Foundation

EVIDENCE REVIEW AND SYNTHESIS METHODS GUIDE

For the Global Safety Evidence Centre

Dr Elizabeth Adjoa Kumah, Ingrid Abreu Scherer, and Nancy Hey

October 2026

With expert input from Prof. Mikela Chatzimichailidou, Dr. Adam Cooper, Prof. Kevin Daniels, Ruth Frankish and Prof. Arnab Majumdar.



Photo by Fahrul Azmi

This document will be reviewed and updated in October 2027

Contents

1. Introduction 5
1.1 Our evidence priorities 5
1.2 High quality evidence 6
1.3 Working with the Centre 7
1.4 The project consultation group 7
1.5 About this guide 7
1.6 How to use this guide 8
2. Reviewing the global safety evidence base 8
2.1 Our approach to evidence synthesis 10
2.3 Types and sources of evidence 10
2.5 Equity considerations in evidence reviews 11
3. Reviewing the evidence 16
3.1 Types of reviews and syntheses 16
3.2 Selecting the right review approach 21
4. Stages of a review 23
5. Formulating a research question 24
5.1 Involving practitioners in developing research questions 24
5.2 Frameworks for developing review questions and inclusion criteria 25
5.3 Other limitations for inclusion 26
6. Developing a review protocol 28
6.1 Registering the review protocol and recording changes 28
6.2 What to include in a review protocol 29
7. Search strategies and sources for safety topics 30
7.1 Developing and testing search strings 30
7.2 Searching for contested concepts 33
7.3 Selecting bibliographic databases for safety topics 33
7.3 Grey Literature and call for evidence 34
7.4 Documenting the search strategy 35

7.4 Additional searches 36
7.5 Using software in searches 36
8. Study selection 37
8.1 Screening and selecting studies 37
8.2 Duplicate publications of the same study 38
8.2 Using Artificial Intelligence tools 38
8.3 Documenting the study selection 39
9. Data extraction 41
9.1 What data to record for each study 41
10. Quality assessment 42
10.1 The quality of included studies 42
10.1.1 Assessing external validity of included studies 44
10.2 Assessing the certainty of review findings 44
10.2.1 Applying GRADE to quantitative synthesis findings 45
11.2.1 Applying GRADE-CERQual to qualitative synthesis findings 47
11. Synthesis 49
11.1 Narrative synthesis 49
11.2 Meta-analysis and quantitative synthesis 50
11.2 Summarising the findings in a table 50
11.4 Developing evidence statements 51
11.3 Developing a theory or model based on the evidence 52
12. Reporting 54
12.1 Using PRISMA reporting standards 54
12.2 Executive summary 55
12.3 Illustrating and contextualising the findings 55
13. Communicating the findings 56
15. Living reviews of safety evidence 58
14. Case Study Synthesis 60
Appendix A: Standards and manuals by review type 63

Appendix B: Assessing the quality of included studies 66
Appendix C: PRISMA 2020 checklist 67
Appendix D: Case Study Synthesis quality framework and practitioner template 71
D.1 Quality of included studies framework 71
D.2 Case Study Template for practitioners 73
References 76
1. Main handbooks and guides 76
2. Other references 76

1. Introduction

> *Evidence is critical to improving safety – but often does not yet exist or is not easily accessible.*
>
> *Across the world and across different sectors, there is a huge opportunity to improve safety outcomes by generating better quality evidence on both the scale and nature of the challenges, and on what works to address them. But simply generating that evidence is not enough – it must also be relevant, understandable, accessible and actionable by those in a position to put it into practice.*
>
> *(Engineering a safer world: Lloyd's Register Foundation Strategy 2024-2029)*

Lloyd's Register Foundation is a global safety charity with a mission to engineer a safer world. We do this by harnessing our heritage to shape a safer and more sustainable ocean economy for the future and finding and sharing the best evidence and insight on what works to improve safety.

The Foundation's Global Safety Evidence Centre (the Centre) was established in 2025 as a comprehensive hub for anyone who needs to know 'what works' to make people safer. The Centre works with a wide range of institutions, teams and practitioners and with global practitioner bodies and international organisations to ensure that the most important research questions are answered in the best ways and in a timely manner.

1.1 Our evidence priorities

The Centre collates and communicates the best safety evidence from the Foundation (including the World Risk Poll), our partners and other sources on both the nature and scale of global safety challenges, and what works to address them.

We focus on safe work, particularly in 'high hazard' industries, but our remit is not limited to occupational safety. We also create and collate evidence on safety systems and processes, particularly in relation to the maritime and critical infrastructure sectors that are the focus of much of the Foundation's work. We support the development of safety science and ways of understanding safety outcomes and effectiveness of safety interventions.

We want to work on understanding and solving safety challenges, that is: foreseeable, solvable or preventable safety issues that are global in nature and/or scale.

Our funding supports independent research projects including evidence synthesis, evaluation and trials, data analysis, and exploring concepts and indicators. We translate and share evidence in accessible and actionable forms, and we work with policy makers, practitioners and partners to identify topical areas of research interest.

The Centre's audiences include:

- **Practitioners:** anyone who can use our evidence in their work to support safe work, including employers, managers and supervisors, HR professionals, regulators, policy makers and campaigners. Practitioners may work in industry or government, public or private sectors, charities and communities.
- **Researchers:** anyone who generates evidence which can be used by others in their work, including university academics and students, evaluators and research consultants.
- **Lloyd's Register Foundation**: teams around the Foundation use our evidence to make strategic funding and other decisions across our portfolios.

1.2 High quality evidence

As a trusted global source of evidence on safety, we support research which is:

1. **Robust and credible:** using tested methods, established standards and transparent reporting so that practitioners and researchers can be confident in our findings.
2. **Relevant and useful:** responding to the evolving needs of practitioners and generating insights they can act on in different contexts around the world.
3. **Communicated well**: through simple and accessible means to reach as many people as possible and help them put the evidence into practice.

We work closely alongside researchers and bring together practitioners and policy makers to make sure any evidence we publish in our Global Safety Evidence Library meets these criteria. The guidance in this guide and in the standards and manuals set out throughout will help ensure our evidence reviews are robust, relevant and communicated well.

1.3 Working with the Centre

Our team will work closely with the research team to help ensure the success of the project and the quality of the research. We connect researchers to practitioners and our wider network, as well as other researchers working in our evidence community.

Our evidence reviews involve experts from research, policy and practice throughout to ensure that the findings are credible and relevant. We work with research teams and audiences to develop recommendations and to translate the evidence into summaries, briefings, infographics and practical tools.

1.4 The project consultation group

We expect all evidence review projects to include a Project Consultation Group to help steer the project and ensure the findings are useful and practical for end users. The Centre will work with the research team to bring together this consultation group for each project. This group should be made up of methods, topic and practice stakeholders, as well as a representative of the Centre's Expert Advisory Panel. The consultation group will meet with the review team at least three times during the course of a project in order to:

- Advise on the review questions and inclusion criteria.
- Advise on the review methods, ensuring that it is appropriate to the topic.
- Identify relevant studies and grey literature.
- Sense-check findings and interpretations.
- Disseminate and communicate findings with their networks.

1.5 About this guide

This guide sets out the approach for conducting evidence reviews for the Centre, including the standards and tools that ensure they are robust and credible, and the processes and people that make them relevant and practical. The guide is structured around the stages of an evidence review, from developing the research question(s) to communicating the results. In each section we set out what we are looking for in our evidence reviews and give examples and suggestions to help reviewers in their work. Where further reading may be useful, we have signposted to the relevant sources.

1.6 How to use this guide

This guide has been produced to support evidence reviewers, academics, researchers, and policy makers to collate, synthesise and communicate evidence on safe work and safety science (including safety systems and processes). The guide includes standard approaches that are applicable to every evidence review (i.e., developing an evidence review protocol, conducting literature searches and selecting studies, data extraction, quality assessment, data synthesis, and interpreting the results). The document also includes guidance on how to apply other approaches such as equity considerations in evidence reviews and conducting case study synthesis.

The document is intended as an introduction to our review approach, not a comprehensive guide. It is not intended to be prescriptive, and we encourage evidence reviewers to adapt methods to the specific context of their projects, and to apply innovative approaches where confident to do so in line with the developments in the field. However, as a minimum, we expect all our evidence reviews to be guided by the methods and approaches set out in this guide.

The guide will be reviewed and updated in October each year, to ensure that our evidence synthesis methods and approaches align with current practices.

2. Reviewing the global safety evidence base

> *'What works' is a method that can be used to improve the impact that research findings have on people's lives.*
> *It is based on the principle that good decision making is underpinned by good evidence, and if that evidence isn't available, robust ways of generating that evidence should be established. 'What works' recognises that research evidence on its own is not enough; you need to know how and why something works, for who, and finally, how to implement what is known.*
>
> (Lloyd's Register Foundation, 2024)

Safety interventions have the potential to reduce harms, accidents and injuries to people around the world, but policymakers and practitioners can't be certain that they are safe and effective without good evidence. Evidence reviews are needed that show interventions are safe, effective and cost effective.

The evidence base on safe work is patchy, with some interventions, sectors and occupations receiving more attention than others. Evidence and data are not equally distributed across the world or for different groups of people.

Evidence on safety is multi- and interdisciplinary, including occupational safety, safety systems, operational safety, engineering safety and environmental safety. Some disciplines are more likely than others to produce evaluations and intervention studies (occupational safety) while others have a greater focus on testing of materials and systems (engineering).

Much of the evidence comes from private organisations, governments and regulatory bodies which publish outside academic journals, for example on organisational websites. Practitioners have essential evidence on how safety cultures and practices work in different settings and workplaces.

Many of these sectors and disciplines have different definitions of safety and different ways of measuring it. All of these challenges make synthesising the evidence challenging and worth doing.

Some challenges with reviewing the safety evidence base

- Lack of conceptual consistency for safety
- Trends in concepts, e.g. use of terms such as 'resilience', 'reliability', instead of 'safety'
- Use of acronyms in titles and abstracts makes searches difficult and may require many variations of search strings to find the relevant studies
- Different disciplines have different reporting standards (engineering, systems safety vs occupational safety)
- Different sectors have different histories and appetites for using evidence in practice - makes making recommendation harder in some reviews than others

2.1 Our approach to evidence synthesis

The Centre awards research grants, directly commissions evidence reviews, conducts internal evidence synthesis, and brings in existing externally conducted evidence reviews into our evidence bank and living reviews. We use established methods and support innovation and testing of new approaches that may be more suited to our topics.

Our evidence reviews involve experts from policy and practice throughout, from developing research questions to interpreting and communicating results. This ensures the evidence is relevant and accessible, and people feel confident making decisions.

We are interested in evidence reviews that:

- Establish intervention effectiveness, and cost effectiveness
- Explore a problem and map the existing evidence base
- Refine important concepts, and assess methods and measures
- Bring together practitioner evidence and expertise / tacit knowledge

2.3 Types and sources of evidence

Evidence on safety comes from a range of sources and takes different forms. Safety science is multidisciplinary, and some safety risks themselves arise from disciplinary divides and silos in research and practice. One aim of our evidence reviews is to search widely for the best evidence and apply robust standards when synthesising it so that practitioners are confident in using it.

Our evidence reviews include a combination of:

- **Quantitative evidence** to establish intervention effectiveness, including for different populations in a range of settings. This evidence may take the form of randomised controlled, quasi-experimental or observational studies (e.g., cohort studies), project evaluations, or secondary data analysis.
- **Qualitative evidence** to understand how and why an intervention works, how people define concepts, how they feel about changes and innovations, and what is needed to implement something successfully. This may take the form of published qualitative studies, implementation and process evaluations, practitioner case studies, after incident reviews and so on.

Our evidence reviews should include both:

- **Peer-reviewed journal articles** sourced from relevant academic databases.
- **Grey literature**, including reports produced by government bodies, regulators, charities, private companies, industry bodies and think tanks, and other evidence which is not published in academic sources.

The role of Grey Literature in reviews of safety topics

Grey Literature plays an essential role in reviews by countering the effects of publication bias and including the experience of practitioners.

When it comes to reviews of safety topics (especially engineering), Grey Literature is particularly important since formal trials, reporting and publishing are less common than in other disciplines.

Types of Grey literature which are key sources of safety-related information include:

- Guidance, standards and white papers published by regulators or government departments.
- Safety impact assessments and technical reports published by industry bodies.
- Training manuals, incidence reports or pilots published by private companies.

You can read more about how to find grey literature later in this guide.

2.5 Equity considerations in evidence reviews

As a global evidence centre, our reviews bring together the best evidence from across the world, mindful of local or regional differences and the disparities in publication and use of evidence. Our reviews consider and analyse the prevalence of safety risks between and within occupations, sectors, regions, populations and contexts. Reviews of interventions look not just for 'what works', but 'how well, for whom, and in what contexts'.

The Centre aims to draw out the distributional impacts of new technologies, interventions, and approaches so that our evidence can help reduce inequalities in outcomes. To achieve this, we aim to search for, synthesise and report on variations in intervention effectiveness across populations and subgroups. We also aim to consider equity in review design and

implementation, and involvement of diverse panel of experts and stakeholders. For the purposes of this guide, *equity* is defined as the absence of unfair and avoidable differences in safety risk and outcomes among populations, regions, and contexts.

Equity consideration in reviews of global safety is a developing field, therefore, the Centre will explore and support reviewers to move the methods forward. Some existing methods and standards, including the PRISMA-Equity extension, the Cochrane PROGRESS-Plus, and the PRO EDI, have been used in other disciplines (such as public health) to integrate equity considerations in evidence reviews. These frameworks can be usefully adapted in reviews of global safety. The text box below provides more information on the PRISMA-Equity, the PROGRESS-Plus, and the PRO-EDI frameworks.

PRISMA-Equity and the Cochrane PROGRESS-Plus

The PRISMA-Equity checklist is an extension of the PRISMA checklist, aimed at providing guidance and support for reviewers to identify, extract, synthesise, and report evidence in systematic reviews with a focus on equity (Welch et al., 2012).

