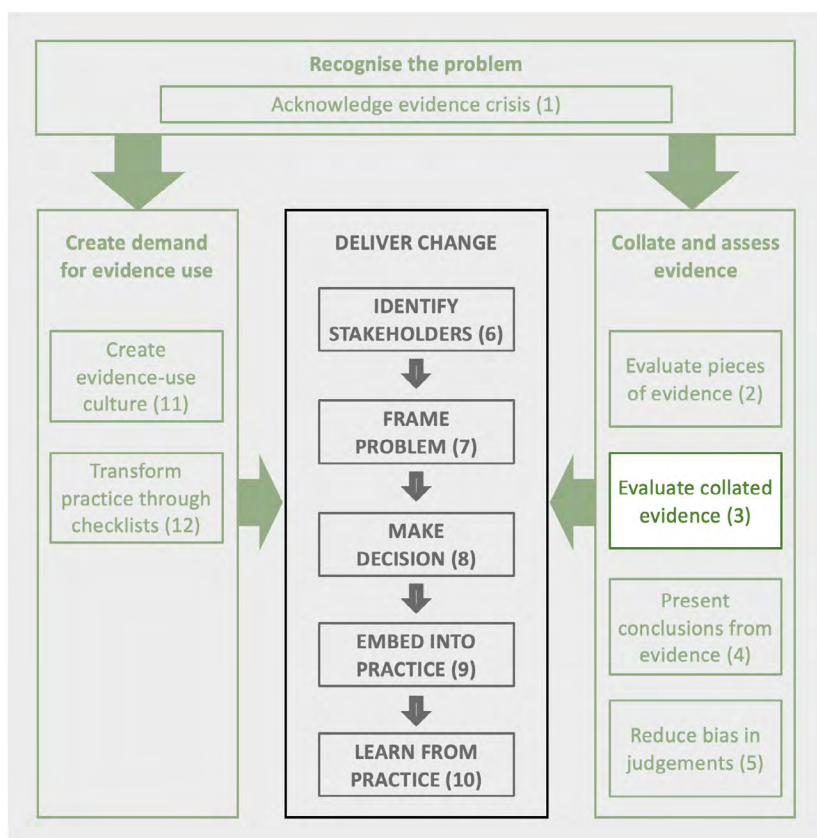


3. Assessing Collated and Synthesised Evidence

William J. Sutherland^{1,2}, Samantha H. Cheng^{3,4}, Alec P. Christie^{1,2,5}, Steven J. Cooke⁶, Julia Koricheva⁷, Nicola Randall⁸, Rebecca K. Smith^{1,2}, Ruth Stewart⁹

Multiple pieces of evidence can be brought together in a variety of ways including systematic maps, subject-wide evidence synthesis, systematic reviews, rapid evidence assessment, meta-analysis and collated open access effect sizes. Each has different uses. This chapter describes each along with suggestions for how to interpret the results and assess the evidence.



¹ Conservation Science Group, Department of Zoology, University of Cambridge, The David Attenborough Building, Pembroke Street, Cambridge, UK

² Biosecurity Research Initiative at St Catharine's College (BioRISC), St Catharine's College, Cambridge, UK.

³ Center for Biodiversity and Conservation, American Museum of Natural History, New York, NY 10026, USA

⁴ Global Science, World Wildlife Fund, Washington, D.C. 20037, USA

⁵ Downing College, Regent Street, Cambridge, UK

⁶ Canadian Centre for Evidence-Based Conservation, Institute of Environmental and Interdisciplinary Science, Carleton University, Ottawa, ON, K1S 5B6, Canada

⁷ Department of Biological Sciences, Royal Holloway University of London, Egham, Surrey, UK

⁸ Centre for Evidence-Based Agriculture, Harper Adams University, Newport, Shropshire, UK

⁹ Africa Centre for Evidence, University of Johannesburg, Bunting Road Campus, 2006, Johannesburg, South Africa

Contents

3.1 Collating the Evidence

3.2 Systematic Maps

3.3 Subject-Wide Evidence Syntheses

3.4 Systematic Reviews

3.5 Rapid Evidence Assessments

3.6 Meta-Analyses

3.7 Open Access Effect Sizes

3.8 Overviews of Reviews

References

3.1 Collating the Evidence

The previous chapter considered how to assess single pieces of evidence. However, there are often multiple pieces of evidence and these should all be assessed. Much of the progress in evidence-based practice has revolved around means of collating and interpreting all the evidence relating to an issue. A range of approaches has been developed to collate evidence for evidence users to use to inform policy, decisions and practice. We summarise those approaches and provide an assessment of how the various outputs can be used and interpreted.

3.2 Systematic Maps

3.2.1 What is a systematic map?

Systematic maps, sometimes known as evidence maps, or evidence gap maps, are reproducible, transparent, analytical methods to collate and organise an evidence base using a decision-relevant framework. Unlike systematic reviews, systematic maps do not aim to answer questions of effectiveness or direction of impact, rather they are particularly useful to understand the extent and nature of the evidence base on a broad topic area (James et al., 2016). Systematic maps can help describe the distribution of existing evidence, highlighting areas of significant research effort and where key gaps exist. This can provide information to guide research, prioritise evaluation, and illustrate where there may be inadequate information to inform decision making (McKinnon et al., 2015). In recent years, systematic maps (and other types of evidence maps) have grown in popularity across different disciplines. In conservation and development, maps are often published in the journal *Environmental Evidence* and are conducted and published by organisations such as the International Initiative for Impact Evaluations (3ie).

3.2.2 How are systematic maps created?

Given the wide range of potential decisions and users that a systematic map may inform, the scope and framework of a systematic map are usually co-developed with a representative group of stakeholders. This is particularly critical for conservation as the challenges facing natural ecosystems require thinking across disciplines – for example linking climate change, nature conservation, and sustainable development. The evidence base for multi-disciplinary topics is likely to be heterogeneous, particularly in the types of terminology that different sectors use to describe interventions and outcomes. Previous work has shown that semantic variability is quite high even within a single discipline of conservation (Westgate and Lindenmayer, 2017), thus developing an agreed-upon framework of interventions/exposures and outcomes is critical to ensure that the finished product is interpretable and salient to the range of potential end users. For example, frameworks of interventions/exposures and outcomes are often grounded in some type of causal theory that describes how one expects these elements are linked (Cheng et al., 2020). This type of causal grounding is important to help the synthesis team interpret findings based on how plausible a causal relationship might be and thus whether an evidence gap is truly a gap or not.

