EU-NESA
Network for Evidence Synthesis in the Agri-food Sector

Introduction to evidence synthesis

Training materials produced by:
Nicola P Randall
nrandall@harper-adams.ac.uk
and
Rafael Jose Vieira
rafaelvieira@med.up.pt

as part of the EU COST funded Network for Evidence Synthesis in the Agri-Food Sector (COST Action No: CA23107)

COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career and innovation

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

Funded by the European Union

Introduction

Nicola Randall
nrandall@harper-adams.ac.uk

Evidence – how can we make good use of it?

- Evidence: "The available body of facts or information indicating whether a belief or proposition is true or valid." Oxford English Dictionary
- The relevance of different types of evidence depends on the question being asked and context in which it is asked.

We often use reviews for collating (research) evidence

- Summarise a large topic in one place
- Provide comprehensive catalogue of interventions, outcomes, methods and theories
- Numerical summary of results from multiple studies (effect sizes with confidence intervals)
- Improve understanding of causality, confounders and effect modifiers
- Provide an 'independent' assessment of the body of evidence

What is an evidence synthesis and what makes it different to a traditional literature review?

- An evidence synthesis follows a structured a-priori methodology that is designed to gather existing research evidence in order to answer a question.
- The methods are designed to reduce bias
- Every choice you make is justified
- Every study that you include (or exclude) is subject to the same criteria for inclusion/data extraction
- Your methods should be able to be followed by others that want to replicate or update your review.

IN OTHER WORDS IT IS DESIGNED AND FOLLOWED LIKE A FIELD OR LAB EXPERIMENT

Traditional Reviews- subject to bias/limitations

- Selection bias
- Lack of comprehensiveness
- Publication bias
- No transparency
- Vote-counting/quality bias
- Discussion bias

Systematic Reviews- minimizing bias and maximizing comprehensiveness

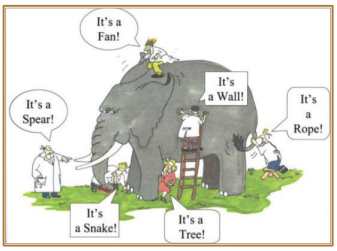
- Exhaustive searching
- Comprehensiveness
- Grey literature
- Transparent methods
- Weight studies
- Synthesis of all relevant studies

Haddaway, N. R., Woodcock, P., Macura, B., and Collins, A. (2015) Making literature reviews more reliable through application of lessons from systematic reviews. Conservation Biology, DOI: 10.1111/cobi.12541.

What is an evidence synthesis and what makes it different to a traditional literature review?

- Literature review
  - Review (and analysis) of current knowledge
- Systematic review
  - Identification, appraisal and synthesis of current research evidence
- Systematic map
  - Searchable database of the state of research
- Meta-analysis
  - Statistically combines evidence from multiple studies (often carried out as part of a systematic review
  - Versions of all the above – eg. Rapid evidence assessments, quick scoping reviews, evidence gap maps

Why do we need reliable evidence?



Lets suppose that we want to know what impact adding fertiliser has to wheat growth…

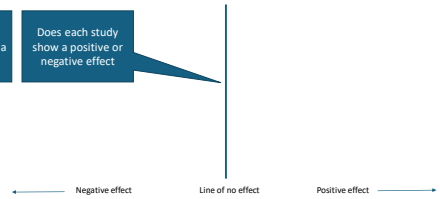
Rather than carrying out new research we could look at existing research

Lets suppose that we want to know what impact adding fertiliser has to wheat growth…

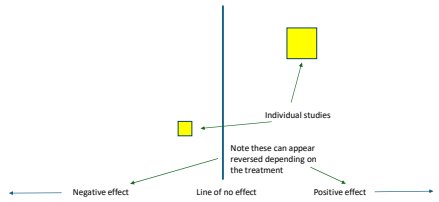
Rather than carrying out new research we could look at existing research

We could plot the studies findings on a graph e.g.

Does each study show a positive or negative effect

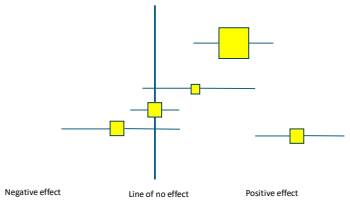


The size of a study relates to how much effect it has overall. Different studies might have different findings

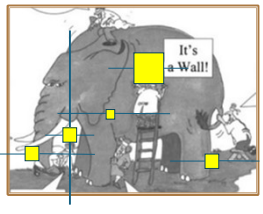


11

We could add confidence intervals to each study and plot lots of studies on our graph if they all ask the same question.

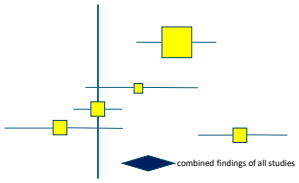


12

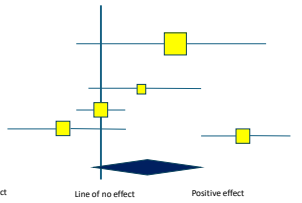


13

And then we could combine data from studies to provide an overall estimate of effect of fertiliser on wheat growth. This is called a forest plot, & is used in meta-analysis



The weight of different studies will influence whether or not we find an overall effect



Why bother?

- 1959 -1st randomised control of thrombolytic drug to treat acute myocardial infarction
- 1973 –the date the cumulative meta-analysis shows a statistically significant reduction of overall mortality of approx. 25% (after eight studies that included 2,432 patients)
- 1988 – the date that thrombolytic drug approved by the Food and Drug Administration (FDA) (only following 62 more placebo-controlled trials involving an additional 45,000 patients).

Original cumulative meta-analysis: Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992) A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 268: 240–248.
Also see: Clarke, M., Brice, A. and Chalmers, I., 2014. Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. PloS one, 9(7), p.e102670.

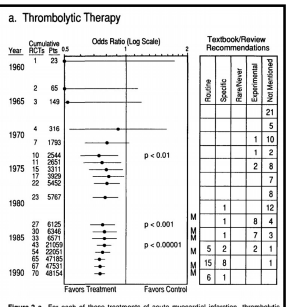


Figure 2 a. For each of these treatments of acute myocardial infarction, thrombolytic therapy (a), prophylactic lidocaine (b), and intravenous magnesium (c), the cumulative meta-analyses by year of publication of RCTs are presented on the left. The cumulative number of trials and patients (Pts) are also presented. On the right, the recommendations of the clinical expert reviewers are presented in 2-year segments, except for the entry in 1986, which represents all previous years. The letter M indicates that at least one meta-analysis was published that year; NS indicates not significant. Reprinted with permis-

Why bother?

- In other words, if a cumulative meta-analyses of existing research had been carried out whenever a new research study was carried out, the drug could have been approved 15 years earlier and saved carrying out a further 62 trials.

