymen, and foremen completed the worker survey (not upper management positions).	An injury/incident reporting system. The project duration of at least 6 months. A project budget of at least $5 million.	Workers completing the surveys from these sites were mostly male (94%), white (72%), and members of unions (88%).
Doherty 2010 [12]	NR but it's reported in the introduction that at its peak the company had >35,000 workers onsite.	NR	NR
Gale 2011 [13]	Drivers (n=NR)	NR	NR
Govender 2022 [42]	NR	NR	NR
Grabowski 2010 [50]	Employees working on the ships (shipboard n=1,599, shoreside n=157)	NR	Shipboard (Mean, SD): Age=40.85 (11.12), experience in current employer=5.26 (5.47), experience in industry=14.87 (10.17) Shoreside (Mean, SD): Age=42.55 (10.38), experience in current employer=7.16 (8.79), experience in industry=17.47 (12.30)
Haas 2018 [49]	Over 450 workers (exact number NR).	NR	NR
Hinze 2013 [39]	NR	Recently completed or ongoing construction objects with over 50,000 worker hours reported.	Ranged in size from 70,000 to 20,000,000 worker hours expended and contract values between $51,000 and $3,600,000,000. Project types were commercial buildings (20%), industrial facilities (30%), infrastructure (13%), energy facilities (23%), institutional (4%), heavy civil (3%), health care (4%), and marine projects (3%).
Lagerstrom 2019 [43]	743	All loggers attending the safety workshops on emergency first-aid in Montana were invited to participate.	Mean (SD, range) age: 45.88 (13.67, 15 to 79) years Mean (SD, range) years in the logging industry: 21.96 (14.11, 0 to 55) 274 (48.1%) reported musculoskeletal symptoms in any body region and 34 (6.0%) reported lost work due to musculoskeletal symptoms 284 (38.2%) were an owner or supervisor (412 [55.5%] were not, 47 [6.3%] NR) 174 (23.4%) were an accredited logging professional (414 [55.7%] were not, 155 [20.9%] NR)

Study	Workers	Eligibility criteria	Key characteristics
			Highest educational level achieved: some high school: 55 (7.4%), high school diploma awarded: 410 (55.2%), some college: 179 (24.1%), bachelor degree or higher: 59 (7.9%), NR: 40 (5.4%) Primary logging system type: conventional: 80 (10.8%), mechanical: 408 (54.9%), both 28 (3.8%), NR; 227 (30.6%)
Laitinen 2010 [29]	NR	"Building sites where the main contractor is a state institution, municipality, small private company or private individual (one-family houses) are outside the scope of the study".	NR
Laitinen 2013 [30]	n=23,399 (n=16,176 of which were "blue collar workers" - blue collar is not defined by the authors).	Only companies with at least 2 years of accident figures and Elmeri+ results from the years 2002 to 2004 were eligible.	70% "blue collar workers" (proportion was highest in the metal industry and lowest in the electronics industry).
Lingard 2017 [5]	NR	NR	NR
López 2013 [14]	Drivers (n NR)	NR	NR
Manjourides 2019 [53]	NR	Data as available and complete from companies with unexpired records on the ConstructSecure database as of 1st October 2015.	NR
Merrick 2014 [54]	915 shipboard employees	NR	Mean age (SD, range): 37.53 (9.27, 20-69) Mean experience in current employer (SD, range): 5.35 (5.74, 0-37) Mean experience in industry (SD, range): 12.85 (8.57, 0-41)
Mohammed 2019 [15]	Construction workers (n NR)	NR	NR
Moore 2022 [31]	2,295 employers	The main participant group consisted of employees insured by Ohio Bureau of Workers' Compensation who completed the SH-26 questionnaire 2 or more times between 2012 and 2015 and experienced at least one compensation claim. A larger group of employees had a less restrictive inclusion criteria - they completed the SH-26 questionnaire at least one and had any number of claims.	NR
Mousavi 2020 [44]	112 survey responses (worker details NR)	Surveys with over 15% of data missing were excluded.	Key characteristics of respondents (but there was only one respondent from each company, so the characteristics do not summarise the characteristics of the workers): Age: 18-24 (n=6), 25-34 (n=27), 35-44 (n=30), 45-54 (n=27), 55-64 (n=16), >65 (n=6) Gender: Male (n=88), female (n=24). Education level: high school (n=5), BS (n=34), MS (n=56), PhD (n=17)