Systematic maps share many characteristics, and therefore methodological processes, with systematic reviews. As the value of these approaches lies in their transparent and reproducible methods, which explicitly aim to account for potential biases, an *a priori* protocol should be developed that details systematic steps for designing and implementing a search strategy: screening, coding, and analysing included articles. This protocol ideally should be critiqued by those with methodological and topic expertise. While the methods for assembling an evidence base for a systematic map are similar to those for a systematic review, there are a few key differences. Systematic maps do not always include a critical appraisal of included studies and, instead of a formal synthesis of findings from included studies, the outputs may include summary tables, heatmaps and searchable databases or spreadsheets of the included evidence. Figure 3.1 shows the typical stages of a systematic map.

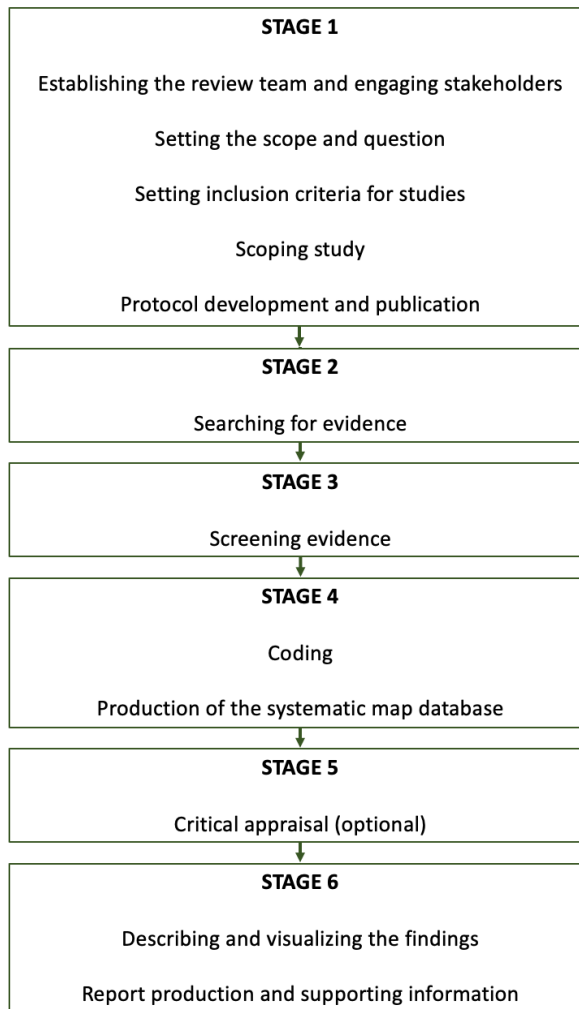


Figure 3.1 Typical stages of a systematic mapping process. (Source: James et al., 2016, CC-BY-4.0)

Searching for studies to be included in systematic maps may involve covering a broader range of sources and evidence types than a typical systematic review. For example, a systematic map may seek to characterise the range of existing knowledge on both interventions and outcomes in the framework, along with important contextual factors or conditions. Thus, this type of map may cover a range of various study types, not only studies focused on impacts, in order to capture all concepts within their framework. Some maps may also seek to characterise novel or emerging topic areas — so they may include evidence types such as datasets, citizen surveys, and practitioner knowledge, in addition to more traditional research studies.

3.2.3 How can systematic maps be used?

Systematic maps tend to have multiple dimensions given the wide range of potential queries that end users may have. Studies are typically categorised across the framework and represented as a ‘heat map’ that illustrates the distribution of studies. These maps are often disaggregated by a number of factors that can help better navigate where evidence exists and where it does not, for example by geography, publication date, demographic variables, and/or ecosystems. These factors can be used to ‘slice’ the data in the systematic map to look only at a subset of studies, such as those from a specific country or ecosystem. They can also be used to enable the reader of the map to drill down into a cell to find out more about the evidence contained within.

Given the potentially endless combination of factors that users may be interested to slice a map by, many have advocated interactive tools that can improve the accessibility and usability of maps for end users. For example, dedicated interactive websites have emerged featuring data from different systematic maps (e.g. Evidence for Nature and People Data Portal; Evidensia; 3ie). In addition, there is a range of different tools that aimed to help evidence synthesis teams produce interactive diagrams such as EPPI-Mapper (<https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3790>) and EviAtlas (Haddaway et al., 2019, <https://estech.shinyapps.io/eviatlas/>). As with question formulation and framework development, scoping and building an interactive platform should be undertaken in collaboration with a representative group of end users through user-centred design to ensure widespread usability.

Throughout this section, we have emphasised that systematic maps, given their broad nature, are useful for different purposes for different users. Thus, interpreting a systematic map can often be challenging for end users. Interpreting the distribution of the evidence base captured in a systematic map should be informed by individual decision needs and calibrated based on a set of key criteria. First, users should determine what volume of evidence is sufficient to inform their decision. For example, if a systematic map reveals that only a few studies exist for a specific linkage in the framework, an end user should consider how risky it might be to base a decision upon that. Second, users should determine how robust evidence needs to be to inform their decision. For example, if a systematic map reveals a high volume of studies but relatively few appear to meet the level of rigour required for a decision, more robust research may be needed. Third, users should consider how plausible that linkage is. Does it represent a causal relationship that is supported by theory? Ultimately, synthesis teams should collaborate with

stakeholders to interpret maps and make clear what types of decisions the interpretations are relevant for and provide guidance for other users to calibrate their interpretations accordingly.

The Collaboration for Environmental Evidence Synthesis Appraisal Tool (CEESAT, <https://environmentalevidence.org/ceeder/about-ceesat/>) enables users to appraise the rigour, transparency and limitations of methods of existing reviews and includes a checklist designed specifically for evidence overviews such as systematic maps. The checklist (3.1) enables users to categorise each stage of the process reported in a systematic map or other evidence overview as either gold, green, amber or red to help inform users of the level of confidence that they may have in the reported findings.