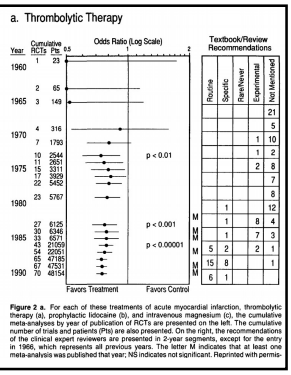
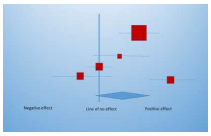


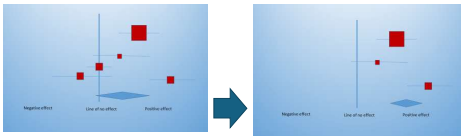
Figure 2 a. For each of these treatments of acute myocardial infarction, thrombolytic therapy (a), prophylactic lidocaine (b), and intravenous magnesium (c), the cumulative meta-analyses by year of publication of RCTs are presented on the left. The cumulative number of trials and patients (Pts) are also presented. On the right, the recommendations of the clinical expert reviewers are presented in 2-year segments, except for the entry in 1986, which represents all previous years. The letter M indicates that at least one meta-analysis was published that year; NS indicates not significant. Reprinted with permis-

Original cumulative meta-analysis: Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC (1992) A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 268: 240–248.
Also see: Clarke, M., Brice, A. and Chalmers, I., 2014. Accumulating research: a systematic account of how cumulative meta-analyses would have provided knowledge, improved health, reduced harm and saved resources. PloS one, 9(7), p.e102670.

That's all very well but what if I choose only the studies that support my point of view?

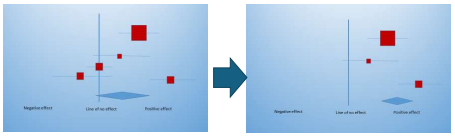


That's all very well but what if I choose only the studies that support my point of view?



That's all very well but what if I choose only the studies that support my point of view?

- How would anybody (including myself) even know?
- *Confirmation bias – tendency to search for and interpret information consistent with one's prior beliefs.*
- *Even published meta-analyses can be more or less susceptible to bias*



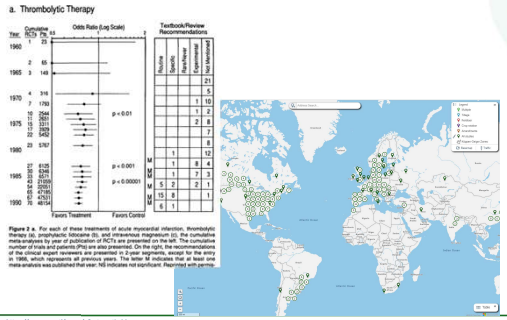
Systematic maps and reviews aim to reduce bias and increase transparency

- **Systematic**: 'done or acting according to a fixed plan or system; methodical'
- **Review**: 'a critical appraisal of a book, play or other work'
- **Map**: 'a diagram or collection of data showing the spatial arrangement or distribution of something over an area'

(Oxford English Dictionary)

Systematic Review or Systematic Map?

- A systematic review usually aims to answer a specific question
  - In order to answer the question, it must be focused
  - Eg. what impact does the application of nitrate have on the growth rate of wheat?
  - The output is usually a statistical or narrative analysis that helps to answer the question in a report.
- A systematic map aims to investigate the state of research
  - It can be used to address broader topic areas, or areas where the evidence is likely to be diverse
  - It may also aim to identify where research gaps are
  - E.g. what interventions aim to improve farmer income in water scarce areas?
  - The output is usually a searchable database of research together with visualisations (heatmaps/geographical maps etc) of the research



Systematic Review process (using the Collaboration for Environmental Evidence process

- Question formulation
- Protocol development
- Searching
- Article Screening
- Data Extraction
- Critical Appraisal
- Synthesis (Meta-analysis)
- Review Findings
- Dissemination (Publication etc)

https://environmentalevidence.org/

Research Question, Protocol Development and Stakeholder Engagement

Rafael José Vieira
rafaelvieira@med.up.pt

Steps of a systematic review

1. **Define the research question**
2. **Plan the eligibility criteria**
3. Search for the primary studies
4. Select the studies
5. Collect the data
6. Assess the methodological quality of the studies
7. Analyse and present the results
8. Assess the certainty of the evidence
9. Interpret the results and draw conclusions

Research Question

Rafael José Vieira
rafaelvieira@med.up.pt

Not all questions are suitable

- Evidence needs are often presented as a broad “problem”
- Initial questions may need breaking down because they:
  - are too broad in scope to answer within a single SR
  - are open-framed
- SMs can generally be based on broader question than SRs.

Stakeholder involvement in question formulation

Start with a broad problem

↓

Explore a number of potential questions

↓

Assess the suitability of questions for synthesis

↓

Prioritisation of these questions

Broad problem

- What is the impact of neonicotinoid pesticides on the environment?
  - Q1 - What is the effect of neonicotinoid pesticides on pollinator performance?
  - Q2 - What is the prevalence of neonicotinoid pesticides in fresh water?
  - Q3 - What is the comparative sensitivity of the tests for presence and concentration of neonicotinoid pesticides?
  - Q4 - What type and severity of impacts do neonicotinoid pesticides have on target pest species?
  - Q5 - What other options are available for pesticides as substitutes for neonicotinoids?

Research question: PICO/PECO

- **Population:** To whom/what does it apply?
- **Intervention / Exposure:** Strategy or exposure of interest
- **Comparator:** Alternative to the intervention/exposure (may be just the absence of intervention/exposure).
- **Outcome**

- Sometimes, also:
  - **Time** (PICOT)
  - **Study design** (PICOS)

Research question: PICO/PECO

P	I	C	O
Population	Intervention/ Exposure	Comparator	Outcome

What is the effect of neonicotinoid pesticides on pollinator performance?

32

Research question: PICO/PECO

P	I	C	O
Pollinator	Neonicotinoid pesticides	Other pesticides/No pesticides	Performance

What is the effect of neonicotinoid pesticides on pollinator performance?

33

Research question: PICO/PECO

P	I	C	O
Population	Intervention/ Exposure	Comparator	Outcome

What is the prevalence of neonicotinoid pesticides in fresh water?

34

Research question: PICO/PECO



What is the prevalence of neonicotinoid pesticides in fresh water?

Purpose

- A **systematic review** usually aims to answer a specific question
  - In order to answer the question, it must be focused
    - Eg. what impact does the application of nitrate have on the growth rate of wheat?
  - The output is usually a statistical or narrative analysis that helps to answer the question in a report.
- A **systematic map** aims to investigate the state of research
  - It can be used to address broader topic areas, or areas where the evidence is likely to be diverse
  - It may also aim to identify where research gaps are
    - E.g. what interventions aim to improve farmer income in water scarce areas?
  - The output is usually a searchable database of research together with visualisations (heatmaps/geographical maps etc) of the research

Purpose

- A **systematic review** usually aims to answer a specific question
  - Often a statistical analysis (meta-analysis)
  - There are other analyses (eg narrative) but you should use the most appropriate analysis
- A **systematic map** aims to investigate the state of research
  - Broader topic areas, research gaps, searchable database and visualisations
  - Apps such as EviAtlas can be used for visualisations

Exercise 1a. Question setting

- This is important as it may influence how you design the rest of your review eg. Where you search, what search terms you use, what data you extract.
- E.g. What are the health impacts of copper?
  - PICO elements?
  - Assumptions?