The purpose of the PRISMA-Equity checklist is to improve completeness and transparency of the conduct and reporting of systematic reviews on equity. It helps reviewers to identify, extract, synthesise, and report evidence on interventions or programmes that:

- target the general population, where it is important to explore the distribution of effects/impacts across different population characteristics, such as those defined by the PROGRESS-Plus or PRO-EDI frameworks;
- focus on at-risk, under-served, or disadvantaged populations; or
- aim to reduce social gradient across population subgroups.

The PRISMA-Equity checklist contains 27 items and recommends the use of the Cochrane PROGRESS-Plus framework to help reviewers to list and define data items related to equity. Further information about the PRISMA-Equity items can be found here.

PROGRESS-PLUS is an acronym for:
PROGRESS**:**

- **P**lace of residence; **R**ace/ethnicity/culture/language; **O**ccupation; **G**ender/sex; **R**eligion; **E**ducation; **S**ocioeconomic status; and **S**ocial capital

PLUS:

- Other personal characteristics associated with inequalities (e.g. age, disability)
- Other instances where a person may be temporarily at a disadvantage (e.g., work-related ill health)

PRO EDI (Equity, Diversity, and Inclusion)

PRO EDI builds on the Cochrane PROGRESS-Plus framework and offers a way for reviewers to collect, report and interpret data on core characteristics that can be associated with inequalities. These include age; sex; gender; sexual identity; race, ethnicity, and ancestry; socio-economic status (SES), level of education, disability, location (country(ies) of data collection, setting/context); and other factors that are relevant to the review.

PRO EDI provides a template to guide data extraction about population characteristics in evidence reviews, which can help reviewers to judge whether the review findings apply equally to all those who could benefit from the intervention or technology being reviewed.

PRO EDI was originally designed for reviews of randomised controlled trials; however, the tool can be useful for reviews of other study designs.

The following table provides an example of how to integrate equity considerations at each step of the review process using the PRISMA-Equity, PROGRESS-Plus, and the PRO-EDI frameworks.

Review step	**Equity consideration**
Review question and inclusion criteria	PICO-C (consider context) • **Population**: consider if the problem is similar across all populations. Define disadvantaged populations clearly • **Intervention**: consider potential for intervention generated inequalities (e.g., ease of access, bias in delivery of safety interventions, etc.) • **Comparator**: consider differences in resources across populations and regions. • **Outcomes**: consider differences across PROGRESS-Plus and PRO-EDI characteristics

Review step	Equity consideration
	• **Context**: context or setting may vary across PROGRESS-Plus/PRO-EDI characteristics which may cause inequity **Study design**: describe the rationale for including particular study designs
Search strategy and filters	• Consider what databases, terms, concepts and search filters are relevant to the review question(s) • Consider including terms relevant to the vulnerable or underserved populations in question
Information sources	Consider information sources (e.g. engineering, occupational health and safety databases and grey literature sources) that would help to address the review question(s)
Data extraction	• Consider how outcomes relevant to underserved populations are extracted and presented (e.g., presenting both absolute and relative differences) • Extract the results by age, ethnicity/race, disability, socio-economic status, etc. • Clearly describe sociodemographic characteristics of included studies. Consider using the PRO-EDI data extraction template as guide.
Critical appraisal	Look for differences when appraising evidence (e.g., attrition rates among population groups, delivery, receipt of, and adherence to intervention)
Data synthesis	The approach should be defined clearly in protocol • Present baseline risks and different relative effects - use additional rows or add a separate 'summary of findings' table • Conduct sub-group analysis to evaluate whether there are any differences in the intervention's effect across distinct sub-set of participants within the included studies, defined by PROGRESS-Plus and PRO-EDI characteristics (e.g., age, gender, race/ethnicity, etc.) • Analyse and present data on gaps, gradients, and targeted interventions • Discuss whether inclusion criteria affect generalisability • Discuss whether the search strategy included terms targeted at the vulnerable or underserved populations in question Discuss the applicability, transferability, and external validity of findings for underserved or vulnerable groups of interest

Review step	Equity consideration
Reporting	Include a section on each of the 27-items of the PRISMA-Equity checklist in review report.

Further reading

1. Welch VA, Petkovic J, Jull J, Hartling L, Klassen T, Kristjansson E, Pardo Pardo J, Petticrew M, Stott DJ, Thomson D, Ueffing E, Williams K, Young C, Tugwell P. Chapter 16: Equity and specific populations [last updated October 2019]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.5*. Cochrane, 2024. Available from www.training.cochrane.org/handbook.
2. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *PLoS Med* 2012;9(10):e1001333. doi: 10.1371/journal.pmed.1001333
3. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, Evans T, Pardo Pardo J, Waters E, White H, Tugwell P. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology*. 2014, 67 (1), pg. 56-64. doi: 10.1016/j.jclinepi.2013.08.005
4. PRO EDI interpretation guidance 2024. Available at: https://www.trialforge.org/trial-diversity/pro-edi/https://www.trialforge.org/trial-diversity/pro-edi/ (Accessed: 12th September 2025).
5. JBI Manual for Evidence synthesis. Equity in systematic reviews, 2025. Available at: https://jbi-global-wiki.refined.site/space/MANUAL/863633421/3.3.2+Equity+in+qualitative+systematic+reviews (Accessed: 12th September 2025).

3. Reviewing the evidence

The Centre supports a range of review methods to answer different research questions and priorities. Regardless of which method is used, all our evidence reviews have these things in common:

- Seek to answer relevant and timely research questions, developed in collaboration with global stakeholders from policy and practice.
- Minimise bias by using transparent, explicit and standardised methods.
- Interpret and share the findings in clear and accessible ways to enable confident and positive decisions to be taken.

3.1 Types of reviews and syntheses

As the practice of reviewing evidence for decision making has grown, so have the number and types of reviews that are conducted. As many as forty-eight different types have been identified that fit into a small subset of 'review families' (Sutton, 2019).

The main types of reviews and syntheses that are suitable for reviewing the evidence on global safety are:

Review type	**Purpose and features**
Systematic Review	A rigorous and comprehensive synthesis of the evidence base to answer one or more focused research questions. These reviews systematically find, collate, appraise and summarise the findings from a body of evidence. They can synthesise qualitative, quantitative, mixed and economic evidence, and Grey Literature. They can look at intervention effectiveness and cost effectiveness, prevalence, risk factors, and implementation. They bring together the findings in a narrative synthesis, usually into distinct evidence statements that answer the research questions. These use strict design, search, quality appraisal and analysis standards.

Review type	**Purpose and features**
Metanalysis	A statistical synthesis of quantitative findings from multiple studies. Metanalyses are often conducted as part of a systematic review.
Scoping Review	A broad review of the evidence base on a specific topic to map the available literature, give a descriptive overview of the evidence, and identify research gaps, priorities or trends. They can be useful to understand emerging or changing research topics, and to determine the need for a more in-depth systematic review. They can look at a variety of study designs, including experimental, observational and theoretical studies. This use widely recognised standardised methods, often adapted from systematic reviews.
Evidence Maps or Evidence Gap Maps	A systematic presentation of the evidence base on a topic or field, usually in visual rather than narrative form. They show the quantity and quality of the evidence base and identify gaps in a matrix format. They often display the evidence in rows and columns using a pre-defined framework based on primary dimensions (such as strength of evidence / effect size; or effectiveness / volume of literature). They may also include secondary dimensions or filters to enable the audience to interact and reorganise the data to focus on specific areas of the evidence (for example population or study design). They display the evidence using coloured bubbles, where the colours can signify the population or study type, and the size of the bubble indicates the number of studies. They can be part of a scoping review or a stand-alone research output.
Rapid Review or Rapid Evidence Assessment	A quicker and more streamlined evidence synthesis to answer a time-sensitive policy or practice question. These reviews can look at a range of study and evidence types,

Review type	Purpose and features
	as well as policy reports, and existing reviews. They provide quicker results by limiting searches or prioritising recent studies to speed up the review process. There are emerging quality standards for conducting these reviews. See the current Cochrane guidance for conducting rapid reviews of effectiveness here.
Review of Reviews or Umbrella Review	These synthesise existing systematic reviews (rather than primary studies) to generate higher level findings and recommendations from them. They are useful when the topic has already been reviewed multiple times from different angles. They bring together common findings and highlight where reviews report conflicting findings or gaps. However, they can miss areas of primary research that have not yet been reviewed. They use established standards for assessing quality and bias in systematic reviews and require rigorous design to avoid double counting primary study results. They can lead to, or inform living umbrella reviews when the field is rapidly developing and expanding. See a novel living umbrella review and knowledge translation approach here.
Conceptual Review	A review of the concepts, definitions, theories and frameworks that are used to explain and research a topic. They can help define a contested concept or 'tangled term' and refine theoretical frameworks. They look at how a concept has changed over time. They are useful when concepts are new and evolving rapidly, or when the same terminology is used differently across sectors, settings and disciplines. They are often carried out alongside a methods or measures review. Although there are no widely accepted standards for this type of review there are frameworks and analytical tools that ensure

Review type	Purpose and features
	it is conducted well.
Methods Review or Measures and Indicators Review	A review of the methodologies and measurement tools that are used to understand a topic. They assess the reliability and validity of instruments and explore how measures work in practice. They are often carried out alongside or following a conceptual review. These make use of standardised appraisal tools to compare and rank measures, although these tools are based on the health literature and may need adapting for safety measures.
Realist Review or Context-Mechanism-Outcomes Review	A theory-led approach that looks at the contexts and mechanisms that affect an intervention's effectiveness. They are useful when exploring how, why and for whom an intervention is effective; when exploring complex interventions or settings, and to inform the implementation of interventions and policies. This approach can be applied to any review which looks at interventions, such as a systematic or rapid review. There are some generally used quality standards for these reviews.
Living Review	A systematic review that continuously or regularly updates findings by incorporating new evidence as it is published or becomes available. They begin as a standard systematic review which is updated at specified intervals. These reviews are especially useful in rapidly evolving fields or contexts (such as healthcare and medicine), or when interventions quickly develop from an innovation stage. The design of the living review is particularly important because it affects all future updates and, unlike other reviews, cannot be corrected with a new review. The research question needs to be focused to make sure the findings remain coherent across time.

Review type	Purpose and features
	The standards for conducting living reviews are in development.
Case Study Synthesis	A method that synthesises evidence from a set of detailed examinations of interventions, projects or approaches in real world (rather than experimental) settings. They rely on rich and detailed qualitative evidence about the contexts, mechanisms, populations and outcomes, often from the perspectives of people involved in their design or delivery. They are not limited by evaluation design and are useful when more formal evidence is lacking. They generate findings that are very applicable to practitioners. This makes them very suitable for projects with few evaluation resources or those testing new interventions or approaches. There are developing standards for conducting these types of syntheses.
Other reviews	Many other types or review approaches exist, though they are not the preferred option for our reviews. Some examples include: • State-of-the-art review: a form of rapid review with heavier focus on interpretation of the current state of knowledge on a topic. • Bibliometric review: analyses bibliographic elements of the literature to identify important authors, papers, networks and connections.

There may be other types of evidence synthesis methods which are not included in this table. We encourage reviewers to use established methods where possible and use more innovative methods when they have experience and good reason to do so. Explaining why the review type and method has been chosen to answer the specific research question is important to ensure credibility and transparency.

More information about the standards and manuals for each review is available in the table in Appendix A.

3.2 Selecting the right review approach

Deciding which review approach is most appropriate will depend on a number of factors, including:

- The research questions,
- The state of the evidence base, including whether the concepts are clear and widely agreed on,
- The audience's needs and priorities, including how quickly they need to make decisions, and
- The resources available to carry out the review.

The following table shows some examples of how different aims and considerations are addressed by review types:

Aims and considerations	**Suggested review method**
Understand how effective an intervention is, (or how interventions compare against each other), or another focused research question.	Systematic review and/or meta-analysis
Broadly map the existing evidence base (including indications of what works and why), identify research gaps, and help prioritise further reviews.	Scoping review
Provide a time-sensitive and policy-relevant synthesis of the evidence.	Rapid Evidence Assessment / Rapid Review
Understand the use and definition of a concept or term, and how it is measured.	Conceptual Review and Methods and Measures Review
Bring together findings from several existing reviews to identify points of commonality or conflict and to prioritise research gaps.	Review of Reviews / Umbrella Review
Understand how an intervention or approach has worked in real life settings, or where evidence is generated by practitioners.	Case Study Synthesis

Generate regular updates on interventions or topics that are developing quickly or where evidence is growing steadily.	Living review

Different reviews will require different resources to complete, including the experience and diversity of skills in the team and the time and budget needed. The complexity of the research question, the state of the evidence, and the funding available all contribute to the scope of the review, and therefore the resources needed.

Systematic reviews take the longest due to their rigour and wide searches, while rapid reviews provide faster results by placing limits on the searches and scope. Other reviews lie somewhere in between the two and depend on the agreed scope and state of the evidence.

The following provides an indication of the length of time a review may take to complete.

Review type	**Approximate duration**
Case Study Synthesis	3 - 9 months
Rapid Review	3 - 6 months
Evidence map, evidence gap map	3 – 12 months
Scoping Review, Conceptual Review, Methods Review, Review of Reviews, Realist Review	6 - 18 months
Systematic Review	6 months to 2 years
Living Review	6 months to 2 years to set up, updates every 1 to 3 months after that

Further reading

6. Sutton, A., Clowes, M., Preston, L. and Booth, A. (2019), Meeting the review family: exploring review types and associated information retrieval requirements. Health Info Libr J, 36: 202-222. https://doi.org/10.1111/hir.12276
7. Campbell, F., Tricco, A.C., Munn, Z. *et al*. Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same but different— the "Big Picture" review family. *Syst Rev* 12, 45 (2023). https://doi.org/10.1186/s13643-023-02178-5

4. Stages of a review

Most of these review types have a common process for planning, designing, conducting, reporting and communicating the findings. The specific steps and standards may vary across the reviews, but the broad stages covered in this guide are as follows.