3.1 Checklists for evidence reviews and systematic maps

From The Collaboration for Environmental Evidence Synthesis Appraisal Tool (CEESAT).

Type of evidence collation	Systematic reviews/evidence reviews	Systematic maps/evidence overviews
Version, date	Version 2.1, 29th July 2021	Version 2.1, 29th July 2021
The review question	<input type="checkbox"/> Are the elements of the review question clear?	<input type="checkbox"/> Are the elements of the review question clear?
The method/protocol	<input type="checkbox"/> Is there an <i>a-priori</i> method/protocol document?	<input type="checkbox"/> Is there an <i>a-priori</i> method/protocol document?
Searching for studies	<input type="checkbox"/> Is the approach to searching clearly defined, systematic and transparent? <input type="checkbox"/> Is the search comprehensive?	<input type="checkbox"/> Is the approach to searching clearly defined, systematic and transparent? <input type="checkbox"/> Is the search conducted in line with the defined search scope?

Type of evidence collation	Systematic reviews/evidence reviews	Systematic maps/evidence overviews
Including studies	<input type="checkbox"/> Are eligibility criteria clearly defined? <input type="checkbox"/> Are eligibility criteria consistently applied to all potentially relevant articles and studies found during the search? <input type="checkbox"/> Are eligibility decisions transparently reported?	<input type="checkbox"/> Are eligibility criteria clearly defined? <input type="checkbox"/> Are eligibility criteria consistently applied to all potentially relevant articles and studies found during the search? <input type="checkbox"/> Are eligibility decisions transparently reported?
Critical appraisal (not assessed for overviews)	<input type="checkbox"/> Does the review critically appraise each study? <input type="checkbox"/> During critical appraisal was an effort made to minimise subjectivity?	
Data extraction/data coding	<input type="checkbox"/> Is the method of data extraction fully documented? <input type="checkbox"/> Are the extracted data reported for each study? <input type="checkbox"/> Were extracted data cross checked by more than one reviewer?	<input type="checkbox"/> Is the method of data coding fully documented? <input type="checkbox"/> Are the coded data reported for each study? <input type="checkbox"/> Were coded data cross checked by more than one reviewer?

Type of evidence collation	Systematic reviews/evidence reviews	Systematic maps/evidence overviews
Data synthesis (not assessed for overviews)	<ul style="list-style-type: none"> <input type="checkbox"/> Is the choice of synthesis approach appropriate? <input type="checkbox"/> Is a statistical estimate of pooled effect (or similar) provided together with a measure of variance and heterogeneity among studies? <input type="checkbox"/> Is variability in the study findings investigated and discussed? 	
Review limitations	<ul style="list-style-type: none"> <input type="checkbox"/> Have the authors considered the limitations of the synthesis? 	<ul style="list-style-type: none"> <input type="checkbox"/> Have the authors considered the limitations of the synthesis?

(Source: <https://environmentalevidence.org/ceeder/about-ceesat>)

These checklists can be downloaded from <https://doi.org/10.11647/OBP.0321#resources>. They can be modified and tailored for specific uses.

A database of systematic maps that have had CEESAT criteria independently applied can be viewed on the CEEDER website (<https://environmentalevidence.org/ceeder-search>). Eligible evidence reviews are rated by a pool of CEEDER Review College members who rate their reliability using CEESAT criteria. Several Review College members apply CEESAT for each evidence overview and disagreements in ratings are resolved by an editorial team.

3.3 Subject-Wide Evidence Syntheses

Subject-wide evidence synthesis was created as a methodology to collate evidence and make it easily accessible to decision makers (Sutherland et al., 2019). It enables the rapid synthesis of

evidence across entire subject areas (comprising tens or hundreds of related review questions), whilst being transparent, objective and minimising bias. End users (practitioners, policy-makers and researchers) are involved in the process to ensure applicability and to encourage uptake.

Using the process of subject-wide evidence synthesis, Conservation Evidence (www.conservationevidence.com) provides a freely accessible, plain-English database, which contains evidence for the effects of conservation interventions (i.e. actions that have been or could be used to conserve biodiversity).

One of the first stages of subject-wide evidence synthesis is solution scanning (Section 7.6.1) to produce a comprehensive list of all actions that have been tried or suggested for the subject of the synthesis and that could realistically be implemented. In Conservation Evidence this list is developed in collaboration with an advisory board and is structured using IUCN Threat Categories and Action Categories (<https://www.iucnredlist.org/resources/classification-schemes>); systematic searches of the literature (e.g. so far 650 academic journals and 25 report series from organisational websites have been searched by Conservation Evidence) are undertaken and relevant studies that test the effectiveness of an action are summarised in a short standardised paragraph in plain English.

The broad diversity of output measures, such as species abundance, species diversity, behaviour, breeding performance, etc., however, represents a serious challenge hence formal analysis, such as meta-analysis, can often be impossible. To overcome this issue, Conservation Evidence uses a panel of experts who assess each action in terms of effectiveness, certainty (strength of the evidence), and harms to the subject under consideration (Sutherland et al., 2021). A modified Delphi Technique (see Section 5.5.1.) is used with at least two rounds of anonymous scoring. Figure 3.2 shows how the resulting median scores are converted into categories of effectiveness.

The main outputs of subject-wide evidence synthesis undertaken by Conservation Evidence are:

- Database of studies testing actions: By August 2022, over 8,400 studies had been summarised. Study summaries, although brief, aim to include sufficient detail of the study context and methods to allow users to begin assessing reliability and relevance to their own system (e.g. species, location, implementation method).
- Database of actions with summaries of studies testing them: By August 2022, over 3,650 actions for the conservation of habitats, species groups, and other conservation issues had been assessed. For each action, background information is provided where necessary to place in context or to refer to additional information other than tests of effectiveness. Key messages provide a brief overview or index to the studies summarised; a location map of studies is also provided. The overall 'category of effectiveness' for each action is generated through an assessment of the summarised evidence by an expert panel (academics, practitioners and policy-makers; Figure 3.2; Sutherland et al., 2021).