Exercise 1a. Question setting

- A better question might be: What are the human health impacts of ingesting copper in the diet?
  - Population: Humans
  - Intervention/exposure: Copper in the diet (this could be deliberate interventions such a copper supplements or could be inadvertent exposure such as in toxicology studies)
  - Comparator: No copper exposure/different levels of copper exposure
  - Outcome: health outcomes (any specified in studies)

Exercise 1b. Your own question



Eligibility criteria

- Rules determining which studies are included in the review
- Based on some or all PICO components (± definition of eligible study designs)
- Any changes to the eligibility criteria after the publication of the protocol must be properly justified in the review.

Population

- Clear definition to identify the subjects of interest
- Two aspects to consider:
  - Condition / characteristic of interest
    - e.g. crop type (wheat, maize, rice), animal breed (dairy cattle vs. beef cattle), or food product (organic vs. conventional).
    - How is it measured or defined (e.g. by laboratory test, farmer record, certification body)?
  - Population and context (setting)
    - e.g. smallholder farms, industrial farms, specific regions, supply-chain stage (production, processing, retail).
    - Any restrictions should have a clear justification.
- Alternative: include all and explore differences through subgroup analyses (e.g. crop type, farming system, geographical region).

Intervention/Exposure

- Describe the intervention/exposure in as much detail as possible:
  - Formulation / type (e.g. fertiliser mix, pesticide formulation, feed composition)
  - Dose / intensity (e.g. kg/ha, concentration, ration amount)
  - Mode of application / delivery (e.g. spraying, irrigation, feed supplement, vaccination)
  - Schedule – start, frequency, duration (e.g. seasonal timing, daily feeding, weekly spraying)
  - Equipment used (e.g. tractors, sprayers, milking machines, storage systems)
  - Human resources – qualifications, training (e.g. skilled labour, agronomist involvement, veterinary supervision)
  - Location and context (e.g. open field, greenhouse, free-range, intensive farming system)
  - Standalone or combined with other interventions (e.g. fertiliser + irrigation, vaccination + biosecurity measures)

Comparator

- Should be defined based on the objectives of the review.
- Active comparators must be described in as much detail as the intervention (e.g. type of fertiliser, pesticide, feed, or management practice used as the alternative).
- Clearly specify what "no intervention" means in the agrifood context:
  - e.g. no fertiliser application, untreated plots, no feed supplement, standard husbandry practices, business-as-usual farming.
- Additional comparators identified during the review can be included, provided they are consistent with the predefined criteria.

Outcomes

- Excluding studies based on reported outcomes can introduce bias:
  - Outcomes may have been selectively reported.
  - Additional information may be available from other sources.
- Restricting by outcome may be appropriate if outcomes are central to the review question.
  - Example: prevention vs. treatment, or interventions applied to more than one crop, animal species, or production stage.
- The protocol should clearly state the decision and justification.

Protocol

EU-NESA Training School

Porto, 28/29 September 2025

Rafael José Vieira
rafaelvieira@med.up.pt

Protocol

- A priori guide to the evidence synthesis with rationale, hypotheses and planned methods
- Protocol should be as detailed and rigorous as possible and include review scoping

Protocol

- Why do you think deciding on a method before you start your review is a good idea?

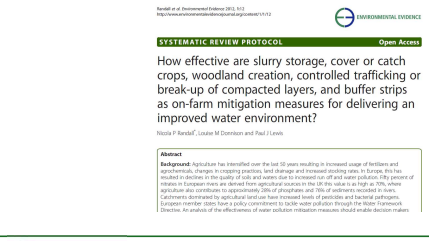
Protocol

- Abstract
- Background
- Objectives
- Methods
  - Search Strategy (search string, databases, search engines, specialist sources, search comprehensiveness assessment)
  - Inclusion Criteria (key elements, consistency checking) and screening methodology
  - Data Extraction
  - Critical Appraisal
  - Synthesis (quantitative/qualitative)
  - Stakeholder engagement
- Sources of Support/Conflicts of Interest
- Acknowledgements
- References

Where to publish

- The protocol may be published through several means:
  - **Scientific journal:** Some journals accept the publication of research protocols, including systematic reviews. In this case, the protocol will be peer-reviewed.
  - **PROSPERO:** The international register of systematic reviews. Although it does not conduct peer review, it is often considered a reference repository for systematic reviews.
  - **Open platforms (e.g. OSF – Open Science Framework):** Allow protocols to be shared quickly, openly, and free of charge, as well as updated over time. These platforms are useful for systematic reviews and also for other types of research, promoting open science.

Where to publish



Randall et al. Environmental Evidence 2013, 2:2
http://www.environmentalevidencejournal.org/content/2/1/2

ENVIRONMENTAL EVIDENCE

SYSTEMATIC REVIEW PROTOCOL Open Access

How effective are slurry storage, cover or catch crops, woodland creation, controlled trafficking or break-up of compacted layers, and buffer strips as on-farm mitigation measures for delivering an improved water environment?

Nicola P Randall*, Louise M Donnison and Paul J Lewis

Abstract

Background: Agriculture has intensified over the last 50 years resulting in increased usage of fertilisers and agrochemicals, changes in cropping practices, land drainage and increased stocking rates. In Europe, this has resulted in declines in the quality of soils and waters due to increased run off and water pollution. Fifty percent of nitrates in European rivers are derived from agricultural sources in the UK this value is as high as 70%, where agriculture also contributes to approximately 28% of phosphates and 75% of sediments recorded in rivers. Catchments dominated by agricultural land use have increased levels of pesticides and bacterial pathogens. European member states have a policy commitment to tackle water pollution through the Water Framework Directive. An analysis of the effectiveness of water pollution mitigation measures should enable decision makers

Stakeholder Engagement

EU-NESA Training School

Porto, 28/29 September 2025

Rafael José Vieira
rafaelvieira@med.up.pt

Key groups involved in evidence synthesis

- **Review Team:** the group that conducts the review; the authors of the review report
- **Commissioners:** policy or practice groups that identify the need for evidence cover the costs of the work
- **Stakeholder Group:** all individuals and organisations that might have a stake in the outcome of the review and may use the findings in decision-making
- **Review or Advisory Group:** a subgroup that may act as expert advisors to the review team but not actually be part of the core team

Stakeholder engagement

- Evidence syntheses often have wide stakeholder communities
- Stakeholders should be engaged through open consultation from early stages to maximise impact
- SH community includes end users and experts
- Open consultation should occur again once review completed

Stakeholder engagement



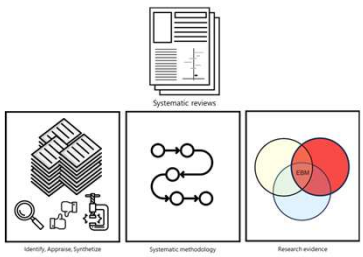
Searching and scoping

Rafael José Vieira
rafaelvieira@med.up.pt

Steps of a systematic review

1. Define the research question
2. Plan the eligibility criteria
3. **Search for the primary studies**
4. Select the studies
5. Collect the data
6. Assess the methodological quality of the studies
7. Analyse and present the results
8. Assess the certainty of the evidence
9. Interpret the results and draw conclusions