The stages of a review

1. Formulating research questions
2. Developing a review protocol
3. Creating a comprehensive search strategy
4. Selecting studies
5. Extracting data
6. Assessing the quality of included studies
7. Assessing the certainty of review-level findings
8. Synthesis of findings
9. Reporting of findings, implications and recommendations
10. Communicating findings and recommendations

Although the middle and latter stages are likely to be the most time and resource heavy, equal importance should be given to all stages, including engaging with stakeholders throughout.

Several comprehensive guides to each of these stages have been written which include much more detail than this guide (see further reading below and Appendix A). The following sections are an introduction to the stages and provide detail for reviews of safety-specific topics.

Further reading

8. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5 (updated August 2024). Cochrane, 2024. Available from www.training.cochrane.org/handbook.
9. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. *JBI Manual for Evidence Synthesis*. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-01
10. Snape D, Meads C, Bagnall A, Tregaskis O, Mansfield L. What Works Wellbeing: a guide to our evidence review methods. *What Works Centre for Wellbeing*; 2019 Apr. Available from: https://whatworkswellbeing.org/wp-content/uploads/2020/02/WWCW-Methods-Guide-FINAL-APRIL-2019a.pdf
11. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. York: University of York, 2009. https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.

5. Formulating a research question

Reviews should articulate clear research questions which, when answered, can provide meaningful evidence for decision making. The Centre supports research questions which are relevant and practical for practitioners, and which develop the fields of safety evidence.

Questions may be narrow and focused (as for systematic reviews), or broad (as for scoping reviews). The nature of the questions will guide the type of review needed to answer them. The research questions will help define the scope for the review and the criteria used to include studies.

5.1 Involving practitioners in developing research questions

The questions should be developed in collaboration with stakeholders from policy and practice. In the case of our reviews, these will generally be a consultation group of experts. This will ensure that the questions address practitioner priorities and generate actionable findings.

5.2 Frameworks for developing review questions and inclusion criteria

Review questions will generally set out the population, interventions and outcomes of interest. One of the following frameworks may be used to ensure the questions fully articulate these factors and start to describe the inclusion criteria.

Framework	Suitable for	Example question
PICOS **P**opulation **I**ntervention **C**omparator/s **O**utcome/s **S**tudy design	Reviews of quantitative evidence about intervention effectiveness. May also be used for reviews of qualitative or mixed-methods evidence.	In construction workers (P) does occupational safety and health training (I) compared to no training (C) reduce workplace accidents (O) in randomised controlled studies (S)?
PECOS **P**opulation **E**xposure **C**omparator/s **O**utcome/s **S**tudy design	Reviews which don't look at an intervention but at exposure to something (for example a risk, or the relationship between two factors).	In dock workers (P) does high exposure to dangerous chemicals (E) compared to low exposure (C) increase serious injury (O) in cohort studies (S)?
SPIDER **S**ample **P**henomenon of **I**nterest **D**esign **E**valuation **R**esearch type	Reviews of qualitative or mixed methods evidence.	According to OSH practitioners (S) what are the barriers to improving workplace safety (PI) through workplace policies (E) as explored in focus groups (D)? Qualitative evidence (R)
PEO **P**opulation **E**xposure **O**utcome	Reviews of qualitative evidence, risk, and aetiology (cause-and-effect)	In seafarers, is there an association between exposure to extreme weather conditions and subsequent development of respiratory diseases?
PICo **P**opulation or Phenomena of **I**nterest	Reviews of qualitative evidence	How do fishers (P) experience workplace safety policies (I) in small fleet vessels in Vietnam (Co)?

Context		

5.3 Other limitations for inclusion

As well as the criteria relating to these frameworks, it is likely you will want to use additional limitations to eligible studies. Any exclusions or limits should be decided with the consultation group and justified and recorded in the protocol.

- **Study design** (if not included in the frameworks above): For some reviews, such as scoping reviews, you may want to have no restrictions on study design. For reviews of intervention effectiveness and systematic reviews, the aim may be to include randomised controlled trials, but you may anticipate that few of these are available and so include cohort and other study design.
- **Publication date:** Such as limiting studies that have been published in the last 20 years. This will depend on the research question, advances in the evidence base and policy priorities. For example, you may want to exclude studies of workplace safety interventions published prior to a change in legislation. However, where there are no previously conducted evidence reviews on the topic, the Centre requires that the review includes eligible studies published from at least 1990 onwards.
- **Geographical location**: As a global evidence centre, our reviews bring together evidence from around the world. However, the published evidence base may favour developed countries in volume, leading to overrepresentation in your review. If your review requires limits on geographical location of interventions, you should set out clearly why this is the case. For example, some regulations only apply in certain countries, or the research question explores regional practices.
- **Language of publication**: In order to address potential language bias, searches need to be broad to start with. Ideally, no restrictions should be made at the search stage, (although they may be needed when selecting studies), and searches should include non-English language journals and databases where possible. You may decide to include all English and non-English language titles and abstracts in screening and then decide if the full-text language can be accommodated. You should only include studies in other languages if someone in the review team can confidently speak that language, and papers should not be translated using AI software.

The priorities for our reviews are that the searches are global and at least in English. The results may be split into high income and low-or-medium income countries.

Further reading

12. Richardson, W. S., Wilson, M. C., Nishikawa, J., & Hayward, R. S. (1995). "The well-built clinical question: a key to evidence-based decisions." *ACP journal club*, 123(3), A12–A13.

13. Morgan RL, Whaley P, Thayer KA, Schünemann HJ. Identifying the PECO: A framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. *Environ Int.* 2018 Dec;121(Pt 1):1027-1031. doi: 10.1016/j.envint.2018.07.015. Epub 2018 Aug 27. PMID: 30166065; PMCID: PMC6908441.

14. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qual Health Res.* 2012 Oct;22(10):1435-43. doi: 10.1177/1049732312452938. Epub 2012 Jul 24. PMID: 22829486.

15. Lefebvre C, Glanville J, Briscoe S, Featherstone R, Littlewood A, Metzendorf M-I, Noel-Storr A, Paynter R, Rader T, Thomas J, Wieland LS. Chapter 4: Searching for and selecting studies [last updated March 2025]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5.1 Cochrane, 2025. Available from www.training.cochrane.org/handbook.

16. Neimann Rasmussen, L., Montgomery, P. The prevalence of and factors associated with inclusion of non-English language studies in Campbell systematic reviews: a survey and meta-epidemiological study. Syst Rev 7, 129 (2018). DOI: 10.1186/s13643-018-0786-6

17. Letter to the editor: Pieper D, Puljak L, et al. Language restrictions in systematic reviews should not be imposed in the search strategy but in the eligibility criteria if necessary. *Journal of Clinical Epidemiology*, Volume 132, 146 - 147 DOI: 10.1016/j.jclinepi.2020.12.027

6. Developing a review protocol

Creating a clear, transparent and well-documented plan for carrying out the review is essential to provide confidence in its findings and recommendations. The review protocol sets out in advance the methods and stages of the research to minimise bias.

6.1 Registering the review protocol and recording changes

In order to maximise transparency and avoid duplication, review protocols should be made public by registering them in an open access location before literature searches begin. Examples of protocol registration platforms include:

- PROSPERO (International Prospective Register of Systematic Reviews) for systematic reviews. Although the focus of PROSPERO is health, it also includes safety, injuries and accidents, or outcomes indirectly affecting health.
- Open Science Framework platform
- University or institutional repositories.

For transparency and consistency, the Centre requires all evidence review protocols to be registered on the Open Science Framework platform.

The review protocol provides a transparent and robust plan, but sometimes changes will be needed to address unforeseen issues. For example, when initial searches do not return important studies, or return too many irrelevant studies, indicating a change in eligibility criteria is needed.

If any modifications to the protocol are needed after starting the work, these should be justified and recorded in the public registry and in the final report.

6.2 What to include in a review protocol

We recommend that review protocols cover the following in some form, with variations according to the type of review which is being planned.

What to include in a review protocol:

- **Background:** the research, policy and practice context; the rationale and objectives for the review; and any key concepts.
- **Research questions:** all the review questions and sub questions.
- **Inclusion criteria:** defined using PICOS/PECOS or other suitable frameworks, as well as any restrictions on publication date, language or geographical eligibility, and other exclusions.
- **Review methods:** including
    - Sources of information, including for grey literature
    - Search strategy, including indicative search strings
    - Study selection approach
    - Data extraction
    - Quality assessment
    - Synthesis and analysis approach, including assessing the confidence of findings
    - Reporting, communicating and dissemination plans
    - Any software used as part of the review process
- **Theoretical model or framework** (recommended).
- **Administrative information**: the authors, institutions, partners and funders involved in the research.

(Adapted from PRISMA-P checklist)

Further reading

18. **PRIMA extension for protocol reporting.** Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA, the PRISMA-P Group. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;349:g7647. doi: 10.1136/bmj.g7647
19. **Guidance on protocol development.** Covidence. Protocol Development for Systematic Reviews: A practical guide. Melbourne: Covidence; 2024 [link].

7. Search strategies and sources for safety topics

The next step is to translate the protocol into a comprehensive search strategy to be used across relevant databases and sources. This stage is intended to identify as many potentially relevant studies as possible and produce a list of retrieved records for screening in the next stage. Ideally, literature searches should be conducted by an expert librarian or information specialist, with quality assurance/peer-review of search strategies and translations available from a second specialist. This is especially important if you are new to reviews.

7.1 Developing and testing search strings

The search strategy needs to balance the requirement to identify relevant studies (sensitivity) with the ability to exclude irrelevant records (specificity). The main starting point for a search strategy is usually the PICOS framework, which includes the relevant terms to include in search strings. For example, a standard search string will include terms for the population, phenomenon of interest, and study design. A broad set of search terms, using keywords and index subject headings should be combined using the Boolean OR/AND operators to ensure the balance of sensitivity and specificity. Test searches will determine whether this balance returns meaningful and manageable volumes of results.

Different review methodologies will require different search strategies. Guidance on how to develop these is available in the technical manuals and guidelines listed in Appendix A.

Example of search strategy for occupational safety review

The following search strategy was used by van der Molen et al. in their 2018 review of interventions to prevent injuries in construction workers, available here.

Search strategy for MEDLINE in PubMed

We ran preliminary searches in PubMed to define useful terms for the search strategy. This revealed that searches could be made sensitive but not specific enough to decrease the total amount of references retrieved to a manageable number, which we set at about 10,000. We developed the definitions described below.

Search terms for types of participants: working at construction sites.

The search term construction is truncated as construction* according to the industry name not as construct*, since many other things can be constructed for example, vectors or plasmids in the biochemistry field. We did not use the terms "construction industry" or "construction worker" so as not to make the search too specific.

Many articles mentioned the word building instead of the term construction, so we added building* as a search term.

There may be articles including neither construction nor building. This is why Koningsveld 1997 used the most important job titles (trades) in their search strategy. In addition, we added the following job titles that appeared many times in the articles found in the preliminary searches: laborer/labourer and contractor.

The terms construction, building and job titles like carpenter are also used for other purposes such as a surname or in a company or street name (location), so the search terms concerning the population are followed by a search tag [tiab] (title abstract) or [tw] (text word).

Search terms for outcome: injury
The primary outcome in the search strategy was defined as an injury, and the term was truncated to injur* to make it sensitive enough.

We also considered the terms accident and safety. Accident was truncated as accident* to make it sensitive enough.

Search terms for interventions
Intervention in the search strategy was defined as any kind of intervention related to safety management, risk management or accident prevention applied to decrease the rate or severity of injuries. Terms resembling these kinds of interventions were selected for this part of the search strategy.

Search terms for study design

For study design, we used two search strategies to find (cluster) randomised controlled trials and prospective non-randomised controlled trials or interrupted time series; for the Discussion section the last strategy, search #7, was also used to find before-after studies and case-reference studies. For randomised controlled trials, we will use the strategy described by Robinson 2002, and for non-randomised studies the strategy described by Verbeek 2005.

We used search terms that covered the concepts of 'construction workers' (participants), 'injury' (primary outcome measure), 'safety' (interventions) and 'study design' to identify studies in the electronic databases.