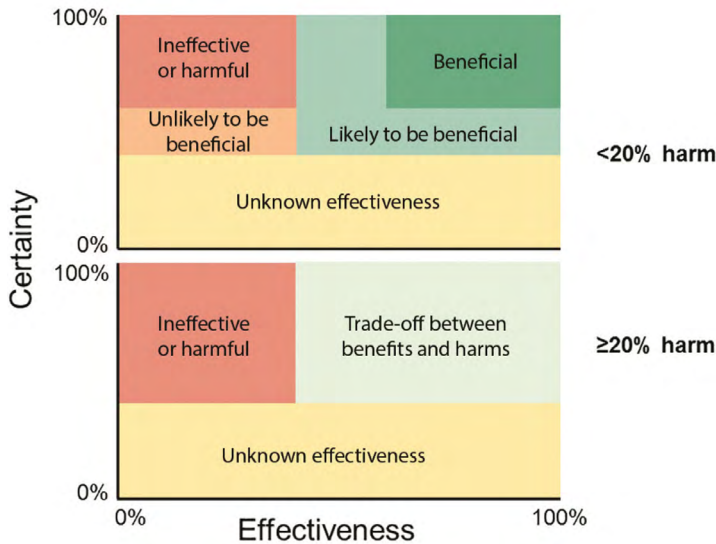


Figure 3.2 Categories of effectiveness based on a combination of effectiveness (the extent of the benefit and harm) and certainty (the strength of the evidence). The top graph refers to interventions with harms scored $<20\%$ and the bottom graph to interventions with harms $\geq 20\%$. (Source: Sutherland et al. 2021, CC-BY-4.0)

- Database of titles of non-English language studies testing actions: By August 2022, studies had been collated by scanning 419,679 paper titles from 330 journals in 16 different languages.
- Synopses: All of the actions relevant to a specific subject are grouped into a subject ‘synopsis’. By August 2022, evidence for 24 different taxa or habitats had been collated including mammal, bird, forest, and peatland conservation.
- *What Works in Conservation*: An annual update of the information on the effectiveness of actions is produced as a book, *What Works in Conservation*. All the information within each update is also presented on the website.

The Conservation Evidence databases can be used for a range of issues, for example:

- *Determining whether a specific action is effective*: e.g. are bat bridges effective? The database can be used to provide an indication of the comparative effectiveness of different actions that could be taken (and where evidence is lacking). However, it is important to look at the evidence for each of the most relevant actions in more detail to determine how relevant it is to the user’s system before making decisions (e.g. species studied, location, implementation methods tested, etc.).
- *Identifying actions to mitigate against a specific threat*: e.g. how could road deaths for large mammals be reduced? Refining a search on the Actions page (<https://>

www.conservationevidence.com/data/index) by ‘Category: Terrestrial Mammal Conservation’, and ‘Threat: Transportation & service corridors’ gives 39 actions with associated evidence. Relevant actions can also be found by using keywords — search terms should be kept simple and variants tried.

- *Identifying possible actions for a specific habitat or taxa:* e.g. peat bogs. There are 125 actions for managing peat bogs, with summaries of studies testing them.
- *Determining what actions have been tested for a specific species.*
- *As a source of literature in 17 languages for a literature review or systematic review.*
- *Identifying studies that relate to a specific habitat, taxa or country:* e.g. refining a search by ‘Habitat: Deserts’ indicates that there are currently 106 studies in deserts, similarly there are 13 studies on horseshoe bats and 14 in Zimbabwe. This might be useful if starting work on a given habitat, taxa or region.

3.4 Systematic Reviews

Systematic review entails using repeatable analytical methods to extract secondary data and analyse it (Pullin et al., 2020) and can also be carried out for qualitative studies (Flemming et al., 2019). In conservation, these are listed in the CCE library (<https://environmentalevidence.org>).

A systematic review typically involves the following eight stages (Collaboration for Environmental Evidence, 2018):

1. Planning a synthesis
2. Developing a protocol
3. Conducting a search
4. Eligibility screening
5. Data coding and data extraction
6. Critical appraisal of study validity
7. Data synthesis
8. Interpreting findings and reporting conduct

ROSES (RepOrting standards for Systematic Evidence Syntheses) provides forms that can be used during the preparation of systematic review and map protocols and final reports. The checklists can be downloaded from <https://www.roses-reporting.com>.

Although systematic reviews are designed to be more robust than traditional reviews, the methods used and the question addressed within a review may impact the confidence that can be placed in the findings of any review. The clarity of the question, the methods used for collating evidence, decisions around critical appraisal and synthesis of included studies, and transparency of reporting can all influence the interpretation of review findings.

The CEESAT checklist (3.1) has been developed to assess the rigour of each stage of any review (systematic and otherwise) in the field of environmental management. The tool can be used by users of reviews to assess the methods used, the transparency of reporting of a review, and the likely limitations that there may be. Criteria for assessing each element of a review are given in Woodcock et al. (2014); also see Section 3.2. A database of evidence reviews that have already had CEESAT criteria independently applied by a group of experts is available for reviews from 2018 onwards on the CEEDER website (<https://environmentalevidence.org/ceeder-search>) and can be easily searched by subtopic area.

Other questions should be asked when applying the results of a review. For example, ask when was the literature searched and does the gap in the literature matter, is there a geographical constraint to the review and how does that match the area of interest, how does the subject match the question being asked?

3.5 Rapid Evidence Assessments

Rapid evidence assessments (also sometimes described using similar terms such as rapid reviews, and rapid systematic reviews) are quick reviews of the evidence when resources are limited or the topic is urgent. Rapid evidence assessments may take a variety of forms, but typically follow the same processes as systematic reviews or systematic maps, with stages omitted or abbreviated (often the *level* of searching, and/or some of the quality checking stages). The abbreviated stages can vary between rapid reviews, since formal guidance for creation and quality appraisal of rapid evidence assessments is less established than for other evidence synthesis methods. Hamel et al. (2021) took definitions from 216 rapid reviews and 90 rapid review methods articles and found that no consensus existed in defining rapid reviews. There were also variations in the way in which each review was streamlined. Collins et al. (2015) outline guidance for creators of rapid reviews in the field of land and water management, but they separate this into processes that are aligned to systematic reviews (which they call rapid evidence assessments) and those that are more similar to systematic maps (which they call quick scoping reviews). The variability in rapid evidence assessment methods may make it difficult for users to assess the reliability and applicability of rapid reviews for their needs, but transparent reporting that outlines all stages of the methodology may help users evaluate individual rapid evidence assessments. The CEESAT checklist (3.1) may aid users in appraising rapid evidence assessments. The PROCEED open-access registry of titles and protocols for prospective evidence syntheses in the environmental sector <https://www.proceedevidence.info/> allows for registration of rapid review protocols. The required structure for prospective authors to follow is based on standard systematic review structures supported by the Collaboration for Environmental Evidence, and although no guidance is provided for streamlining processes, the use of standardised formats such as this may help authors with reporting.