Introduction



Introduction

Identification of the maximum number of eligible studies from multiple sources

- ≠ Traditional narrative reviews
- ↓ Biases
- Better estimates of effect and uncertainty

Where to search

Rafael José Vieira
rafaelvieira@med.up.pt

Where to search

- Electronic databases and search engines
  - Bibliographic databases
  - Ongoing studies and unpublished data
- Additional search sources
  - "Hand" searching, experts, references, and citations

Bibliographic databases

- Web of Science
- Medline/PubMed
- Scopus
- Agricola
- AGRIS (FAO)
- Academic Search Premier
- Biological Abstracts
- CAB Abstracts

62

Bibliographic databases

- How to access?
  - Subscription based
    - e.g. through your institution
  - Free

63

Bibliographic databases

- Web of Science (Subscription)
- Medline/PubMed (Free)
- Scopus (Subscription)
- Agricola (Free)
- AGRIS (FAO) (Free)
- Academic Search Premier (Subscription)
- Biological Abstracts (Subscription)
- CAB Abstracts (Subscription)

64

Bibliographic databases

- It is desirable that searches in **"general" databases** are complemented with searches in "**specific" databases**.
- Typically: search in two "general" databases and one "specific" database.

Grey literature

- Theses and dissertations
  - Open Access Theses and Dissertations (OATD)
  - Open Repositories of Universities
- Conference and congress abstracts
  - Google Scholar, WoS (if the congress journal is indexed)
- Technical reports, newspaper articles, …
- Organisational websites

Additional sources

- Experts
  - (in information science, or in the research field)
- References and citations
  - Forward and backward citation searching
- Searching in "general" search engines

How to search

Rafael José Vieira
rafaelvieira@med.up.pt

How to search

- Structure a search string
  - Concepts, filters, previous queries, and missing terms
- Boolean operators, truncation, descriptors, and MeSH
- Other operators

Aspects to consider

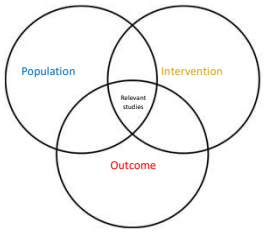
- We intend to include studies rather than study reports.
  - However, it is more convenient to search for articles/publications.
- The aim of the search is to maximise sensitivity.
  - Trade-off with specificity.

Structuring a search query

- PICO + study design + eligibility criteria
  - Some concepts may not be well described in the title or abstract, or may not be well indexed.
  - Therefore, the inclusion of a comparator or outcome, for example, should be assessed on a case-by-case basis.

Structuring a search query

- E.g. How do draining and re-wetting affect carbon stores and greenhouse gas fluxes in peatland soils?
  - Population: Peatland soils
  - Interventions: Draining and re-wetting
  - Outcome: Change in carbon stores and GHG fluxes



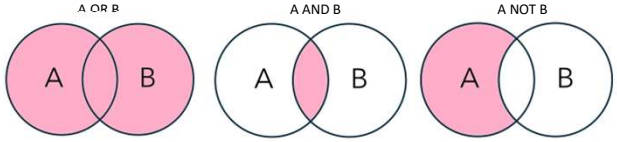
Designing a search query

- How do draining and re-wetting affect carbon stores and greenhouse gas fluxes in peatland soils?

Population (P)		Intervention (I)	Outcome (O)	
Peat*	Tourbe	Flood*	Carbon	CH4
Bog*	Tourbièr*	Drain*	Greenhouse gas*	CO2
Muskeg	Suo	Restor*	Green-house gas*	N2O
Pocosin*	Fen	Grip block*	GHG*	Nitrous oxide
Quag*	Torfmoor	Rewet*	Methane	Dissolved organic carbon
Mire*	Muck	Re-wet*	DOC	Dissolved organic matter
Slough	Swamp	Plough*	DOM	Soil organic matter
Aapa*	Hochmoortorf	Ditch*	SOM	Organic matter
Turvesuo	Sedge	Drought	Organic content	

Boolean operators

#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10	animals [mh] NOT humans [mh]
#11	#9 NOT #10



74

Designing a search query

Combine search terms:
(population terms) AND (intervention terms) AND (outcome terms)

(Peat* OR Bog*OR Muskeg OR Pocosin*OR Quag*OR Mire OR Slough OR Aapa*OR Turvesuo OR Tourbe OR Tourbièr* OR Suo OR Fen OR Torfmoor OR Niedermoortorf OR Hochmoortorf OR Palsa OR Swamp OR Carr OR Mor OR Sedge OR Muck)

AND

(Flood* OR Drain* OR Restor* OR Grip block* OR Rewet* OR "Re-wet*" OR Plough* OR Ditch* OR Drought OR paludiculture)

AND

(Carbon OR "Greenhouse gas*" OR "Green-house gas*" OR "GHG*" OR Methane OR "Organic matter" OR "Organic content" OR CO2 OR CH4 OR N2O OR Nitrous Oxide OR DOM OR DOC OR SOM)

75

Truncation

Syst*	System
	Systemic
	Systematic
	SYST2 *[Bacillus subtilis strain]*
Syst*2	System
Syst*6	Systematic

76

Search operators

- ADJ
  - *patient adj3 education*
    - Patient education
    - Education for patients
    - Education for cancer patients
    - Patient-focused health education
- *Wildcard*
  - *wom*n*
    - Woman
    - Women

Structuring a search query

- Filters and previous *queries*
  - Search strategies targeted at identifying specific records
  - InterTASC (ISSG)
  - Previous systematic reviews

Scoping the literature

- You need to check whether your work is viable
- This would usually come before/alongside your search strategy design
- Test some basic search terms on one or two databases

Managing your references

- Use a reference manager that:
  - Groups references into libraries
  - Identifies type of reference (e.g. paper, conference proceedings, book etc.)
  - Can be added manually or imported from other databases/catalogues
  - Enables multiple citations
  - Provides further details for each reference (E.g. abstracts)

Study selection

Rafael José Vieira
rafaelvieira@med.up.pt

Steps of a systematic review

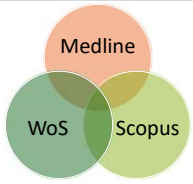
1. Define the research question
2. Plan the eligibility criteria
3. Search for the primary studies
4. **Select the studies**
5. Collect the data
6. Assess the methodological quality of the studies
7. Analyse and present the results
8. Assess the certainty of the evidence
9. Interpret the results and draw conclusions