We used the following search strategy adapted as appropriate to the specifications of each database:

#1 construction*[tiab] OR building*[tw] OR builder*[tiab]OR laborer* [tw] OR labourer* [tw] OR contractor* [tw] OR supervisor*[tw] OR "machine driver"[tw] OR "machine drivers"[tw] OR "machine operator"[tw] OR "brick mason"[tw] OR "pile driver"[tw] OR "pile drivers"[tw] OR "concrete worker"[tw] OR "concrete workers"[tw] OR "metal worker"[tw] OR "metal workers"[tw] OR "road builder"[tw] OR "road builders"[tw] OR "pipe driver"[tw] OR "pipe drivers"[tw] OR "tower crane"[tw] OR fitter*[tw] OR carpenter* [tw] OR rammer* [tw] OR scaffolder* [tw] OR bricklayer* [tw] OR pointer* [tw] OR plasterer* [tw] OR plasterpainter* [tw] OR roofer* [tw] OR plumber* [tw] OR glazier* [tw] OR screeder* [tw] OR electrician* [tw] OR tiler* [tw] OR painter* [tw] OR paviour* [tw] OR pavier*[tw] OR ironwork*[tw] OR metalwork*[tw] OR asphalt*[tw] OR roofing[tw] OR painting[tw] OR "construction materials"[MeSH] OR "facility design and construction"[MeSH]

#2 injur*[tw] OR accident*[tw] OR "accidents, occupational"[MeSH] OR "wounds and injuries"[MeSH] OR harm*[tw] OR wound*[tw] OR fall[tw] OR falling*[tw] OR burn*[tw] OR slipper*[tw] OR poison*[tw] OR fatal*[tw] OR "injuries"[MeSH Subheading]
#3 Safety[MeSH] OR "Safety Management"[MeSH] OR "prevention and control"[MeSH Subheading] OR safet*[tw] OR prevent*[tw] OR control*[tw] OR risk[tiab] OR "risk"[MeSH Term] OR "risk management"[MeSH Terms] OR "accident prevention"[MeSH Terms]

#4 = #1 AND #2 AND #3

#5 (randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized controlled trials[mh] OR random allocation[mh] OR double-blind method[mh] OR single-blind method[mh] OR clinical trial[pt] OR clinical trials[mh] OR "clinical trial"[tw] OR ((singl*[tw] OR doubl*[tw] OR trebl*[tw] OR tripl*[tw]) AND (mask*[tw] OR blind*[tw])) OR "latin square"[tw] OR placebos[mh] OR placebo*[tw] OR random*[tw] OR research design[mh:noexp] OR comparative study[mh] OR evaluation studies[mh] OR follow-up studies[mh] OR prospective studies[mh] OR cross-over studies[mh] OR control*[tw] OR prospectiv*[tw] OR volunteer*[tw]) NOT (animal[mh] NOT human[mh])

#6 = #4 AND #5
#7 (effect* [tw] OR control* [tw] OR evaluation* [tw] OR program* [tw]) NOT (animal[mh] NOT human[mh])
#8 = #4 AND #7
#9 = #6 OR #8

7.2 Searching for contested concepts

The right search strategies are particularly important when terms or concepts are less developed or where categories and concepts are contested, as is often the case in safety. There are trade-offs between sensitivity and specificity in all reviews, but where concepts or definitions don't have universal definitions, or are used differently in different disciplines, the balance between these two aspects will be especially important. Using iterative searches or designing a review with different search stages may be useful approaches.

Study-type limits or filters should be generally avoided, due to the broad nature of safety evidence. Existing databases may not provide adequate indexing by study design, and the quality of indexing for – and the vocabulary used in – study methodologies and designs varies extensively and, in some instances, is poor. However, if you are confident that the disciplines you are searching have good indexing practices, and you are focusing on experimental or intervention studies, then limits may be appropriate (and tested via dummy searches).

7.3 Selecting bibliographic databases for safety topics

Searching for evidence across the main safety topics is likely to involve a number of core and subject-specific databases, as well as other sources.

The multidisciplinary nature of the topics and the global application of interventions means there is no single dedicated source for our evidence. The study design or type of output sought will also determine the relevant databases. Identifying the most useful databases should be done with support from the consultation group and academic or specialist librarians/information specialists.

Databases and sources for safety topics:

- Scopus
- Citation indexes via Web of Science
- PubMed/Medline
- Cochrane Library
- Campbell Collaboration reviews
- Embase
- SafetyLit (nb. SafetyLit stopped being updated in August 2024)
- PsychLIT
- EconLit
- For engineering: the IET Digital Library, Inventory of Carbon and Energy (ICE), IEEE Xplore.

There are an increasing number of web-based resources that index available web content, including open access peer-reviewed literature. These may prove to be suitable alternatives to support review activities but are still to be fully evaluated for this purpose.

7.3 Grey Literature and call for evidence

Grey literature is research and information that has been produced outside of academic settings and publishing, and is therefore not formally peer reviewed. It's produced by governments, sector bodies, think tanks, or private and third sector organisations for a variety of reasons. It can include research or evaluation reports, business reports, manuals and standards, and white papers. It also includes not yet published academic evidence, such as conference papers or preprints.

Grey literature is important to include in reviews of global safety because it provides important evidence of real-world implementation and innovation, it contains topical information on policies, and it helps mitigate against publication bias of positive academic studies. Grey literature can both supplement evidence of intervention effectiveness and provide information about implementation in different contexts. Grey literature provides evidence from across the evidence pipeline - from case studies, to evaluations, and more formal studies.

Grey literature should be sought and included for all our reviews. Where grey literature is not included, the decision should be justified and reported. Grey literature should only be included in a review when it meets the inclusion criteria of the review, and when it is publicly available and the source can be cited (i.e. the authors and publisher or institution).

Much of the safety-relevant grey literature will be produced by private companies. Reviewers should be mindful of potential commercial interests as well as other biases and report them fully in their analysis.

Since grey literature is unlikely to be included in bibliographic databases, finding it requires a different approach:

- A **call for evidence** should be published on the Centre website and disseminated using social media and professional networks.
- The expert **consultation group** should signpost to relevant grey literature publications or sources, and help disseminate the call for evidence.

- Dedicated **grey literature databases** can be searched. These include OpenAlex, CORE, BASE (Bielefeld Academic Search Engine), OAIster, Dimensions, Overton and Google Scholar.
- **Internet search engines** can be used with specific search strings to find studies in large government or third sector websites, or specific countries, for example using Advanced Google Search.
- Targeted searches of **relevant organisation websites.**
- **Policy, manuals and guidelines repositories** may be searched, for example WHO Iris.
- **Lloyd's Register Foundation evidence bank and knowledge repository** should be searched for relevant publications.

7.4 Documenting the search strategy

The search strategy should be documented and reported for transparency, and so that searches can be re-run at a later date, for example when updating a previous systematic review.

Documenting the search for evidence

Information sources: Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.

Search strategy: Present the full search strategies for at least one databases including any filters and limits used, so it can be repeated.

Additional searches: Describe the broad terms used to address equity questions. Describe grey literature information sources that were searched.

(Adapted from PRISMA and What Works Centre for Wellbeing)

7.4 Additional searches

It may be necessary to conduct additional searches to identify evidence using supplementary methods such as citation searching or 'related records' functions in databases. The decision to use these methods would be made by the review group depending on the topic being investigated. Additional searches should be recorded in the protocol.

7.5 Using software in searches

It is recommended that reference management software be used to manage the retrieved references. These simplify processes such as deduplication and formatting and provide efficiencies for the later stages of the review, such as study selection and report writing. Several proprietary tools are available (such as Endnote or EPPI-Reviewer) as are free web tools (Zotero or Mendeley). The functionality of these varies greatly and should be investigated for suitability. Tools such as EPPI-Reviewer provide functions beyond the collation of the literature and support the entire review process.

Further reading

20. **Cochrane Handbook for Systematic Reviews of Interventions.** Chapter 4: Searching for and selecting studies. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. London: Cochrane; 2024 [cited 19 June 2025]. Available from: https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-04
21. **ISSG Search Filter Resource.** Glanville J, Lefebvre C, Manson P, Robinson S, Brbre l and Woods L, editors. York (UK): The InterTASC Information Specialists' Sub-Group; 2006 [updated 4 June 2025; cited 4 June 2025]. Available from https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home
22. **On reporting searches**. PRISMA-S checklist and explanatory paper: Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, Koffel JB; PRISMA-S Group. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021;10(1):39. doi: 10.1186/s13643-020-01542-z
23. **On searching in reviews of reviews.** Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Development of an efficient search filter to retrieve systematic reviews from PubMed. J Med Libr Assoc. 2021 Oct 1;109(4):561-574.

8. Study selection

The search stage will return many more records than are eligible for inclusion in the review. The first step in reducing this number is to remove duplicate studies using the recording software, followed by screening against the inclusion criteria.

8.1 Screening and selecting studies

The process of identifying and selecting the eligible studies should start by working through some examples from the searches against the inclusion criteria. Members of the team should work through the examples separately and then compare results to ensure the criteria is uniformly applied and to highlight any issues.

The studies should then be selected using the following screening steps:

1. **Title and abstract screening:** the titles and abstracts of studies should be double screened (independently assessed by two reviewers). Any disagreements should be resolved in discussion with other reviewers. If the disagreement cannot be resolved the study should be retained.
2. **Full-paper screening:** following the title and abstract screening the reviewers should assess the full text of each study. Dedicated full-paper screening tools such as EPPI-Reviewer or Covidence may be used. Studies should normally be assessed independently by two reviewers, and any differences resolved between them or by consulting the wider team.

If the volume of returns is very large, a random selection (e.g. 20%) may be double screened, with the remainder being single screened, though you should conduct statistical tests to ensure the consistency of single coding is high enough. In Rapid Reviews a single reviewer may undertake the screening, with a second reviewer screening a much smaller sample. This speeds up the review process but increases the risk of bias or missing relevant studies.

8.2 Duplicate publications of the same study

Some research projects publish their results differently in multiple journals, so that your searches may return several different records for the same findings. If these publications are treated as separate studies in the review, there is a risk of double counting or giving the findings undue weight overall. It can be difficult to determine if multiple publications refer to the same study since they may include different authors or report different outcomes.

Multiple reports of the same study should be treated as a single study (with reference made to all publications), and this should be recorded in the study table. Identifying duplicate studies isn't always straightforward. Some techniques include looking for studies with matching sample sizes, designs and measures, authors or grant numbers.

8.2 Using Artificial Intelligence tools

The use of Artificial Intelligence (AI) in conducting reviews is still at an early stage of development, with no commonly agreed standards and methods.

Some examples of AI tools used in reviews include:

- Elicit, Connected Papers and Research Rabbit which can help to track and identify papers and visualise the relationships between them.
- Scite which can help evaluate the credibility of research by showing how and why a paper is cited.
- Semantic Scholar which can filter large volumes of search results from databases like PubMed and Google Scholar.

AI tools may be used to help develop search strategies, identify and screen studies, extract and analyse data and report findings. AI tools have the potential to speed up certain stages of reviews and make large volumes of data more manageable, as well as spot patterns and themes that may be missed by reviewers. AI tools might play an important part in reviewing Grey Literature and evidence that doesn't appear in research databases. It's likely that AI tools will be increasingly used in evidence reviews in the coming years.

As an experimental approach it carries risks, primarily to the transparency, accuracy and completeness of reviews. Reviewers need to carefully assess and report the bias and weaknesses they pose.

AI tools should support and augment the role of human reviewers, not replace them, and reviewers should conduct thorough external validation of AI-generated findings. Reviewers should also address the potential lack of trust from practitioners and policy makers when AI is used in research (Clark, 2025).

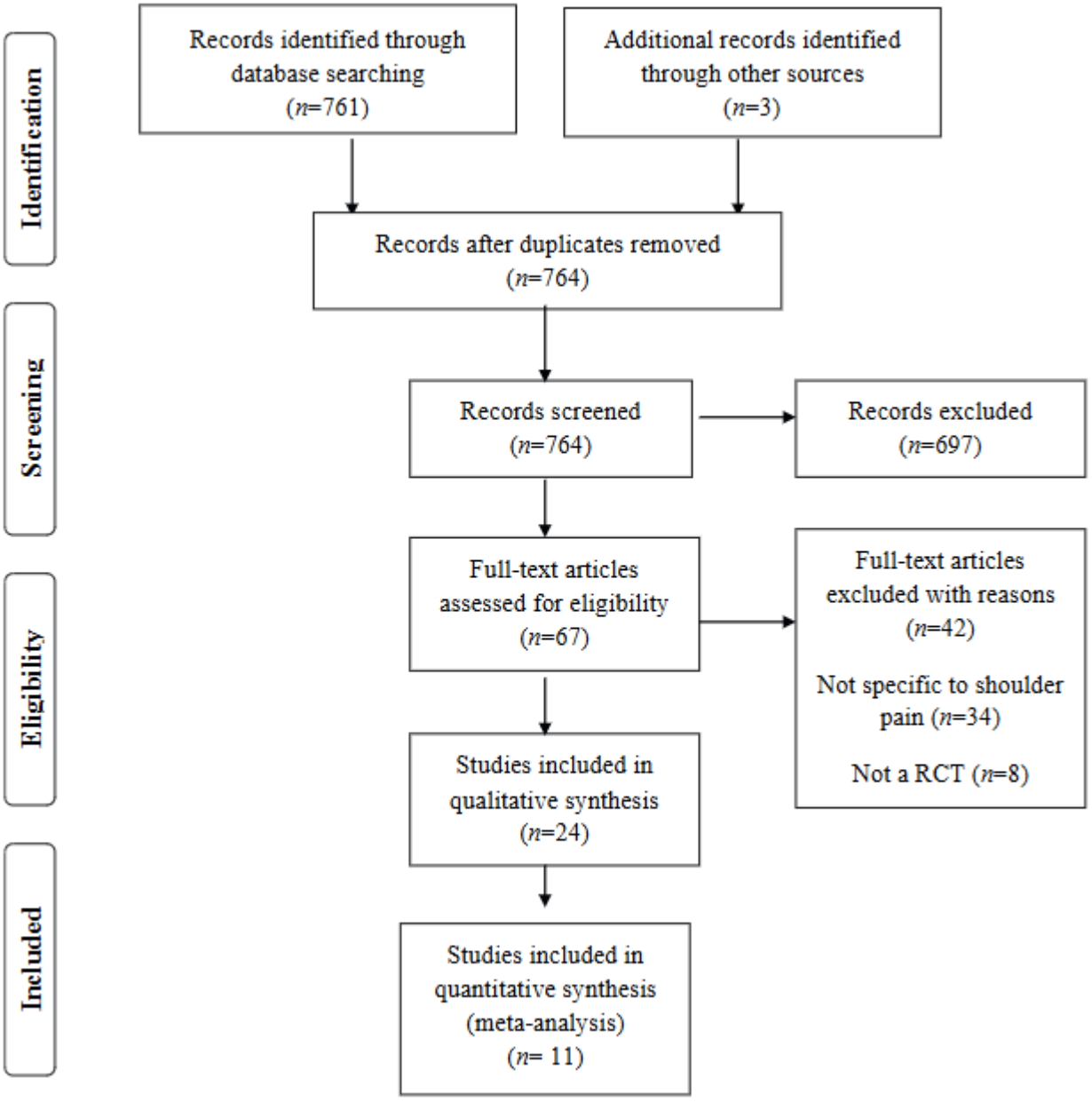
The Responsible AI in Evidence Synthesis (RAISE) project is developing guidance and recommendations on the use of AI in reviews, and NICE has issued a position statement on the use of AI in evidence generation (NICE position statement). King's College London's LibGuide on Systematic Reviews has a useful summary on the role of AI. Recently, major evidence synthesis organisations – including Cochrane Collaboration, Campbell Collaboration, Joanna-Briggs Institute and the Collaboration for Environmental Evidence – have issued a joint Position statement on artificial intelligence (AI) use in evidence synthesis.