Study	Workers	Eligibility criteria	Key characteristics
Pereira 2017 [16]	Construction workers (n NR)	NR	NR
Quaigrain 2023 [55]	NR	NR	NR
Rajendran 2013 [56]	NR	NR	NR
Robson 2017 [33]	NR	Audit "end-date" between 1st January 2007 and 15th December 2010. Results from a 3rd or 4th audit were not eligible. Results from a 2nd audit with no corresponding results from the 1st audit were not eligible.	WorkWell audits are for companies with 20 or more employees and "poor safety records".
Sá 2023 [11]	NR	Organisations that responded to the questionnaire and reported that they had implemented Lean tools.	NR
Salas 2016 [40]	Oil and gas industry contractors (261)	NR	Average work hours provided by contractors=2.82 million (2012-2013) and 3.05 million in 2014 TRincR ranged from 0.00-8.79 (2012-2013) and 0.00-6.78 in 2014. SR values ranged from 0.00-2.20 in 2012-2013 and 0.00-3.39 in 2014. A majority of the contractors provided construction (20%), mechanical (17%) and environmental (16%) types of work to clients.
Schiavi 2013 [17]	n=>2,200	NR	The company has 12 unions. Supervisors are also in a specific union although they are considered management.
Sheehan 2016 [37]	n=3,578 employees	Respondents were deleted if they did not identify as belonging to a specific workplace as they could not be included in the multi-level analysis.	24% middle managers or line managers, 76% worked in other roles 77% were employed on a continuing basis 57% were working full-time 61% male 52% had been working for their organisation for five years or less
Stough 2012 [18]	NR	NR	NR
Tang 2017 [26]	172 health, safety and environmental professionals	Ideally have offshore work experience (either technical or administrative).	NR
Tang 2018 [35]	NR	NR	NR
Tauseef 2012 [19]	NR	NR	NR
Thananan 2014 [20]	Employees and contractors	NR	NR

Study	Workers	Eligibility criteria	Key characteristics
Van Derlyke 2022 [47]	Workers	NR	Length of safety responsibilities: > 1 year (n=5), 1-2 years (n=11), 3-5 years (n=24), 6-10 years (n=26), > 11 years (n=2) Safety and Health Certifications held: None(n=57), Certified Safety Professional (n=13), Graduate Safety Practitioner (4), Associate Safety Professional (3), Others (3), Occupational Hygiene and Safety Technician (1), Safety Management Specialist (1), Certified Safety Director (1), Multiple Certifications (1)
Versteeg 2019 [6]	Workers	The construction contact was either fixed bid, design build or construction management. The project was completed with no additional collection of administrative data. The project was recorded constantly across departments.	NR
Vosoughi 2021 [45]	11 experts in the automotive industry	"Experts with sufficient experience".	Age:30-40 (n=8), <40 (4) Level of education: MS (8), associate professor (4) Work experience: 7-12 (5), >17 (3) Type of work: Auditor (4), OHS officer (4), academic staff (4)
Watcher 2014 [36]	Study 1: mean number of employees per establishment was 632 (208,560 employees*) Study 2: 650	NR	NR
Wei 2020 [21]	n=587 construction workers	Surveys with significant missing data were removed.	Mean age (SD): Texas: 35.13 (10.34), Ontario: 37.40 (11.25) (Mann-Whitney test shows a statistically significant, ≤0.05, difference between the 2 regions) Mean years in construction (SD): Texas: 13.09 (9.71), Ontario: 14.34 (10.33) Mean years with current employer (SD): Texas: 5.63 (6.13), Ontario: 6.07 (6.45) Mean number of projects in previous 3 years (SD): 17.25 (25.31), Ontario: 9.50 (21.82) (Mann-Whitney test shows a statistically significant, ≤0.05, difference between the 2 regions) Mean work hours per week (SD): Texas: 49.21 (11.94), Ontario: 43.55 (7.01) (Mann-Whitney test shows a statistically significant, ≤0.05, difference between the 2 regions) Gender (male): Texas 96.8%, Ontario 97.9%
Winge 2019 [34]	NR	Projects with high or low safety performance. Projects similar in size, building type and contractual arrangements.	NR