Selection process

Document each step with detail

Selection process

Combine search results + Remove duplicates

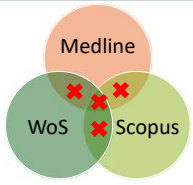


EN

rayyan

Selection process

Combine search results + Remove duplicates

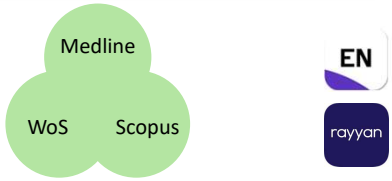


EN

rayyan

Selection process

Combine search results + Remove duplicates



EN

rayyan

Eligibility criteria

- Avoid "**selection bias**" - be transparent, repeatable, objective
- Defined set of inclusion criteria (*a priori* in protocol)
- Based on population, intervention, comparator and outcome
- Decisions to be made include
  - Optional use of comparator element of question? (title & abstract)
  - Should all criteria be filled or just some? (title & abstract)
  - Include a criteria for appropriate study design to the question
  - Justified limitations? e.g date range

Eligibility criteria

- Are marine protected areas effective tools for sustainable fisheries management? (Biodiversity impact of marine reserves in temperate zones)

P	I	C	O
All marine biota	Temperate marine protected areas (geographically defined areas subject to fishing restrictions/no take zones)	Less protected or unprotected area	Density (or other abundance measure), biomass, species richness

Selection process

Combine search results + Remove duplicates

Title and abstract screening



Title and abstract screening

- Assess for relevance and eligibility at title, abstract and full-text (include/exclude)
- If uncertain retain for next level of screening.

Title and abstract screening

- Assess for relevance and eligibility at title, abstract and full-text (include/exclude)
- If uncertain retain for next level of screening.
- Easier if criteria are well-defined.
- Trial and include all who will do this task.
- Assess level of agreement between screeners.
- If disagreement refine definitions (and repeat assessment).
- Record process.

Selection process

- Combine search results + Remove duplicates
- Title and abstract screening
- Full text retrieving

Selection process

- Combine search results + Remove duplicates
- Title and abstract screening
- Full text retrieving
- Combine multiple reports

Selection process

- Combine multiple reports

- The unit of analysis of a systematic review is the **study**.
- Multiple reports from a same study must be identified.

Selection process

Combine multiple reports

- Study identification number
- Authors' names
- Location
- Intervention details
- Number of participants and baseline data
- Date and duration of the study
- Contact authors

Selection process

Combine multiple reports

- Study identification number
- Authors' nam
- Location
- Intervention
- Number of p
- Date and dur
- Contact authors

Choose and justify which report is used as the source of the results!

Ultimately, it may be necessary to collect data from different reports.

Selection process

Review full article

Contact authors

Decide which studies to include

Identify ongoing studies

Selection process



Review full article

- Studies should be included in the systematic review regardless of whether the data can be used.



Selection process



Contact authors

- Missing / raw data
- Measures of dispersion
- Methodological uncertainties
- Intervention details
- Possible errors
- Multiple reports

Selection process



Decide which studies to include

- The study selection process should be carried out independently by two people.
- The process for resolving conflicts should be defined from the outset.
- Record reasons for exclusion.

Selection process

Decide which studies to include

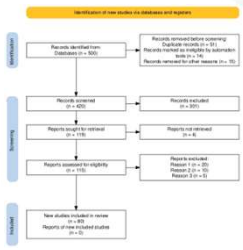


Selection process

- Support software
- Automatisation

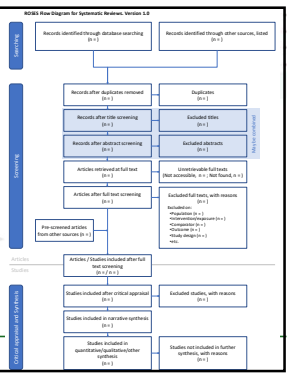
PRISMA flowchart

Document each step



ROSES flowchart

Document each step



Example

Can biomanipulation improve water quality in eutrophicated lakes?

- 28,329 search results
- 14,552 articles
- 551 (after title and abstract screening)
- 22 full texts not found
- 231 (after full text screening)

Data extraction

Rafael José Vieira
rafaelvieira@med.up.pt

Steps of a systematic review

1. Define the research question
2. Plan the eligibility criteria
3. Search for the primary studies
4. Select the studies
5. **Collect the data**
6. Assess the methodological quality of the studies
7. Analyse and present the results
8. Assess the certainty of the evidence
9. Interpret the results and draw conclusions

Data extraction

The aim of the data collection process is to gather information in sufficient detail to allow the **creation of a table of study characteristics and to conduct a narrative/quantitative synthesis** of study results.

As with study selection, data extraction should be carried out independently by two people.

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Source of data

- Extractor names
- Extraction date
- Source identification
  - Full article/Abstract
  - Authors
  - Date

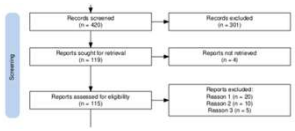
Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Eligibility criteria

- Confirm study inclusion
- Exclusion reasons



Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Methods

- Study design
- Recruitment and sampling
- Study duration
- Information required for critical appraisal/risk of bias assessment
- Statistical methods
- Funding and conflicts of interest

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Participants/Population

- Setting
- Characteristics

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Intervention/Exposure

- Intervention/Exposure and comparator
- Details on procedures

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Outcomes

- Outcome definition
- How outcomes were assessed, including scales and units of measurement
- Time

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Results

- Number of participants overall and by group
  - Included in the analysis
  - Those who withdrew, were lost to follow-up, or were excluded
- Summary data by group
- Effect size estimate with measure of dispersion or precision (confidence interval, standard error, standard deviation, *p*-value, etc.)
- Subgroup analyses

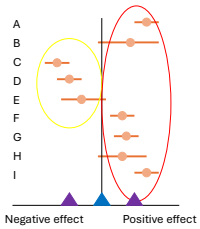
Which data to collect?

Results: Effect Size

- A single estimate of the size and direction of effect for each study
- The "effect size" is how we summarise the results from each study, the strength of the relationship between two variables
- The focus is on the magnitude and direction of effect
- Through 'standardisation', it allows comparison of results of different studies
- The effect size makes meta-analysis possible.
- There are several different ways of calculating the effect size for a study, each suited to different research situations…

Which data to collect?

Result	*N* studies
Significant -	2
Significant +	4
Non-significant	3



Which data to collect?

Results

Type	Data to collect
Dichotomous	• Number of participants with the event in each group • Total number of participants in each group
Continuous	• Mean value of the outcome in each group • Standard deviation of the outcome in each group • Sample size in each group
Counts and rates	• Number of events in each group • Person-time of follow-up in each group (e.g. person-years)

Which data to collect?

	Intervention			Comparator		
	Mean	SD	n	Mean	SD	n
Study 1	94	22	60	92	20	60
Study 2	98	21	65	92	22	65
Study 3	98	28	40	88	26	40
Study 4	94	19	200	82	17	200
Study 5	98	21	50	88	22	45
Study 6	96	21	85	92	22	85

Which data to collect?

- Source of data
- Eligibility criteria
- Methods
- Participants
- Intervention / Exposure
- Outcomes
- Results
- Other information

Which data to collect?

Other informations

- Main conclusions
- References to other relevant studies
- Need to contact the authors
- Comments from the study authors and/or the review

Common issues

Multiple reports

- The same study may lead to several publications (e.g. partial results, follow-up, secondary analyses).
- There is a risk of data duplication if these are treated as separate studies.
- There may be inconsistencies between publications (e.g. number of participants, outcomes assessed, results).
- It is not always clear that they refer to the same study.