In line with the above guidance, reviewers conducting evidence reviews for the Centre can use appropriate AI tools if they can show that it will not compromise the integrity and methodological rigour of the review. Reviewers should adhere to the Responsible AI in Evidence Synthesis (RAISE) recommendations at all times when applying AI in our evidence reviews.

8.3 Documenting the study selection

The searches and study selection should be documented, including details relating to the inclusion criteria. The search and selection process should also be summarised using a PRISMA flow diagram to ensure accurate and transparent reporting. See an example of a PRISMA flow diagram on the next page.

Example of PRISMA flowchart showing study inclusion



From Lowry, V., Desjardins-Charbonneau, A., Roy, J.-S., Dionne, C. E., Frémont, P., MacDermid, J. C., & Desmeules, F. (2017). Efficacy of workplace interventions for shoulder pain: A systematic review and meta-analysis. *Journal of Rehabilitation Medicine*, *49*(7), 529–542. https://doi.org/10.2340/16501977-2236

9. Data extraction

The data from every included full paper should be extracted into a pre-agreed template or table. This should be done by a single reviewer, but a sample checked by one or two additional reviewers to check for accuracy. The evidence table should list and define all the variables for which data was extracted - including PICOS, numerical results, funding sources - as well as any assumptions made

9.1 What data to record for each study

Where available, exact p-values and confidence intervals should be reported, as well as the tests used to obtain them. Where p-values are not adequately reported or not given in studies, this should also be recorded. Any descriptive statistics or analysis should be recorded. We would generally consider a p-value of ≤0.05 statistically significant, and reviewers should set out their thinking if other thresholds are used.

What to include in a table of included studies

- Bibliographic information (e.g. authors, publications date)
- Study design (e.g. RCT, case–control)
- Population (characteristics, eligible and selected)
- Geographical location and setting
- Dates of study (start and end dates, length of follow up)
- Intervention/Exposure, if applicable (components, intervener, duration or length of exposure, method, mode or timing of delivery, fidelity of implementation)
- Method of allocation to study group (if applicable)
- Numbers of participants in each group at baseline and at follow up (if applicable)
- Outcomes (primary and secondary and whether measures were objective, subjective or otherwise validated, adverse or unexpected outcomes)
- Measurement tools (name and description, upper and lower limit and thresholds of scales, timing of outcome measures)
- Key numerical results (including proportions experiencing relevant outcomes in each group, means and medians, standard deviations, ranges and effects sizes, statistical methods used)
- Inadequately reported or missing data.

The Centre is developing a map of global safety evidence, so reviews should draw out information that can help inform this. For example, the location and population in the studies. Reviewers should also report when there is a growing evidence base in a language other than English, so this can be recorded in the evidence maps. Where studies have used or developed theoretical models, these should also be reported.

Further reading

24. Li T, Higgins JPT, Deeks JJ. Chapter 5: Collecting data [last updated October 2019]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.5. Cochrane, 2024. Available from cochrane.org/handbook.

10. Quality assessment

Quality assessments in reviews occur at two distinct phases. First each included study is assessed, and second the overall certainty of review-level findings is assessed. The two tasks require different tools.

For some types of reviews, such as scoping reviews or evidence maps, comprehensive quality assessment may not be suitable or necessary. However, reviews that don't carry out some form of quality assessment or description of the strength of the evidence base will be of limited use for our audiences. We encourage the use of quality assessment in scoping reviews where possible.

10.1 The quality of included studies

The studies in the global safety evidence base vary greatly in rigour. Assessing the risk of bias and quality of design and analysis of each included study is essential in determining the overall confidence of the review findings and recommendations.

The focus of quality assessment is determining a study's internal validity (the extent to which it accurately shows that the observed effects are caused by the intervention or exposure, not by other factors). The different strengths and weaknesses of each study should be assessed, and an overall comment on the risk of bias for the study should be recorded in the study summary table.

A number of checklists and tools are available to ensure this process is done systematically according to the type of study included. Where studies use different types of evidence, each aspect of the study should be assessed separately. The research protocol should specify which tools you will use for this stage.

Assessing the quality of included studies

Type of evidence	Type of study	Checklist or tool
Quantitative	Randomised-control trials	Risk of Bias 2 (Rob2): A revised Cochrane risk-of-bias tool for randomized trials
	Non-randomised studies	Risk Of Bias In Non-Randomized Studies of Interventions (ROBINS-I V2)
	Cross-sectional	JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies
Qualitative	All qual studies	CASP (Critical Appraisal Skills Programme) for Qualitative Research
		JBI Critical Appraisal Checklist for Qualitative Research
	Case studies	See Appendix D
Mixed methods	Mixed-method studies	Mixed Methods Appraisal Tool (MMAT)
Other	Systematic reviews	AMSTAR-2 (or ROBIS)

The above are our recommended tools for use in our reviews. A number of older approaches (such as Newcastle-Ottawa Scale (NOS) for non-randomised studies) are still used, but they have been superseded by more rigorous tools, so we do not recommend them here.

Some reviews may use other standards of evidence when the evidence base is unsuitable for these tools. This may be the case for engineering or technical reports in the global safety evidence base. Some other tools include the NESTA standards of evidence and the What Works Centre for Wellbeing Quality in Quality framework. Ideally, you should use one of the recommended checklists, but if necessary other approaches may be used.

10.1.1 Assessing external validity of included studies

The checklists above focus primarily on assessing a study's risk of bias to determine its internal validity. This is the priority of this stage of quality assessment, since it is the most important factor in determining accurate and reliable results (Cochrane Handbook, ch.7).

However, our evidence reviews are intended to produce applicable recommendations for decision makers in different sectors and settings around the world. Understanding how generalisable the included studies are is important to achieve that aim.

External validity (also known as generalisability) is how well the evidence in the research you are assessing can be relevant to the situation in practice. Some research may not be locally relevant because, for example, the setting is completely different, or the intervention might not be locally acceptable. This is very much a matter of judgement and if in doubt, you may need to come to a consensus within the team.

10.2 Assessing the certainty of review findings

In order to help decision makers act with confidence on the findings of a review, each finding should be given a certainty rating based on the quality of the evidence it is based on.

The two recommended approaches to be used for this assessment are:

- GRADE (Grading of Recommendations, Assessment, Development, and Evaluations) for quantitative evidence synthesis findings.
- GRADE-CERQual (Confidence in the Evidence from Reviews of Qualitative Research) for qualitative evidence synthesis findings.

Although both approaches were developed for use in systematic reviews of health interventions, they are widely used in other disciplines as well as other review types including scoping, rapid and umbrella reviews.

10.2.1 Applying GRADE to quantitative synthesis findings

GRADE is a system for grading certainty of evidence which has been adopted by over 100 organisations worldwide.

> *[…] The grade approach defines the certainty of a body of evidence as the extent to which one can be confident that an estimate of effect or association is close to the quantity of specific interest. [it] involves consideration of within- and across-study risk of bias (limitations in study design and execution or methodological quality), inconsistency (or heterogeneity), indirectness of evidence, imprecision of the effect estimates and risk of publication bias […] The grade system entails an assessment of the certainty of a body of evidence for each individual outcome.*
> *(Cochrane Handbook).*

The GRADE Working Group has a number of resources to help you apply the approach in your review, including the GRADE Handbook. The GRADE working group has also published an update (Montgomery, 2019) on how to apply the approach in reviews of complex interventions or settings (such as public health, or in our case, global maritime systems).

The GRADE approach has four quality ratings (from GRADE Handbook, 5)

Grade	**Definition**
High	We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

In keeping with other widely accepted evidence ratings, study design is the main factor that determines rating. So that the 'high' rating is generally used for randomised-control studies or trials, while robust observational studies would be rated as 'moderate' or 'low'.

However, an RCT or other 'high' rated study may be downgraded according to five considerations:

1. Study limitations (risk of bias)
2. Inconsistency of results
3. Indirectness of evidence
4. Imprecision
5. Publication bias

Similarly, evidence initially given a 'low quality' rating (such as evidence from observational studies) can be graded upwards if there is:

1. A very large magnitude of effect
2. A dose-response gradient; and
3. All plausible biases would reduce an apparent treatment effect.

Alternatively, non-randomised studies can be upgraded to 'high' initial certainty if the ROBINS-I tool, primarily designed for cohort studies, has been used to assess the risk of bias in the studies.

Justifications for upgrading or downgrading the certainty of the evidence can be found in chapter 14 of the GRADE Handbook.

GRADEpro GDT is AI powered software designed to implement the GRADE approach. It allows reviewers to create the summary of findings table including calculating relative effects and absolute risks from data from controlled trials. This software, as all artificial intelligence-based tools, should be used with caution.

For adaptations of the GRADE approach to different types of reviews and data see Additional file 5B of Kolaski, 2023.

11.2.1 Applying GRADE-CERQual to qualitative synthesis findings

GRADE-CERQual provides an approach for assessing how much confidence to place in the findings of a qualitative evidence synthesis. It was developed by the GRADE Working Group and is widely used in reviews worldwide.

CERQual has four quality ratings based on considerations of methodological limitations, coherence, adequacy and relevance (Lewin, 2018).

Level	Definition
High	It is highly likely that the review finding is a reasonable representation of the phenomenon of interest.
Moderate	It is likely that the review finding is a reasonable representation of the phenomenon of interest.
Low	It is possible that the review finding is a reasonable representation of the phenomenon of interest.
Very Low	It is not clear whether the review finding is a reasonable representation of the phenomenon of interest.

For information on applying the approach in complex interventions or settings, see Lewin, 2018.

Further reading

25. Kolaski, K., Logan, L.R. & Ioannidis, J.P.A. Guidance to best tools and practices for systematic reviews. *Syst Rev* 12, 96 (2023). https://doi.org/10.1186/s13643-023-02255-9
26. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, Bohren MA, Tunçalp Ö, Colvin CJ, Garside R, Carlsen B, Langlois EV, Noyes J. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci.* 2018 Feb 1;13(Suppl 1):2. doi: 10.1186/s13012-017-0688-3. PMID: 29384079; PMCID: PMC5791031.
27. For more information on Nesta Standards of Evidence see Puttick, R., & Ludlow, J. (2013). *Standards of evidence: An approach that balances the need for evidence with innovation*. Nesta. https://media.nesta.org.uk/documents/standards_of_evidence.pdf
28. For more information on assessing qualitatative evaluations see Breckon, J., & Puttick, R. (2021). *Quality in Qual: A proposed framework to commission, judge and generate good qualitative evaluation in wellbeing impacts*. What Works Centre for Wellbeing. https://whatworkswellbeing.org/resources/quality-in-qualitative-research/
29. Schünemann H, Brożek J, Guyatt G, Oxman A, editors. *GRADE Handbook: Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach*. GRADE Working Group; 2013. Available from: https://gdt.gradepro.org/app/handbook/handbook.html.
30. Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, Bohren MA, Tunçalp Ö, Colvin CJ, Garside R, Carlsen B, Langlois EV, Noyes J. *Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series*. Implement Sci. 2018 Jan 25;13(Suppl 1):2 FULL TEXT
31. Montgomery P, Movsisyan A, Grant SP, Macdonald G, Rehfuess EA. Considerations of complexity in rating certainty of evidence in systematic reviews: a primer on using the GRADE approach in global health. *BMJ Glob Health*. 2019 Jan 25;4(Suppl 1):e000848. doi: 10.1136/bmjgh-2018-000848. PMID: 30775013; PMCID: PMC6350753.
32. Hempel, Susanne. *Systematic Reviews for Occupational Safety and Health Questions: Resources for Evidence Synthesis*. RAND 2016 10.13140/RG.2.1.2597.0167.

11. Synthesis

Synthesising the studies involves combining and summarising the evidence in narrative and numerical form.

- **Narrative synthesis** draws together the results in a descriptive and analytical narrative, with groupings and subgroupings relating to the research question.
- **Quantitative synthesis** involves the use of statistical techniques such as meta-analysis, which is primarily relevant for systematic reviews and reviews of evidence effectiveness.

11.1 Narrative synthesis

The starting point of synthesis should be a clear descriptive summary of all the studies in the form of a narrative. This narrative groups together studies and analyses the commonalities and differences between them in order to arrive at an overall set of evidence statements or findings.

This narrative summary can be organised in different ways but should include details about the intervention or phenomena of interest, the population, and the outcomes including effect size and direction. In other words, the narrative should reflect the interests and priorities of decision makers and other audiences of the review.

The groupings and clusters in the narrative should be organised to help answer the research question. The narrative may be organised by intervention type, with subgroups for the outcomes for different populations or settings, or by outcomes of a single intervention, with subgroups by population or setting.

Where studies concur or differ in the size or direction of effects, the narrative should explore why this may be the case. For example, by referring to the similarity or otherwise of the population group, settings, comparators or outcome measures. For populations, it's important to describe the differences in outcomes between different population groups within each study and between studies, to draw out potential inequalities.

The most useful narrative syntheses for safety topics are those that describe and test theoretical frameworks of how or why different interventions are intended to work. This

kind of synthesis creates practical recommendations for practitioners and clear gaps for future researchers to follow.

For practical guidance on developing a narrative synthesis, including how to establish a framework for synthesis, see Popay, 2006.

11.2 Meta-analysis and quantitative synthesis

There is technical guidance on combining quantitative results from two or more studies in chapter 10 of the Cochrane Handbook and in the JBI Manual.