3.6 Meta-Analyses

Meta-analysis is a set of statistical methods for combining the magnitude of outcomes (effect sizes) across different studies addressing the same research question. The methods of meta-analysis were originally developed in medicine and social sciences (Glass et al., 1981) and then introduced in ecology in the early 1990s (Arnqvist and Wooster, 1995). Fernandez-Duque and Valeggia (1994) brought meta-analysis to the attention of conservation biologists and outlined several advantages of meta-analysis over narrative reviews. In particular, meta-analysis allows better control of type II errors (e.g. assuming that a particular human action has no effect when it in fact does), which can have more serious consequences for conservation than making a type I error (e.g. assuming that an action has an effect when it really does not). As a small sample size is a common limitation in conservation studies, many individual primary studies have low statistical power and might fail to demonstrate the effect even when it exists. Fernandez-Duque and Valeggia (1994) have demonstrated that meta-analysis enables demonstration of the overall effect even when this effect was not apparent from the individual studies included in the meta-analysis, thus increasing the power of primary studies.

Nowadays, meta-analyses are often conducted as part of systematic reviews in conservation biology (as in step 7 – data synthesis, see Section 3.4). One of the main applications of meta-analysis in conservation biology is to assess the effectiveness of management interventions. This is achieved by combining weighted effect sizes, with more weight given to some studies than others, from individual primary studies to calculate the overall effect and its confidence interval (CI). The sign (positive/negative effect), the magnitude, and the significance of the overall effect can then be assessed. The results of meta-analyses often challenge the conventional wisdom about management effectiveness (Stewart, 2010; Côté et al., 2013). Meta-analysis also allows an assessment of the degree of heterogeneity in effect sizes, revealing the factors causing variation in effect among studies. Such distinction is important for establishing the conditions under which the management interventions are effective.

As an example, to assess whether tree retention at harvest helps to mitigate negative impacts on biodiversity, Fedrowitz et al. (2014) conducted a meta-analysis of 78 primary studies comparing species richness and abundance between retention cuts and either clearcuts or unharvested forests. They found that overall species richness was higher in retention cuts than in clearcuts and unharvested forests. However, effects varied between different species groups (Figure 3.3). Retention cuts supported higher species richness and abundance of forest species compared to clearcuts, but lower species richness and abundance of open-habitat species. In contrast, species richness and abundance of forest specialists were lower in retention cuts than in unharvested forest while species richness and abundance of open-habitat species were higher than in unharvested forests. Species richness and abundance of generalists did not differ between retention cuts and clearcuts. The results support the use of retention forestry since it moderates negative harvesting impacts on biodiversity. However, retention forestry cannot substitute conservation actions targeting certain highly specialised species associated with forest-interior or open-habitat conditions.

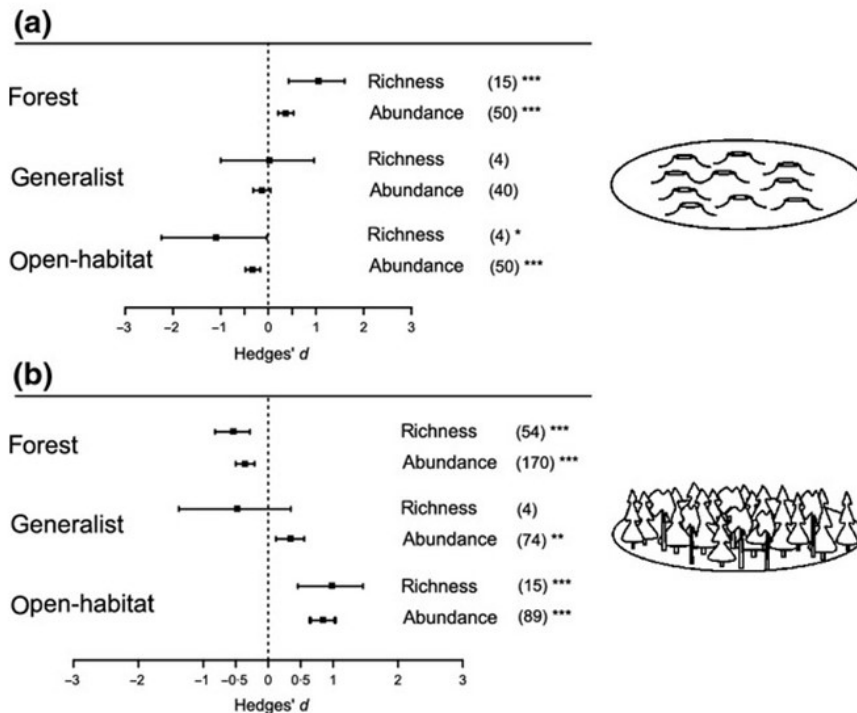


Figure 3.3 Effects of retention cuts (mean effect size 95% CI) on species richness and abundance of forest, generalist and open-habitat species when using (a) clearcut, or (b) unharvested forest as the control. Number of observations are stated in brackets. Effect size measure is a standardised mean difference (Hedges' d) between species richness and abundance in retention cuts and clearcut (a) or unharvested forest (b). Positive effects indicate higher species richness and abundance on retention cuts whereas negative effects indicate lower species richness and abundance on retention cuts. Effects are not significantly different from 0 when 95% CIs include 0. For significant effects, P-values are shown as * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (Source: Fedrowitz et al., 2014, CC-BY-NC-3.0 <https://creativecommons.org/licenses/by-nc/3.0/>)

While the popularity of meta-analyses in ecology and conservation biology has grown considerably over the last two decades, repeated concerns have been raised about the quality of the published meta-analyses in these fields (Gates, 2002; Vetter et al., 2013; Koricheva and Gurevitch, 2014; O'Dea et al., 2021). Of particular relevance to policy and decision making in conservation, it has been argued that rapid temporal changes in magnitude, statistical significance and even the sign of the effect sizes reported in many ecological meta-analyses represent a real threat to policy making in conservation and environmental management (Koricheva and Kulinskaya, 2019). To improve the quality of meta-analyses in ecology and evolutionary biology, Koricheva and Gurevitch (2014) developed a checklist of quality criteria for meta-analysis (Checklist 3.1). A more extensive checklist covering both systematic review and meta-analysis criteria has been recently developed by O'Dea et al. (2021) as an extension of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) to ecology and evolutionary biology (PRISMA-EcoEvo).