Common issues

Multiple reports

- Group publications from the same study (e.g. by registration number, authors, methods).
- Select one main publication and use the others as supplementary sources for data extraction.
- Document decisions and maintain transparency in data selection.

Common issues

Outcomes reported heterogeneously

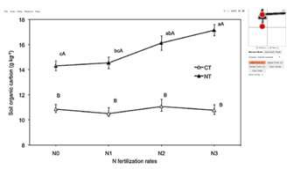
- Different studies may report a given outcome in different ways.
- If possible, convert it into a common measure.
- If conversion is not possible, report it in a table and text — do not exclude it from the review.

Outcomes for crop yield	*N* studies
Yield per hectare (kg/ha)	10
Yield adjusted for moisture content (%)	2
Total harvest weight (tonnes)	4

Common issues

Data reported in figures

- Contact authors
- Software supported extraction



Common issues

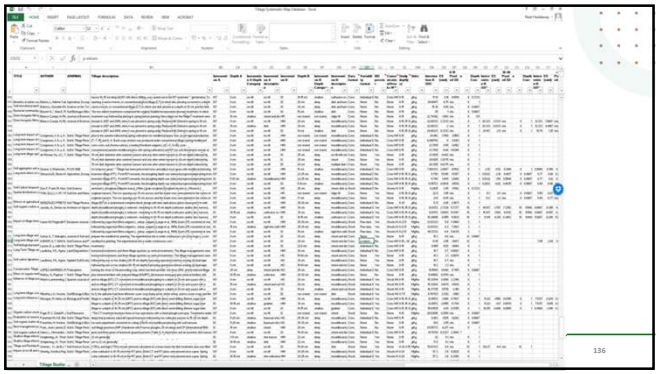
Missing data

- In the presence of missing data, the possibility of calculating the missing values from the data provided should be assessed.
  - It is possible to calculate the standard deviation from the standard error and sample size.
  - It is possible to calculate the standard error from the confidence interval or the p-value.
- If this is not possible, the authors should be contacted.

Common issues

Missing data

- If the authors do not respond, several approaches are possible:
  - Algebraic approximations (for example, calculating the mean and standard deviation from the median and interquartile range, provided the data are approximately symmetrically distributed).
  - Imputation of data from other studies with similar characteristics.
- It is important to conduct sensitivity analyses in the meta-analysis, particularly when the data are approximated or imputed.



Critical Appraisal

Nicola Randall
nrandall@harper-adams.ac.uk

What is it?

Transparent, objective, repeatable assessment of study reliability

Helps us:

- Identify *potential bias* in individual studies (internal validity)
- Judge how closely results reflect the *true effect*
- Assess *uncertainty* in individual studies and the overall synthesis
- Decide which studies should carry more *weight* in a synthesis
- Highlight gaps and improve future research design (eg identify deficiencies in experimental design)

how much confidence we can place in findings & why?

Why it matters in agri-food reviews

•Agricultural and ecological studies are often:

•Complex (multiple interventions, interacting variables)
•Context-specific (soil type, climate, crop/livestock)
•Resource-limited (small sample sizes, short timeframes)

•Critical appraisal can help ensure that:

•Findings are reliable despite field complexity
•Evidence is applicable across regions and systems
•Reviews do not overstate effects from weak studies

How to do it

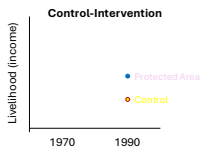
- Use or adapt a critical appraisal checklist/tool
- Consider:
  - What study design best addresses the review question?
  - Which design elements are essential (replication, randomisation, controls)?
  - How do real-world studies differ from the "ideal" design?
- Assess for factors that may reduce reliability or introduce bias
- Record all appraisal decision transparently

Internal validity (study reliability)

- **Study design** – BA, CI, BACI, RCT, observational, experimental
  - BACI designs are often strongest in field ecology/agriculture
  - BA/CI designs may miss confounders
- **Replication** – enough plots, farms, or time points?
- **Intervention assignment** – randomised? blocked? opportunistic? (*selection bias*)
- **Comparators** – were they equivalent in management, soil, climate? (*performance bias*)
- **Outcome measures** – were they appropriate and accurate?
  - e.g. direct SOC stocks vs proxies like soil colour (*detection bias*)
- **Missing data/dropouts** – were they reported and handled properly? (*attrition bias*)
- **Reporting bias** – selective publication of "positive" outcomes only

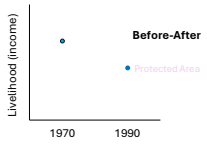
Factors that affect study reliability

Controls and Baseline measurements



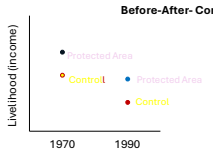
143

Factors that affect study reliability



144

Factors that affect study reliability



145

External validity (generalisability)

- **Population and setting**
- Crops, livestock, soils, regions representative of review scope?
  - eg applicability of N Europe wheat studies to sub-Saharan Africa maize
- **Intervention**
- Applied in ways that reflect real-world practice?
  - e.g. tillage with farmer machinery vs experimental micro-plots
- **Confounders and modifiers**
- Fertiliser history, rainfall, pest pressure, market access
- **Scales of measurement**
- Soil sampling depth (10 cm vs 30 cm)
- Experiment duration (1 year vs decades)
- Farm/plot scale vs landscape scale

Tools and frameworks for Critical Appraisal

- Many standard tools (CASP, EPHPP, JBI) originate from health/social sciences — useful as starting points
- CEE Critical Appraisal Tool (v0.3, prototype)
  - Designed for environmental/agri-food studies (interventions / exposures)
  - Focuses on *risk of bias* (internal validity)
  - **Does not yet assess external validity** environmentalevidence.org **also CEESAT (via CEEDER)** appraises review-level reliability (syntheses) rather than individual primary studies
- Custom / modified tools
  - Adapt checklists to include agri-food relevant criteria (e.g. soil depth, cropping systems, temporal replication)
  - Pilot-test modifications and record decision logic
- Key principles
  - Pre-specify in protocol
  - Use multiple reviewers with calibration
  - Justify all modifications and report transparently

What is the effect of prescribed burning in temperate and boreal forest on biodiversity, beyond pyrophilous and saproxylic species? A systematic review