11.2 Summarising the findings in a table

The findings with their quality assessment should be summarised in a table organised according to the intervention or outcome of interest. The strength of the evidence against each outcome or intervention should also be recorded.

What to include in a synthesis or summary of findings table

- Outcome or intervention or exposure
- Number of studies, type or studies
- Number of subjects or study participants
- Findings, including effect size and direction
- Quality assessment of studies, bias and limitations
- Quality assessment of finding (GRADE, CERQual)

Based on GRADE and Hempel (2016)

The table may include symbols (such as arrows to indicate the direction of an effect) or short phrases and descriptions. It should enable the reader to understand the evidence behind each intervention or outcome finding at a glance. Evidence quality or evidence certainty may be reported with symbols or statements.

Example of a summary of findings table

Table 4: Overview of behavioural interventions for the prevention of occupational injuries

Intervention	Comparison	Number and study design	Population/ Setting	Method of data synthesis	Outcome	Result	Quality of evidence	Author and year
Active training (face-to-face instruction)	Passive training (folders or videos)	1 retrospective cohort study	Health or hospital staff	Individual study	Noncompliance with PPE use	◀▶	Very low	Verbeek et al. 2016 [40]
Active training (face-to-face instruction)	Passive training (folders or videos)	1 retrospective cohort study	Health or hospital staff	Individual study	Noncompliance with doffing guidance	▲	Very low	Verbeek et al. 2016 [40]
Drug-free workplace programme	No intervention	1 ITS	Construction workers	Individual study	A: non-fatal injuries in the year following implementation B: non-fatal injuries in the years thereafter	A: ▲ B: ▲	Very low	van der Molen et al. 2018 [35]
Safety campaign	No intervention	1 ITS	Construction workers	Individual study	A: initial decrease in injuries at the company level B: sustained decrease in injuries at the company level C: initial decrease in injuries at the regional level D: sustained decrease in injuries at the regional level	A: ◀▶ B: ▲ C: ◀▶ D: ▼	Very low	van der Molen et al. 2018 [35]
Safety training interventions	No intervention	1 ITS, 1 CBA	Construction workers	Individual studies	A: non-fatal injuries, immediate effect B: non-fatal injuries, trend	A: ◀▶ B: ◀▶	Very low	van der Molen et al. 2018 [35]
Educational interventions	No intervention	3 RCTs	Agricultural workers	MA	Injuries	◀▶	n. a.	Rautiainen et al. 2008 [34]
Insurance premium discount program	No intervention	1 ITS	Agricultural workers	Individual study	A: Injuries, immediate effect B: Injuries, progressive effect	A: ▲ B: ◀▶	n. a.	Rautiainen et al. 2008 [34]

Abbreviations: CBA = controlled before-after studies; ITS = interrupted time series; MA = meta-analyses; n. a. = not available; RCTs = randomised controlled trials

Explanations for column Result: Effect direction: up arrow = positive health effects, down arrow = negative health effects, sideways arrows = mixed effects / contradictory results; statistical significance: black arrow: $p < 0.05$; grey arrow: $p > 0.05$; empty arrow: no statistical data reported

(From Teufer, 2019)

11.4 Developing evidence statements

In order to make the findings accessible and meaningful for our audiences, the review should generate clear evidence statements that answer the research questions.

These evidence statements are clear and succinct sentences which summarise the evidence findings, including the studies they are based on and the quality and certainty levels for each.

Examples of evidence statements

The following are from Robson et al. A systematic review of the effectiveness of occupational health and safety training. Scand J Work Environ Health. 2012 May;38(3):193-208:

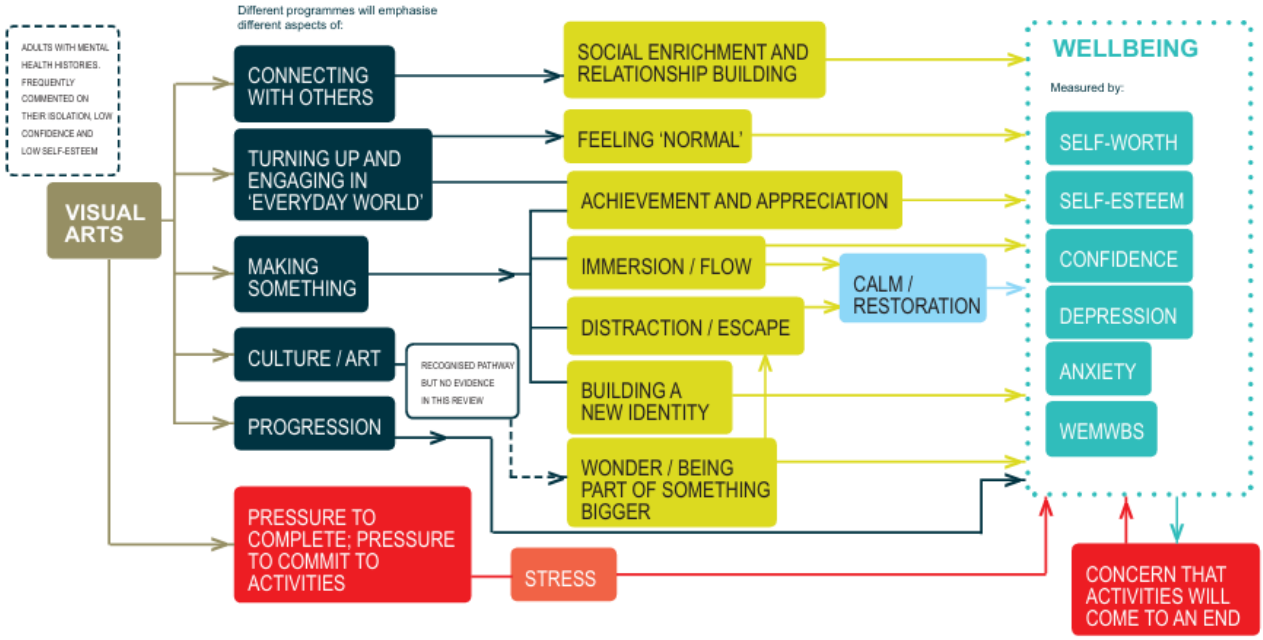
- Ten studies showed generally positive effects on worker behaviours following occupational health and safety training, and the evidence was deemed strong.
- Five trials showed a positive effect of occupational health and safety training on knowledge, but three studies were deemed poor quality so the evidence for this outcome was deemed insufficient.
- Ten studies that reported on health showed mixed/limited effects, so this evidence was insufficient.
- Three studies showed mixed/limited effects of training on attitudes and beliefs, which was also insufficient evidence.
- There was limited evidence on the effectiveness of higher versus lower engagement.

11.3 Developing a theory or model based on the evidence

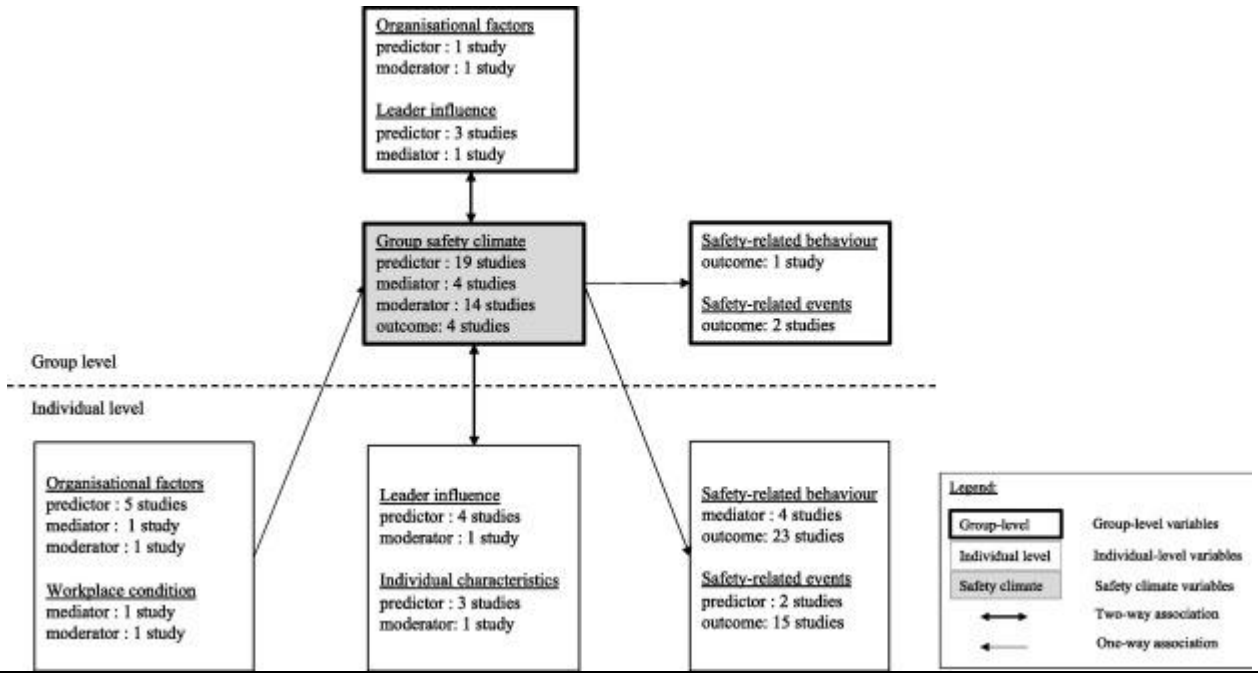
As part of the synthesis process the review team may be able to build a theory or model of the intervention or phenomena of interest. This may mean developing a theory of change showing how the intervention works, in what contexts and for whom. This theoretical model will help decision makers in designing or improving their interventions, and in identifying the contexts or mechanisms involved in successfully replicating effects.

Examples of review-based theories or models

Example of theory of change based on a systematic review of visual arts interventions for adults with mental health issues. This was developed by the What Works Centre for Wellbeing in 2018.



There are fewer examples of theories of change or models in the safety literature. One example shows the role of safety climate in workplace safety from Syed-Yahya (2022):



Further reading

33. Deeks JJ, Higgins JPT, Altman DG, McKenzie JE, Veroniki AA (editors). Chapter 10: Chapter 10: Analysing data and undertaking meta-analyses [last updated November 2024]. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.5. Cochrane, 2024. Available from cochrane.org/handbook.

12. Reporting

A full technical report should be prepared for all reviews, setting out the research questions, methods, findings and conclusions. Transparency, clarity and consistency are important to ensure that the findings are credible.

The main audience of our reviews are people who make decisions about safety in high-risk sectors across the world. These include policy makers, regulators, employers and supervisors, workers unions, charities and private companies. Although not all of them will read the technical report, it should be written so that it can be understood by someone with basic knowledge of the topic.

12.1 Using PRISMA reporting standards

Reviews should use PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) to guide their reporting. This is a set of guidelines to support the transparent and robust reporting of systematic reviews and meta-analyses. The main tools used in PRISMA reporting are a checklist and a flow diagram.

The PRISMA 2020 standards include a 27-item checklist for reporting a systematic review (and abstract), and a flow diagram to describe searches, screening and inclusion of studies. Although PRISMA was designed for use in systematic reviews, the principles are applicable across all types of reviews. The main checklist and flow diagram should be used for all reviews, with any amendments recorded. You can find the PRISMA checklist at the link or in Appendix C of this guide.

There are also several PRISMA extensions for other review types, including PRISMA-COSMIN for outcome measurement instruments, PRISMA-ScR for scoping reviews, and

PRISMA-LSR for living reviews. There are several extensions under development, including PRISMA-RR for rapid reviews.

12.2 Executive summary

All technical reports should include a short executive summary which will form the basis for the Centre's evidence briefings and other communications tools.

This executive summary should be written in plain English, for a global audience of lay people who are generally informed about the topic but not technically knowledgeable about the methods. In your summary you should define any unfamiliar terms and write out all abbreviations.

The executive summary should show which countries, groups of people and settings are included in the evidence, so that practitioners can judge how relevant the findings are for their contexts.

What to include in an executive summary

- Introduction
- Project aims/objectives and research questions
- Description of methods
- Findings and how they answer the research questions
- Implication and recommendations for research
- Implications and recommendations for practice

12.3 Illustrating and contextualising the findings

Evidence reviews by their nature show a broad picture rather than focusing on single studies or evaluations. This makes it hard for practitioners to visualise what the interventions or risks look like in practice. We recommend reviewers include a case study of an intervention or project as part of the executive summary to help readers visualise the context and implementation in real work settings.

What to include in an illustrative case study

- Title of the project or intervention
- Organisations involved
- Location, setting and context
- Populations involved
- Details of the project or intervention
- Details of implementation or delivery
- Details of the evaluation or study
- Outcomes and other learning

Other ways to contextualise the findings include bringing in the voices of practitioners or policy makers, key statistics or surveys, and initiatives that support evidence use in the field.

Further reading

34. Page M J, McKenzie J E, Bossuyt P M, Boutron I, Hoffmann T C, Mulrow C D et al. *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews* BMJ 2021; 372 :n71 doi:10.1136/bmj.n71

13. Communicating the findings

The Centre will publish the technical report and a briefing on the Global Safety Evidence Library as long as it meets the quality criteria set out in this guide.

The research team may choose to submit a paper based on this research to an open access journal or publish it on other websites.

The responsibility for communicating findings of a review is shared between the research team the Centre:

- **The research team:** will share the report and other publications through their academic and practitioner networks, conferences and other media.

- **The project consultation group:** will share the report and other publications with their networks and especially with individuals who are able to disseminate and use the evidence across high-risk sectors.
- **The Centre:** will develop a briefing summarising the findings of the review and setting it in a broader context. Promoting it through social media, press releases, blogs, videos and other communications.

Example of a Global Centre for Safety Evidence briefing:



Lloyd's Register Foundation, "Occupational safety and health interventions: The state of the evidence (Briefing)," Lloyd's Register Foundation, 2025. doi: 10.60743/SFJ8-KM98. Available on the Global Safety Evidence Library.