3.2 Checklist of quality criteria for meta-analyses

- Has formal meta-analysis been conducted (i.e. combination of effect sizes using standard meta-analysis methodology) or is it simply a vote count (comparison of % of statistically significant and nonsignificant findings)?
- Have standard metrics of effect size been used (e.g. standardised mean difference, correlation coefficient, response ratio, etc.) or, if non-standard metrics have been employed, is the distribution of these parameters known and have the authors explained how they calculated variances for such metrics?
- If more than one estimate of effect size per study was included in the analysis, has the potential non-independence of these estimates been taken into account?
- Have effect sizes been weighted by study precision or has the rationale for using an unweighted approach been provided?
- Has the statistical model for meta-analysis and the software used been described?
- Has heterogeneity of effect sizes between studies been quantified?
- Have the causes of existent heterogeneity in effect sizes been explored by meta-regression?
- If the effects of multiple moderators have been tested, have potential non-independence of and interactions between moderators been taken into account?
- If meta-analysis combines studies conducted on different species, has the phylogenetic relatedness of species been taken into account?
- Have tests of publication bias been conducted?
- If any meta-analysis combined studies published over a considerable time span then have possible temporal changes in effect size been tested?
- Has sensitivity analysis been performed to test the robustness of the results?
- Has the dataset used for meta-analysis, including effect sizes and variances/sample sizes from individual primary studies and moderator variables, been provided in an electronic appendix?

(Source: Koricheva and Gurevitch, 2014)

This checklist can be downloaded from <https://doi.org/10.11647/OBP.0321#resources>. It can be modified and tailored for specific uses.

3.7 Open Access Effect Sizes

Metadataset (<https://www.metadataset.com/>), a database of effect sizes for actions, has been created, which, so far, contains over 15,000 effect sizes on invasive species control, cassava farming and aspects of Mediterranean agriculture, along with a suite of analytical tools (Shackelford et al., 2021). Metadataset enables interactive evidence synthesis, browsing publications by intervention, outcome, or country (using interactive evidence maps). It also allows filtering and weighing the evidence for specific options and then recalculating the results using only the relevant studies, a method known as ‘dynamic meta-analysis’ (Shackelford et al., 2021).

The database was created following a systematic review (Martin et al., 2020; see Section 3.4), systematic map (Shackelford et al., 2018; see Section 3.2) and subject-wide evidence synthesis (Shackelford et al., 2017; see Section 3.3) methodologies. Metadata was extracted from all studies that met each selection criteria. This included the mean values of treatments (e.g. plots with management interventions) and controls (e.g. plots without management interventions or which used alternative management interventions) and, if available, measures of variability around the mean (standard deviation, variance, standard error of the mean, or confidence intervals), number of replicates, and the P value of the comparison between treatments and controls.

Metadataset enables both narrative synthesis, in which the studies that met the inclusion criteria are described, and quantitative synthesis of the relevant results. The mean effects of each intervention on each outcome were calculated using standard meta-analytic methods (Borenstein et al., 2011; Martin et al., 2020). With the dynamic meta-analysis, users can filter the data to define a subset relevant to their situation and then the results for that subset are calculated using subgroup analysis and/or meta-regression (Shackelford et al., 2021; see Section 3.6 for further details). Figure 3.4 illustrates how Metadataset operates.

Individuals or organisations want to know whether an action will be effective for a specific option, for example for a particular species, but evidence summaries often provide general conclusions, for example for a group of species (Martin et al., 2022). Dynamic synthesis (Shackelford et al., 2021) was designed to help overcome this problem by enabling a process of creating an analysis tailored to the specific option, for example just considering studies within the user’s region.

Dynamic meta-analysis enables analysis only using the subset selected by the user (subgroup analysis) or by analysing all of the data but calculating different results for different subsets while accounting for the effects of other variables (meta-regression). Dynamic meta-analysis also includes ‘recalibration’, a method of weighting studies based on their relevance, allowing users to consider a wider range of evidence, not just data that is completely relevant (as none may exist; Shackelford et al., 2021). Some ‘critical appraisal’ (i.e. deciding which studies should be included in the meta-analysis, based on study quality) and ‘sensitivity analysis’ (i.e. permuting the assumptions of a meta-analysis, to test the robustness of the results) is also possible. For example, where evidence was limited, a user could decide to include further relevant, but lower quality, studies.