Question/criterion	Yes	Partly	No	Unclear
Did the study have a temporal and/or spatial control?	BACI study	BA study or CI study	N/a (not eligible according to inclusion criteria)	Lacking sufficient information to judge
Degree of replication appropriate and representative? (to outcome measure)	Replicated burns, OR a single large burn area, with samples considered to be representative and spatially independent	Single burn with limited sampling within burn	A single/small burn area with very limited sampling within burn	Lacking sufficient information to judge
Does treatment allocation account for spatial heterogeneity? (e.g. blocked or randomised) and/or intervention and comparator sites well-matched i.e. similar at baseline (e.g. aspect, soil type, forest type)	Treatment allocation (multiple burn) or replication (within a single burn) accounts for spatial heterogeneity i.e. appropriately randomised, randomised and blocked or stratified OR stated/clear attempt at matching OR highly likely to be similar	Spatial heterogeneity not fully accounted for i.e. partial randomisation/blocking/stratification OR Moderately matched	Purposive treatment allocation that clearly does not account for spatial heterogeneity AND/OR Do not match	Lacking sufficient information to judge
No severely confounding factors present (factors likely to confound the effect of the intervention on outcome)? apart from those present at baseline (e.g. other interventions)	Confounding factors likely to be minimal, were minimised or lacking	Some confounders present, likely to have moderate impact on outcome	Subject to confounders with major impact on outcome (such that outcomes are not clearly the effect of the intervention)	Lacking sufficient information to judge
Intervention was likely appropriately and realistically applied?	Typical of a prescribed burn in terms of temperature, area, low intensity/severity of burn, or described as a prescribed burn and likely to be a typical prescribed burn	Partly typical of a prescribed burn in one (or more) aspect/s but not in others	Not typical of a prescribed burn (e.g. inappropriate temperature, very high intensity/severity)	Lacking any information/description of burn

Studies were assigned Low validity if any of the following factors applied	
	Any of these questions answered with "No" or "Unclear"
	• Did the study have a temporal and/or spatial control? • Degree of replication appropriate and representative?
OR	**Any of these questions answered with "No"**
	• Does treatment allocation account for spatial heterogeneity? and/or Intervention and comparator sites well-matched • No severely confounding factors present? apart from those present at baseline • Intervention was likely appropriately and realistically applied? • Outcome measure method was appropriate? • Study methodology and results are generalisable to other prescribed burns in temperate or boreal forest
Studies that were not assigned Low validity were considered to have Medium validity or Medium (unclear) validity if any of the following factors applied	
	Any of these questions answered with "Partly" • Did the study have a temporal and/or spatial control? • Degree of replication appropriate and representative? (to outcome measure)
OR	**Any of these questions answered with "Partly" or "Unclear":**
	• Does treatment allocation account for spatial heterogeneity? and/or Intervention and comparator sites well-matched • No severely confounding factors present? apart from those present at baseline • Intervention was likely appropriately and realistically applied? • Outcome measure method was appropriate? • Study methodology and results are generalisable to other prescribed burns in temperate or boreal forest

149

Recording and Transparency

- Pre-specify criteria in a **protocol**
- Keep detailed records of appraisal for each study (spreadsheets, coding sheets if not using an online review tool)
- Calibrate among reviewers to check consistency
- Critical appraisal can be:
- Done alongside full-text screening
- Integrated with data extraction (recording quality + results simultaneously)

150

How to use appraisal results

- **Exclude** very low-reliability studies (e.g. single replicate, poor design)
- **Weight** studies qualitatively (high/medium/low) in narrative synthesis
- **Sensitivity analyses**
  - e.g. Compare SOC results including vs excluding <2 year studies
- Helps test robustness of conclusions
- **Quantitative weighting** (meta-analysis) still based on study precision, but appraisal informs interpretation
- **Use to qualify conclusions** – transparent about confidence in findings
- **Identify research needs**
  - e.g. "Most studies short-term and shallow soil depths; long-term monitoring needed"

151

What is the effect of prescribed burning in temperate and boreal forest on biodiversity, beyond pyrophilous and saproxylic species? A systematic review

Eligibility for meta-analysis

Studies were considered unsuitable for meta-analysis (and no outcome data were extracted from them) if any of the following applied:

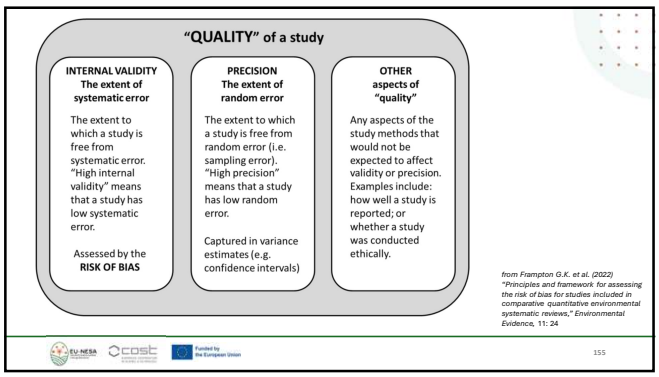
- The study provided quantitative data that were already provided in another relevant article (in cases of such redundant data, studies providing more information were selected for further synthesis, but missing information was filled in from linked studies).
- Measures of outcome variability and/or data on sample sizes were not available (and not possible to calculate from raw data)—effect sizes could not be calculated.
- Effects of burning were compared with effects of alternative levels of burning (rather than no burning). These studies were of limited value because they could not be compared with other studies in a quantitative analysis.
- Multiple interventions were applied concurrently in comparison with no intervention, e.g. thinning and burning compared with no intervention.
- Additional interventions (such as thinning or manipulation of grazing) had been carried out across the study areas (in both burned and unburned plots).

Critical Appraisal
Summary

- Conducted to assess internal and external validity of individual studies in a SR
- Based on study design and generalisability of findings
- Reporting important for transparency
- Need decide how critical individual study's susceptibility to bias/validity is before using study results to inform the review findings

Bias Class	Summary
1. Confounding bias	Distortion caused by uncontrolled variables affecting both exposure and outcome.
2. Selection bias	Systematic differences in who or what is included after exposure; includes confirmation bias and bias from missing data.
3. Misclassification of exposure	Inaccurate or imprecise measurement/definition of exposure or comparator, leading to misleading associations.
4. Deviation from planned exposure (performance bias)	Changes or non-adherence to intended treatments or exposures after the study begins.
5. Missing data (attrition bias)	Loss of data that differs between groups or relates to exposure/outcome, potentially distorting results.
6. Measurement bias (detection bias)	Systematic differences in outcome measurement between groups, intentionally or unintentionally.
7. Reporting bias	Selective reporting of outcomes, subgroups, or analyses that favour certain findings.
8. Statistical analysis bias	Errors or inappropriate statistical methods affecting validity (e.g. wrong tests, unblinded analysts).
9. Other biases	Any additional or design-specific biases not captured above (e.g. test accuracy studies).

Adapted (condensed) from Frampton G.K. et al. (2022) "Principles and framework for assessing the risk of bias for studies included in comparative quantitative environmental systematic reviews," Environmental Evidence, 11: 24



Types of Synthesis

Rafael José Vieira
rafaelvieira@med.up.pt

Types of synthesis

- A **systematic map** aims to investigate the state of research
  - Broader topic areas, research gaps, searchable database and visualisations
- A **systematic review** usually aims to answer a specific question
  - Narrative synthesis
  - Quantitative synthesis (meta-analysis)

157

Types of synthesis

- A **systematic map** aims to investigate the state of research
  - Broader topic areas, research gaps, searchable database and visualisations
- A **systematic** [illegible] question
  - Narrative synthesis
  - Quantitative synthesis (meta-analysis)

What research has been done?

What is known from what has been done?