15. Living reviews of safety evidence

Living reviews are a type of systematic review that is continually updated, incorporating relevant evidence as it becomes available (Cochrane, 2019). The purpose of these reviews is to ensure decision makers always have the best available evidence available, especially on topics that are developing rapidly or where the evidence base is growing quickly.

Living reviews begin with a full systematic review, which is then updated regularly (usually on a monthly basis) by searches and analysis of new findings. Existing systematic reviews can also be converted to living reviews, though a new protocol will need to be published.

Living reviews can be costly and time consuming and require long-term commitments. Living reviews are worth doing when:

- The research question is a priority for decision makers,
- There is uncertainty in the existing evidence,
- There is likely to be emerging evidence on the topic.

As the Living Review Network puts it: "Embarking on an LSR is not a life sentence. It will be appropriate to cease this form of updating when the conditions specified above no longer hold" (Elliott, 2017). However, evidence from Covid-19 suggests that many living reviews are not updated due to waning commitment, loss of funding and the burden of screening. They often experienced delays in publishing updates and difficulty in communicating updates to stakeholders (De Silva 2025, Chen 2022, Zheng 2022). These factors should be weighed against the benefits on initiating a living review.

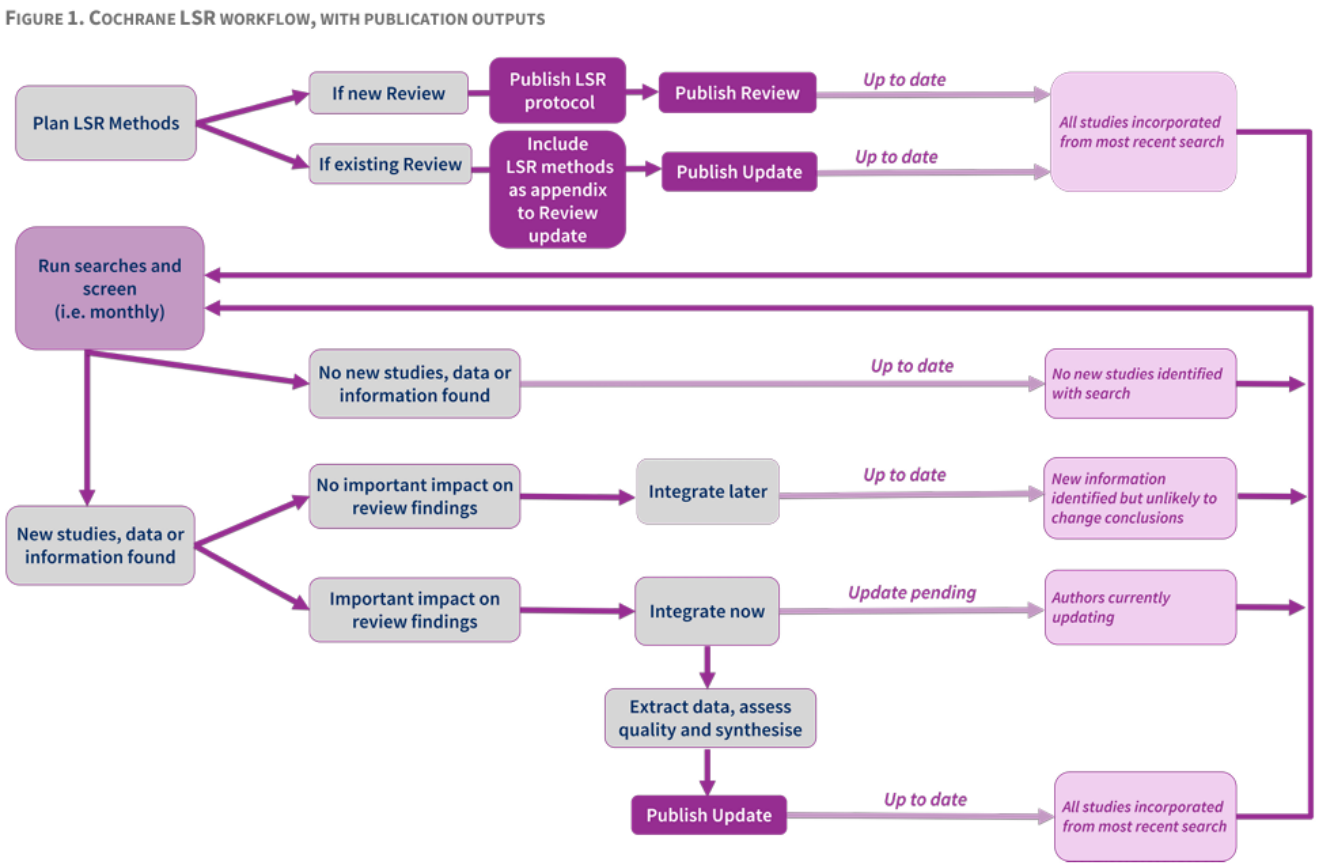
Living reviews are of particular interest for the Centre as several topics in global safety meet these criteria. The Centre team and external experts have a wealth of knowledge which can be used to determine which topics are suitable for living reviews.

Standards for living reviews are still being developed by the Living Systematic Review Network and others. It's likely that artificial intelligence tools will be used to automate updates and make living reviews interactive, though this approach is experimental and not widely tested. The Cochrane and Living Evidence Network Guidance should be a starting point for planning and the PRISMA-LSR extension should be used in reporting.

As well as the other aspects of the systematic review, the living review protocol should specify:

- How frequently evidence will be sought and screened,
- When and how new evidence is incorporated into the review,
- How and how often the findings will be communicated to the audience.

An example of the process involved in conducting and updating a living review is set out in this diagram from the Cochrane guidance:



Further reading

35. Cochrane Collaboration. Guidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode (Version December 2019). Cochrane Collaboration. https://community.cochrane.org/review-development/resources/living-systematic-reviews
36. De Silva, K., Turner, T. and McDonald, S. (2025), Cochrane's COVID-19 Living Systematic Reviews: A Mixed-Methods Study of Their Conduct, Reporting and Currency. Cochrane Evidence Synthesis and Methods, 3: e70024. https://doi.org/10.1002/cesm.70024
37. Z. Chen, J. Luo, S. Li, et al., "Characteristics of Living Systematic Review for COVID-19," *Clinical Epidemiology* **14** (2022): 925–935.
38. Q. Zheng, J. Xu, Y. Gao, et al., "Past, Present and Future of Living Systematic Review: A Bibliometrics Analysis," *BMJ Global Health* **7**, no. 10 (2022): e009378.

14. Case Study Synthesis

Case Study Synthesis is a systematic, transparent and pragmatic approach to synthesising evidence from practice. It's an emerging methodology which has not been widely applied to safety settings, so we encourage testing and development of the methods.

Safety project case studies are developed from the experience or tacit knowledge of practitioners implementing interventions, rather than a systematic process of enquiry. They are usually written by practitioners and people involved in the design and delivery of an intervention or activity, and contain a description of a project, information about organisations and participants, settings and context, as well as outcomes and impact. They may be based on an independent evaluation, internal monitoring data, or observation and feedback. They use a narrative structure and emphasise key learning around design, mechanisms of change, delivery and scaling, or unexpected outcomes (South, 2024; Brown et al., 2025).

Case studies are an underused but valuable form of evidence in safety reviews since they provide rich data on small-scale projects, pilot initiatives or innovation programmes - areas where published research is sparse or lags behind fast-moving

practice. Case study synthesis allows researchers to collate, review and synthesise this evidence in a systematic and robust way (Hardoon, 2021 and South, 2024).

The steps of a safety case study synthesis are very similar to those of other reviews:

1. **Developing research questions and protocol**: this should be done as described in sections 5 and 6.
    a. **Developing a conceptual framework**: identifying a theory or conceptual framework that helps define, categorise and select interventions or projects will be useful in finding relevant evidence and analysing results.
2. **Searching for evidence**: in case study banks or collections, or by hand-searching relevant websites (funders, sector bodies or practitioners). Some synthesis projects may involve commissioning new case studies from practitioners following a set template (Appendix D).
3. **Selecting studies**: the research protocol should set out the criteria for inclusion, and studies should be selected against this.
4. **Extracting data**: a structured template should be used to extract and organise data. This may be based on the template fields in Appendix D or on those in Table 1 of South 2024.
5. **Assessing the quality of included studies**: using a checklist of dimensions including integrity, transparency, completeness, responsibility, format, and learning reported. A sample framework is included in Appendix D. Results may be reported in aggregate rather than for each individual case study.
6. **Synthesis**: different methods may be applied depending on the research question and data quality. These may include framework analysis and cross-case analysis to identify patterns between cases while preserving within-case contextual data. An explanatory or conceptual framework may be used to interpret the findings.
7. **Reporting**: a narrative report of the results should include an overview of included studies and their content, a summary or table of quality assessment, a thematic report with contextual detail and direct quotations to answer the research questions.

Further reading

39. Brown, S.D., Dahill, D., Smith, S., Abreu Scherer, I. & King, D. (2025). Learning from Innovation: Case Study Synthesis of Safetytech Accelerator Projects. Nottingham: Nottingham Trent University. https://www.lrfoundation.org.uk/publications/learning-from-innovation
40. Hardoon, D., South, J., Southby, K., Freeman, C., Bagnall, A.-M., Pennington, A., & Corcoran, R. (2021). A guide to synthesising case studies. What Works Centre for Wellbeing. Available from: https://whatworkswellbeing.org/wp-content/uploads/2021/04/Guide-to-synthesising-case-studies-2021-FINAL-1.pdf
41. South J., Southby K., Freeman C., Bagnall A., Pennington A., Corcoran R. (2021). Community wellbeing case study synthesis. Technical report. What Works Centre for Wellbeing. Available from: https://whatworkswellbeing.org/wp-content/uploads/2018/05/Community-wellbeing-case-study-synthesis-Technical-Report-2021v2.pdf
42. South, J., Southby, K., Freeman, C., Bagnall, A.-M., Pennington, A., & Corcoran, R. (2024). Synthesising Practice-Based Case Study Evidence from Community Interventions: Development of a Method. International Journal of Qualitative Methods, 23. https://doi.org/10.1177/16094069241276964

Appendix A: Standards and manuals by review type

Review type	**Manuals and guides**	**Quality of included studies**	**Certainty (or quality) of review evidence**	**Reporting**
Systematic Review	Cochrane Handbook (Higgins 2024) JBI Manual (Aromataris, 2024)	(See table in section 10.1 or Appendix B)	GRADE GRADE-CERQual	PRISMA 2020
Scoping Review	JBI Scoping Reviews (Peters, 2020) Arksey & O'Malley (2005)	Not mandatory but preferred for our reviews. (See table in section 10.1 or Appendix B)	Not mandatory.	PRISMA-Scr (An update is under development)
Rapid Review or Rapid Evidence Assessment	WHO (2017) NCCMT/McMasters (Dobbins 2020) Cochrane Rapid Review Methods (Garrity, 2024)	(See table in section 10.1 or Appendix B)	GRADE GRADE-CERQual (modified). see Guidance on assessing certainty (Gartlehner, 2024).	PRISMA 2020, with adaptations. (A PRISMA-RR extension is under development)
Review of Reviews or	JBI Umbrella Reviews	AMSTAR-2 (or ROBIS)	GRADE (if included	PRISMA 2020, with

Umbrella Review	(Aromataris, 2020)		reviews have used GRADE).	adaptations.
Conceptual Review	No commonly used guideline but see Schreiber, 2022.	No commonly used guideline but see Schreiber, 2022.	No commonly used guideline but see Schreiber, 2022.	PRISMA 2020, with adaptations.
Methods or Measures Review	No specific guidance for safety measures. Suggest as for scoping or systematic reviews plus COSMIN Guideline	COSMIN risk of bias tool	Modified GRADE (see COSMIN guideline, section 6).	PRISMA-COSMIN (designed for health outcome measures, so may need adaptations).
Realist Review or Context-Mechanism-Outcomes Review	Adapted method for systematic reviews, or guidance from RAMESES Project.	RAMESES training materials.	RAMESES publication standards.	Consider PRISMA-CI extension for complex interventions. RAMESES publication standards.
Living Review	Cochrane and Living Evidence Network guidance	As for systematic reviews: (See table in section 10.1 or Appendix B)	As for systematic reviews: GRADE GRADE-CERQual	PRISMA-LSR
Case Study Synthesis	See South, 2024).	See Appendix D.	Not recommended.	See South 2024.

Appendix B: Assessing the quality of included studies

Type of evidence	Type of study	Checklist or tool
Quantitative	Randomised-control trials	Risk of Bias 2 (Rob2): A revised Cochrane risk-of-bias tool for randomized trials
	Non-randomised studies	Risk Of Bias In Non-Randomized Studies of Interventions (ROBINS-I V2)
	Cross-sectional	JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies
Qualitative	All qual studies	CASP (Critical Appraisal Skills Programme) for Qualitative Research
		JBI Critical Appraisal Checklist for Qualitative Research
	Case studies	See Appendix D
Mixed methods	Mixed-method studies	Mixed Methods Appraisal Tool (MMAT)
Other	Systematic reviews	AMSTAR-2 (or ROBIS)

Appendix C: PRISMA 2020 checklist

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	

Section and Topic	Item #	Checklist item	Location where item is reported
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	
Study characteristics	17	Cite each included study and present its characteristics.	
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	
Results of individual	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its	

Section and Topic	Item #	Checklist item	Location where item is reported
studies		precision (e.g. confidence/credible interval), ideally using structured tables or plots.	
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	
	23b	Discuss any limitations of the evidence included in the review.	
	23c	Discuss any limitations of the review processes used.	
	23d	Discuss implications of the results for practice, policy, and future research.	
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

Section and Topic	Item #	Checklist item	Location where item is reported
other materials			

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. This work is licensed under CC BY 4.0.
More info on the PRISMA site.