Study	Workers	Eligibility criteria	Key characteristics
		Projects that were finished or more than halfway finished.	
Zahoor 2017 [46]	n=426 Frontline worker=85 (19.95) Foreman=26 (6.1) Supervisor=58 (13.62) (37 supervisors, 21 surveyors) Site engineer=82 (19.25) Construction manager=98 (23) (55 construction managers, 26 resident engineers) Safety official=77 (18.08) (31 safety officers and 46 safety inspectors).	Employees at an under-construction multi-storey (at least 70 metres) building projects in Pakistan. Incomplete and "unengaged" survey responses were not included in the analysis.	Age (years) - ≤20: 93 (21.83) 21-30: 105 (24.65) 31-40: 94 (22.06) 41-50: 79 (18.55) 51-60: 43 (10.09) >60: 12 (2.82) Education level - Below primary: 21 (4.93) Primary: 32 (7.51) Middle: 41 (9.62) Secondary: 17 (3.99) Diploma: 135 (31.69) Degree or higher: 180 (42.25) Type of employer/organisation - Client/owner: 77 (18.08) Main contractor: 88 (20.66) Subcontractor: 133 (31.22) Consultant: 86 (20.19) Academia: 42 (9.86) Service with current employer - <1 year: 174 (40.85) 1-5 years: 213 (50) 6-10 years: 24 (5.63) 11-15 years: 10 (2.35) >15 years: 5 (1.17) Work experience in the construction industry - <5 years: 133 (31.22) 6-10 years: 81 (19.01) 11-15 years: 106 (24.88) 16-20 years: 68 (15.96) >20 years: 38 (8.92)

Abbreviations: BMI – Body mass index, BS – Bachelor of science, DKK – Danish Crone, HSE – Health, safety and environment, LTSA – Long-term sickness absence, MS – Master of science, NR – Not reported, PhD – Doctor of philosophy, SD – Standard deviation, SH-26 – Safety management self-assessment questionnaire, TRincR – Total recordable incident rate, UK – United Kingdom, US – United States.

Appendix H: Summary of Brandt 2023

Figure H.1: Summary of Brandt 2023 [27]

Study design: Cohort study.

Objective: To investigate safety climate at work as a predictor for the risk of long-term sickness absence in the general working population.

Methods: Data from the Work Environment and Health in Denmark Study (WEHD) on 63,500 workers was combined with long-term sickness absence data from the Danish Register for Evaluation of Marginalisation (DREAM) and analysed. This WEHD data included demographic details as well as responses to a shortened version of the NOSACQ-50 safety climate survey. The survey version of the Cox proportional hazard model was then used to calculate hazard ratios of long-term sickness absence for different numbers of safety climate problems.

Results: "The presence of safety climate problems increases the risk of LTSA [long-term sickness absence] in the general working population and could be seen as an early warning sign to invest in promoting OSH in the workplace."

Strengths and weaknesses: Population characteristics of the included workers were reported, though limited data were provided on sampling and recruitment methods. As a cohort study, this was one of the few comparative studies identified in this review. Despite sparse reporting of some methods, it was judged to be at low risk of bias because potential confounders were identified and controlled for using weighted Cox regression, data were collected through appropriate methods (the survey used has been shown to be valid and reliable and national registry data was also used), and appropriate statistical tests were used in data analysis. Furthermore, the large sample size and inclusion of workers from any industry mean the results are likely to be applicable to workplaces across Denmark.