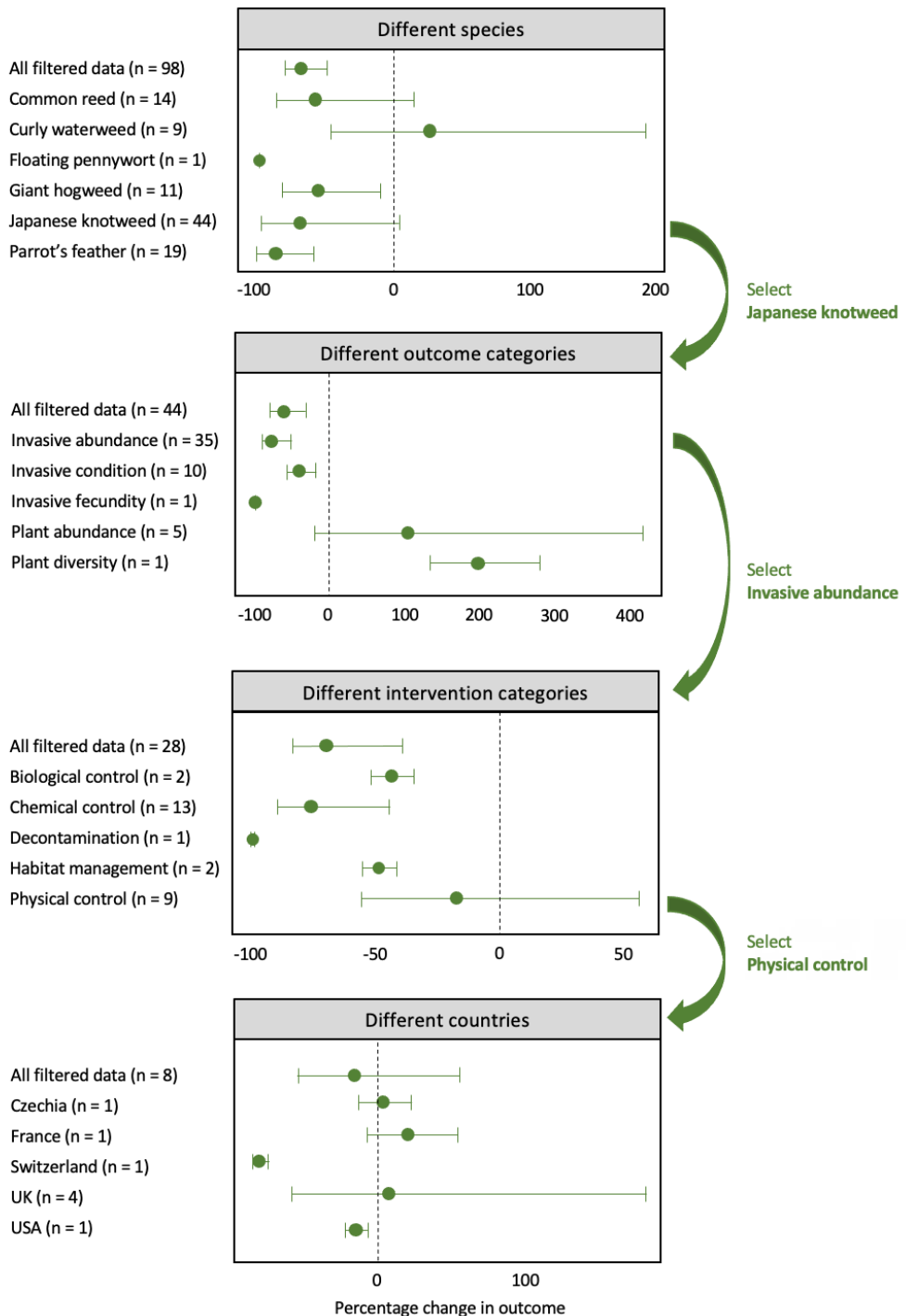


Figure 3.4 The process of adjusting analyses using Metadataset. From top shows the effect of all actions on all outcomes for all 17 invasive species and the six with the most data; Japanese knotweed is selected. Then shows the effect of all six most data rich outcome categories on Japanese knotweed; plant abundance is selected. Then shows the effect of all six actions on Japanese knotweed outcomes; physical control is selected. Finally shows the overall effect and the effect in five countries of physical control on Japanese knotweed. Number of effect sizes (n). (Source: authors)

3.8 Overviews of Reviews

Overviews of reviews (or reviews of reviews) use systematic methodologies to search for and identify systematic reviews within a topic area. The principles are similar to those used in traditional systematic reviews and maps, but instead of primary research, reviewers search for and identify existing systematic reviews, extracting, sometimes re-analysing, and reporting on their combined data.

In disciplines where multiple systematic reviews may be available within a topic area, overviews of reviews are recognised as useful methods of synthesis. For example, in the Cochrane handbook they are recommended for addressing research questions that are broader in scope than those examined in individual systematic reviews, for saving time and resources in areas where systematic reviews have already been conducted, and where it is important to understand any diversity present in the existing systematic review literature (Pollock et al., 2002). Overviews of reviews can also be used as part of multi-stage approaches to a question. For example, in the field of international development, Rebelo de Silva et al. (2017) carried out an overview of reviews to investigate the systematic review evidence base for the effects of interventions for smallholder farmers in Africa on various food security outcomes. This was used to inform a systematic map to assess the gaps and overlaps in a more focused sub-topic area, which was then used to inform a systematic review investigating the effects of specific interventions.

In the field of conservation, as the number of systematic reviews increases, overviews of reviews may become a useful tool for the future.

References

- Arnqvist, G. and Wooster, D. 1995. Meta-analysis: Synthesizing research findings in ecology and evolution. *Trends in Ecology & Evolution* 10: 236–40, [https://doi.org/10.1016/S0169-5347\(00\)89073-4](https://doi.org/10.1016/S0169-5347(00)89073-4).
- Borenstein, M., Hedges, L.V., Higgins, J.P., et al. 2011. *Introduction to Meta-Analysis* (Chichester: John Wiley and Sons, Ltd), <https://doi.org/10.1002/9780470743386>.
- Cheng, S.H., McKinnon, M.C., Masuda, Y.J., et al. 2020. Strengthen causal models for better conservation outcomes for human well-being. *PLoS ONE* 15: e0230495, <https://doi.org/10.1371/journal.pone.0230495>.
- Collaboration for Environmental Evidence. 2018. *Guidelines and Standards for Evidence synthesis in Environmental Management. Version 5.0*, ed. by A.S. Pullin, et al., <https://environmentalevidence.org/information-for-authors/>.
- Collins, A., Coughlin, D., Miller, J., et al. 2015. *The Production of Quick Scoping Reviews and Rapid Evidence Assessments: A How to Guide* (UK Defra; Joint Water Evidence Group), <https://nora.nerc.ac.uk/id/eprint/512448/1/N512448CR.pdf>.
- Côté, I.M., Stewart, G.B. and Koricheva, J. 2013. Contributions of meta-analysis to conservation and management. In: *Handbook of Meta-Analysis in Ecology and Evolution*, ed. by J. Koricheva, et al. (Princeton: Princeton University Press), pp. 420–25, <https://doi.org/10.23943/princeton/9780691137285.003.0026>.