Appendix D: Case Study Synthesis quality framework and practitioner template

D.1 Quality of included studies framework

The quality assessment framework comprises of five domains, each reflecting a key dimension of quality relevant to practice-based case studies. Each domain may be scored on a scale from 0 to 4, with higher scores indicating stronger performance in that domain. The maximum score for each case study is 20 points (5 domains x 4 points each).

Quality Assessment Tool Structure
The assessment framework comprised five domains, each reflecting a key dimension of quality relevant to practice-based case studies. Each domain was scored on a scale from 0 to 4, with higher scores indicating stronger performance in that domain. The maximum possible score for each case study was 20 points (5 domains × 4 points each).

Domains and Scoring Criteria

1. Integrity (0–4 points)
    - Assesses the accuracy, honesty, and reliability of the case study.
    - Criteria include: clear description of context, transparency about methods, and avoidance of selective reporting.
2. Transparency (0–4 points)
    - Evaluates the openness with which the case study reports its processes and findings.
    - Criteria include: explicitness about data sources, clarity in reporting outcomes, and disclosure of limitations.
3. Completeness (0–4 points)
    - Measures the extent to which the case study provides a full account of the intervention or innovation.
    - Criteria include: coverage of background, implementation, outcomes, and lessons learned.
4. Responsibility (0–4 points)
    - Examines the ethical and social responsibility demonstrated in the case study.
    - Criteria include: attention to participant consent, safeguarding, and consideration of wider impacts.
5. Format and Learning Reported (0–4 points)
    - Assesses the accessibility and usefulness of the case study for learning and future application.

- Criteria include clarity of writing, use of illustrative examples, and articulation of transferable lessons.

Scoring
Each domain should be scored independently by two reviewers, using a rubric that specified what constituted a score of 0 (absent), 1 (limited), 2 (adequate), 3 (good), or 4 (excellent) for each criterion. The total quality score for each case study is calculated out of a possible 20 points. Reviews are then compared and moderated.

Reporting
Since case studies may be developed by practitioners for a number of reasons not related to research, it may not be appropriate to report the scores for each case study individually. It may be more suitable to report on overall trends across the included case studies and reflect on the dimensions which were stronger or weaker among the set under study.

This quality assessment approach was developed by Brown et al. (2025) for a synthesis of SafetyTech innovation case studies, and based on original methods from South (2024). Further development and adaptation may be needed to tailor this approach to other safety topics and practice.

D.2 Case Study Template for practitioners

Some synthesis projects may involve commissioning new case studies from practitioners following a template. The following may be used or amended as necessary.

1. Name of your organisation
2. Your name and role in the project Who is writing this case study?
3. Project title
4. Project dates When the project took place, including start and finish dates, key design and delivery periods, or whether the project is ongoing. The date you wrote this case study.
5. Project resources and costs (100 words) How much did the project cost and who funded it? What other resources did you have for the project (e.g. staff with technical expertise, access to buildings and workplaces)
6. Setting (100 words) The specific location/s and context/s in which the project took place.
7. Safety challenges and goals (100 words)

The safety issue or problem your project sought to address, and how you identified it. The aims and objectives of your project and the difference you wanted to make. What does 'safety' mean and look like in this setting and for these people?
8. Activities and interventions (200 words) The activities and delivery of the project (including new technology used, training delivered, equipment tested, etc.)
9. People and partners (150 words) Who took part in the project: the number of people and information about them (e.g. age, ethnicity, gender, occupation)? How did they get involved with your project, and what were their motivations for taking part? Which organisations were involved in delivering or supporting the project, and what were their roles? What other stakeholders were involved?
10. Project outcomes and impact (200 words) What changed as a result of your project in the short and long term? What impact did the project have on your or partner organisations, on the people directly affected by the safety issue, or on the wider community or sector? Did some people benefit more than others? Did some outcomes last longer than others? Were there any unexpected or negative outcomes?
11. Evaluation and data collection (150 words) Was the project evaluated (if so, how)? What data did you collect to help you understand what changed? Who collected and analysed the data (you or an independent body)? What other evidence informed this case study (your observation, official data, participant feedback)?

12. Next steps and sustainability (150 words) What will happen as a result of the project? Is the project continuing or completed? Will the project be rolled out or scaled? If you developed a new product or technology, what is the next step in developing it for market?
13. Key learning (200 words) What is the most important thing you have learned from this project? The barriers and enablers that determined the success of your project? What do you want other practitioners working in this field to know?
14. Further information Links to project or partner websites, evaluation reports, media or other materials, etc. Your contact details if you're happy for people to get in touch to find out more.

References

1. Main handbooks and guides

Cochrane: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.5 (updated August 2024). Cochrane, 2024. Available from www.training.cochrane.org/handbook.

JBI: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-01

GRADE: Neumann I, Schünemann H (Editors). The GRADE Book version 1.0 (updated September 2024). The GRADE Working Group. Available from: https://book.gradepro.org

GRADE-CERQual: Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, Bohren MA, Tunçalp Ö, Colvin CJ, Garside R, Carlsen B, Langlois EV, Noyes J. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci.* 2018 Feb 1;13(Suppl 1):2. doi:10.1186/s13012-017-0688-3. PMID:29384079; PMCID: PMC5791031. Available here.

PRISMA: 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. This work is licensed under CC BY 4.0.
Available from PRISMA site.

2. Other references

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Aromataris E, Fernandez R, Godfrey C, Holly C, Khalil H, Tungpunkom P. Umbrella Reviews (2020). Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-08

Brown, S.D., Dahill, D., Smith, S., Abreu Scherer, I. & King, D. (2025). Learning from Innovation: Case Study Synthesis of Safetytech Accelerator Projects. Nottingham: Nottingham Trent University. https://www.lrfoundation.org.uk/publications/learning-from-innovation.

Campbell, F., Tricco, A.C., Munn, Z. et al. Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same but different — the "Big Picture" review family. *Systematic Review,* 12, 45 (2023). https://doi.org/10.1186/s13643-023-02178-5

Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York: University of York (2009). https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf.

Z. Chen, J. Luo, S. Li, et al., "Characteristics of Living Systematic Review for COVID-19," *Clinical Epidemiology* **14** (2022): 925–935.

Clark, Justin, Belinda Barton, Loai Albarqouni, Oyungerel Byambasuren, Tanisha Jowsey, Justin Keogh, Tian Liang, Christian Moro, Hayley O'Neill, and Mark Jones. "Generative Artificial Intelligence Use in Evidence Synthesis: A Systematic Review." *Research Synthesis Methods*, 2025, 1–19. https://doi.org/10.1017/rsm.2025.16.

Cochrane Collaboration. Guidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode (Version December 2019). Cochrane Collaboration. https://community.cochrane.org/review-development/resources/living-systematic-reviews

De Silva, K., Turner, T. and McDonald, S. (2025), Cochrane's COVID-19 Living Systematic Reviews: A Mixed-Methods Study of Their Conduct, Reporting and Currency. Cochrane Evidence Synthesis and Methods, 3: e70024. https://doi.org/10.1002/cesm.70024

Dobbins M. Rapid review guidebook: steps for conducting a rapid review. National Collaborating Centre for Methods and Tools. Available from: https://www.nccmt.ca/uploads/media/media/0001/01/a816af720e4d587e13da6bb307df8c907a5dff9a.pdf.

Elliott, J. H., Synnot, A., Turner, T., Simmonds, M., Akl, E. A., McDonald, S., Salanti, G., Meerpohl, J. J., MacLehose, H., Hilton, J., Tovey, D., Shemilt, I., & Thomas, J., for the Living Systematic Review Network. (2017). Living systematic review: 1. Introduction — The why, what, when, and how. Journal of Clinical Epidemiology, 91, 23–30. https://doi.org/10.1016/j.jclinepi.2017.08.010

Garritty C, Hamel C, Trivella M, Gartlehner G, Nussbaumer-Streit B, Devane D et al. Updated recommendations for the Cochrane rapid review methods guidance for rapid reviews of effectiveness. *BMJ* 2024;384:e076335. doi:10.1136/bmj-2023-076335

Gartlehner G, Nussbaumer-Streit B, Devane D on behalf of the Cochrane Rapid Reviews Methods Group, et al. Rapid reviews methods series: Guidance on assessing the certainty of evidence. *BMJ Evidence-Based Medicine* 2024;29:50–54.

Hardoon, D., South, J., Southby, K., Freeman, C., Bagnall, A.-M., Pennington, A., & Corcoran, R. (2021). A guide to synthesising case studies. What Works Centre for Wellbeing. Available from: https://whatworkswellbeing.org/wp-content/uploads/2021/04/Guide-to-synthesising-case-studies-2021-FINAL-1.pdf

Hempel, S., Xenakis, L., Danz, M. (2016). Systematic reviews for occupational safety and health questions: Resources for evidence synthesis. RAND Corporation. https://doi.org/10.13140/RG.2.1.2597.0167

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). Cochrane, 2024. Available from www.training.cochrane.org/handbook.

Kolaski, K., Logan, L.R. & Ioannidis, J.P.A. Guidance to best tools and practices for systematic reviews. *Syst Rev* 12, 96 (2023). https://doi.org/10.1186/s13643-023-02255-9

Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin CJ, Gülmezoglu M, Noyes J, Booth A, Garside R, Rashidian A. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Med*. 2015 Oct 27;12(10):e1001895. doi:10.1371/journal.pmed.1001895. PMID: 26506244; PMCID: PMC4621438.

Montgomery P, Movsisyan A, Grant SP, Macdonald G, Rehfuess EA. Considerations of complexity in rating certainty of evidence in systematic reviews: a primer on using the GRADE approach in global health. *BMJ Glob Health*. 2019 Jan 25;4(Suppl 1):e000848. doi:10.1136/bmjgh-2018-000848. PMID:30775013; PMCID:PMC6350753.

Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil, H. Scoping Reviews (2020). Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. JBI; 2024. Available from: https://synthesismanual.jbi.global. https://doi.org/10.46658/JBIMES-24-09

Popay, Jennie & Roberts, Helen & Sowden, Amanda & Petticrew, Mark & Arai, Lisa & Rodgers, Mark & Britten, Nicky & Roen, Katrina & Duffy, Steven. (2006). Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC Methods Programme. 10.13140/2.1.1018.4643.

Robson LS, Stephenson CM, Schulte PA, Amick BC 3rd, Irvin EL, Eggerth DE, Chan S, Bielecky AR, Wang AM, Heidotting TL, Peters RH, Clarke JA, Cullen K, Rotunda CJ, Grubb PL. A systematic review of the effectiveness of occupational health and safety training. Scand J Work Environ Health. 2012 May;38(3):193-208. doi: 10.5271/sjweh.3259. Epub 2011 Nov 1. PMID: 22045515.

Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Development of an efficient search filter to retrieve systematic reviews from PubMed. *J Med Libr Assoc*. 2021 Oct 1;109(4):561–574.

Schreiber, F., & Cramer, C. (2022). Towards a conceptual systematic review: proposing a methodological framework. *Educational Review*, *76*(6), 1458–1479. https://doi.org/10.1080/00131911.2022.2116561

Snape D, Meads C, Bagnall A, Tregaskis O, Mansfield L. What Works Wellbeing: a guide to our evidence review methods. What Works Centre for Wellbeing; 2019 Apr. Available from: https://whatworkswellbeing.org/wp-content/uploads/2020/02/WWCW-Methods-Guide-FINAL-APRIL-2019a.pdf.

South J., Southby K., Freeman C., Bagnall A., Pennington A., Corcoran R. (2021). Community wellbeing case study synthesis. Technical report. What Works Centre for Wellbeing. Available from: https://whatworkswellbeing.org/wp-content/uploads/2018/05/Community-wellbeing-case-study-synthesis-Technical-Report-2021v2.pdf

South, J., Southby, K., Freeman, C., Bagnall, A.-M., Pennington, A., & Corcoran, R. (2024). Synthesising Practice-Based Case Study Evidence From Community Interventions: Development of a Method. International Journal of Qualitative Methods, 23. https://doi.org/10.1177/16094069241276964

Sutton, A., Clowes, M., Preston, L. and Booth, A. (2019), Meeting the review family: exploring review types and associated information retrieval requirements. Health Info Libr J, 36: 202-222. https://doi.org/10.1111/hir.12276

Syed-Yahya SNN, Idris MA, Noblet AJ. The relationship between safety climate and safety performance: A review. *Journal of Safety Research*. 2022;83:105–118. Available from: https://doi.org/10.1016/j.jsr.2022.08.008

Q. Zheng, J. Xu, Y. Gao, et al., "Past, Present and Future of Living Systematic Review: A Bibliometrics Analysis," *BMJ Global Health* **7**, no. 10 (2022): e009378.

Teufer B, Ebenberger A, Affengruber L, et al. Evidence-based occupational health and safety interventions: a comprehensive overview of reviews. *BMJ Open* 2019;9:e032528. doi:10.1136/bmjopen-2019-032528

Tricco AC, Langlois EV, Straus SE, editors. *World Health Organization*. (2017). Rapid Reviews to Strengthen Health Policy and Systems: A Practical Guide. Available from: https://iris.who.int/bitstream/handle/10665/258698/9789241512763-eng.pdf

van der Molen HF, Basnet P, Hoonakker PLT, Lehtola MM, Lappalainen J, Frings-Dresen MHW, Haslam R, Verbeek JH. Interventions to prevent injuries in construction workers. Cochrane Database of Systematic Reviews 2018, Issue 2. Art. No.: CD006251. DOI: 10.1002/14651858.CD006251.pub4. Appendix 1.

Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, et al. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *PLoS Med* 2012, 9(10):e1001333. doi: 10.1371/journal.pmed.1001333

Wong, G., Greenhalgh, T., Westhorp, G. *et al*. RAMESES publication standards: realist syntheses. *BMC Med* **11**, 21 (2013). https://doi.org/10.1186/1741-7015-11-21