- Fedrowitz, K., Koricheva, J., Baker, S.C., et al. 2014. Can retention forestry help conserve biodiversity? A meta-analysis. *Journal of Applied Ecology* 51: 1669–79, <https://doi.org/10.1111/1365-2664.12289>.
- Fernandez-Duque, E. and Valeggia, C. 1994. Meta-analysis: A valuable tool in conservation. *Conservation Biology* 8: 555–61, <https://doi.org/10.1046/j.1523-1739.1994.08020555.x>.
- Flemming, K., Booth, A., Garside, R., et al. 2019. Qualitative evidence synthesis for complex interventions and guideline development: Clarification of the purpose, designs and relevant methods. *BMJ Global Health* 4: e000882, <https://doi.org/10.1136/bmjgh-2018-000882>.
- Gates, S. 2002. Review of methodology of quantitative reviews using meta-analysis in ecology. *Journal of Animal Ecology* 71: 547–57, <https://doi.org/10.1046/j.1365-2656.2002.00634.x>
- Glass, G.V., McGaw, B. and Smith, M.L. 1981. *Meta-Analysis in Social Research* (Beverly Hills, CA: SAGE Publications).
- Haddaway, N.R., Feierman, A., Grainger, M.J., et al. 2019. *EviAtlas*: A tool for visualising evidence synthesis databases. *Environmental Evidence* 8: Article 22, <https://doi.org/10.1186/s13750-019-0167-1>.
- Hamel, C., Michaud, A., Thuku, M., et al. 2021. Defining rapid reviews: A systematic scoping review and thematic analysis of definitions and defining characteristics of rapid reviews. *Journal of Clinical Epidemiology* 129: 74–85, <https://doi.org/10.1016/j.jclinepi.2020.09.041>.
- James, K.L., Randall, N.P. and Haddaway, N.R. 2016. A methodology for systematic mapping in environmental sciences. *Environmental Evidence* 5: Article 7, <https://doi.org/10.1186/s13750-016-0059-6>.
- Koricheva, J. and Gurevitch, J. 2014. Uses and misuses of meta-analysis in plant ecology. *Journal of Ecology* 102: 828–44, <https://doi.org/10.1111/1365-2745.12224>.
- Koricheva, J. and Kulinskaya, E. 2019. Temporal instability of evidence base: A threat to policy making? *Trends in Ecology & Evolution* 34: 895–902, <https://doi.org/10.1016/j.tree.2019.05.006>.
- Martin, P.A., Shackelford, G.E., Bullock, J.M., et al. 2020. Management of UK priority invasive alien plants: A systematic review protocol. *Environmental Evidence* 9: Article 1, <https://doi.org/10.1186/s13750-020-0186-y>.
- Martin, P.A., Christie, A.P., Shackelford, G.E., et al. 2022. Flexible synthesis to deliver relevant evidence for decision-making. Preprint, <https://doi.org/10.31219/osf.io/3t9jy>.
- McKinnon, M., Cheng, S., Garside, R., et al. 2015. Sustainability: Map the evidence. *Nature* 528: 185–87, <https://doi.org/10.1038/528185a>.
- O’Dea, R.E., Lagisz, M., Jennions, M.D., et al. 2021. Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: A PRISMA extension. *Biological Reviews* 96: 1695–1722, <https://doi.org/10.1111/brv.12721>.
- Pollock, M., Fernandes, R.M., Becker, L.A., et al. 2022. Chapter V: Overviews of Reviews. In: *Cochrane Handbook for Systematic Reviews of Interventions, version 6.3* (updated February 2022), ed. by J.P.T. Higgins, et al. (Cochrane, 2022), www.training.cochrane.org/handbook.
- Pullin, A., Cheng, S., Cooke, S., et al. 2020. Informing conservation decisions through evidence synthesis and communication. In: *Conservation Research, Policy and Practice*, Ecological Reviews, ed. by W.J. Sutherland, et al. (Cambridge: Cambridge University Press), pp. 114–28, <https://doi.org/10.1017/9781108638210.007>.

- Rebello Da Silva, N., Zaranyika, H., Langer, L., et al. 2017. Making the most of what we already know: A three-stage approach to systematic reviewing. *Evaluation Review* 41: 155–72, <https://doi.org/10.1177/0193841X16666363>.
- Shackelford, G.E., Haddaway, M.R., Usieta, H.O., et al. 2018. Cassava farming practices and their agricultural and environmental outcomes: A systematic map protocol. *Environmental Evidence* 7: Article 30, <https://doi.org/10.1186/s13750-018-0142-2>.
- Shackelford, G.E., Kelsey, R., Robertson, R.J., et al. 2017. *Sustainable Agriculture in California and Mediterranean Climates: Evidence for the Effects of Selected Interventions*, Synopses of Conservation Evidence Series (Cambridge: University of Cambridge), <https://www.conservationevidence.com/synopsis/pdf/15>.
- Shackelford, G.E., Martin, P.A., Hood, A.S.C., et al. 2021. Dynamic meta-analysis: A method of using global evidence for local decision making. *BMC Biology* 19: Article 33, <https://doi.org/10.1186/s12915-021-00974-w>.
- Stewart, G. 2010. Meta-analysis in applied ecology. *Biology Letters* 6: 78–81, <https://doi.org/10.1098/rsbl.2009.0546>.
- Sutherland, W.J., Dicks, L.V., Petrovan, S.O., et al. 2021. *What Works in Conservation 2021* (Cambridge: Open Book Publishers), <https://doi.org/10.11647/OBP.0267>.
- Sutherland, W.J., Taylor, N.G., MacFarlane, D., et al. 2019. Building a tool to overcome barriers in the research-implementation space: The Conservation Evidence database. *Biological Conservation* 238: 108199, <https://doi.org/10.1016/j.biocon.2019.108199>.
- Vetter, D., Rucker, G. and Storch, I. 2013. Meta-analysis: A need for well-defined usage in ecology and conservation biology. *Ecosphere* 4: Article 74, <https://doi.org/10.1890/ES13-00062.1>.
- Westgate, M.J. and Lindenmayer, D.B. 2017. The difficulties of systematic reviews. *Conservation Biology* 31: 1002–07, <https://doi.org/10.1111/cobi.12890>.
- Woodcock, P., Pullin, A.S. and Kaiser, M.J. 2014. Evaluating and improving the reliability of evidence syntheses in conservation and environmental science: A methodology. *Biological Conservation* 176: 54–62, <https://doi.org/10.1016/j.biocon.2014.04.020>.