158

Systematic maps

EU-NESA Training School

Porto, 28/29 September 2025

Rafael José Vieira
rafaelvieira@med.up.pt

Systematic maps

- Systematic maps are methodical overviews of the quantity and quality of evidence in relation to a broad (open) question of policy or management relevance.
- The process and rigour of the mapping exercise is the same as for systematic review except that no evidence synthesis is attempted to seek an answer to the question.

160

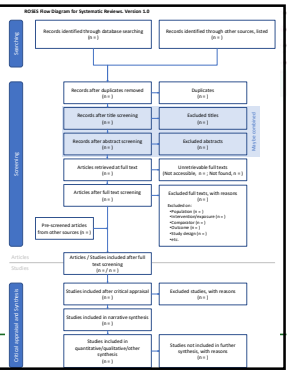
Any systematic map should:

- Be comprehensive/representative (avoid selective reporting)
- Visualise predefined variables (a priori design)
- Not present study findings, systematic maps are not a full synthesis
- Describe the evidence base, i.e. **synthetise**
  - Don't just describe each study individually

Any systematic map should contain:

- Systematic map methodology (flow diagram)
- Description of the evidence base using one or more of:
  - Summary tables
  - Summary figures
  - Heat maps
  - Evidence atlases
  - Interactive database

ROSES Flowchart

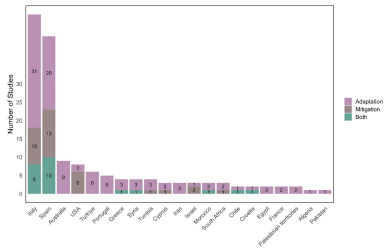


Summary tables

Service type	Ecosystem service	Number of reviews	Primary studies per review
Provisioning services	Food	23	1.76
	Fresh water	15	3.07
	Fiber and fuel	24	1.08
	Biochemical products	1	0.00
	Genetic materials	3	1.33
Regulating services	Climate regulation	15	3.20
	Hydrological regimes	30	6.50
	Pollution control	70	4.47
	Erosion protection	47	2.51
	Natural hazards	2	7.50
	Pest regulation	15	6.87
	Nutrient cycling	10	29.70
Cultural services	Spiritual and inspirational	0	–
	Recreational	6	1.67
	Aesthetic	8	1.38
	Educational	0	
Supporting services	Biodiversity	51	12.96
	Pollination	6	5.17
	Soil formation	7	20.57

Outcome type	Common categories of indicators used
Monetary income (direct sale)	Cash income from sales
	Dependency on harvest or sale income (percent of total income)
	Income inequality (Gini coefficient)
	Perceived impacts on cash income
	Poverty status (wealth rankings, poverty indices)
	Transaction costs (cost-benefits)
Monetary income (wage labor)	Access to forest-based wage labor
	Availability of forest-based employment (number of jobs)
	Change in income from wage labor
	Change in number individuals employed
	Perception of change in benefits
	Poverty status (wealth rankings, poverty indices)
	Transaction costs (cost-benefits)

Summary figures



Heatmaps



Evidence atlases





Knowledge gaps

- Identification of missing/under-represented groups
  - Countries
  - Populations
  - Interventions/exposures
  - Outcomes
  - Study designs

Number of studies

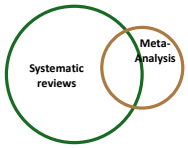
	Birds	Mammals	Reptiles
China	1	26	18
Hungary	16	14	17
Romania	8	0	7
Switzerland	9	12	13

Systematic reviews

Rafael José Vieira
rafaelvieira@med.up.pt

Systematic reviews

- All systematic reviews should include a narrative synthesis.
- If suitable, systematic reviews may also include quantitative synthesis (meta-analysis).



Narrative synthesis

- The textual approach to synthesising the findings.
- Must go beyond description of findings of individual studies:
  - What does the 'body of evidence' say?
  - How robust are findings?

Narrative synthesis

- Narrative synthesis should include:
  - Results and description of individual studies
  - Effect size for each comparison and outcome of interest
  - Description of context:
    - Study design / methodology
    - Intervention / exposure
    - Population / setting
  - Summary of methodological robustness

Narrative synthesis

- Separate tables/figures for main categories
  - Intervention/exposure
  - Outcomes
  - Species
- Use the tables/groupings to identify:
  - Themes and trends
  - Consistency of effect
  - Study clusters
  - Patterns of factos
- Consistency of findings across studies, if not explore why

Narrative synthesis

- Narrative synthesis alone when:
  - Insufficient data for meta-analysis
  - Large number of different outcome measures
  - Heterogeneity of intervention, measures or populations
  - When a combined mean would not make sense
- Still includes quantitative results from primary studies, effect sizes and confidence in those results

Meta-analysis

- It is not a specific type of study, but rather a set of statistical methods that allow the quantitative synthesis of scientific evidence.
- Although frequently used in the context of systematic reviews, there can be systematic reviews without a meta-analysis (and this may be the most appropriate option – conducting a meta-analysis requires that the primary studies are sufficiently similar).

Meta-analysis

Study reference	No. events treatment	No. entered control	
U.S. Collaborative 1984	9/ 200	15/ 206	RR=0.62
Smolders 1990	2/ 64	2/ 58	RR=0.91
Howie 1988	12/ 139	15/ 111	RR=0.64
■ Total (95% CI)	23/ 403	32/ 375	

How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

Meta-analysis

Study reference	No. events treatment	No. entered control	
U.S. Collaborative 1984	9/ 200	15/ 206	RR=0.62
Smolders 1990	2/ 64	2/ 58	RR=0.91
Howie 1988	12/ 139	15/ 111	RR=0.64
■ Total (95% CI)	23/ 403	32/ 375	

How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

But are all primary studies equal?

Meta-analysis

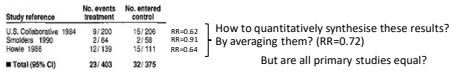
Study reference	No. events treatment	No. entered control	
U.S. Collaborative 1984	9/ 200	15/ 206	RR=0.62
Smolders 1990	2/ 64	2/ 58	RR=0.91
Howie 1988	12/ 139	15/ 111	RR=0.64
■ Total (95% CI)	23/ 403	32/ 375	

How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

But are all primary studies equal?

- Meta-analysis encompasses a set of statistical methods that allow the calculation of weighted measures based on quantitative data from multiple primary studies.

Meta-analysis



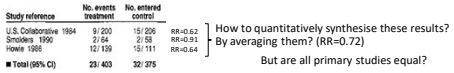
How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

But are all primary studies equal?

- Meta-analysis encompasses a set of statistical methods that allow the calculation of weighted measures based on quantitative data from multiple primary studies.

How to weight the studies?

Meta-analysis



How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

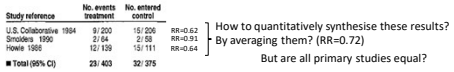
But are all primary studies equal?

- Meta-analysis encompasses a set of statistical methods that allow the calculation of weighted measures based on quantitative data from multiple primary studies.

How to weight the studies?

By their methodological quality?

Meta-analysis



How to quantitatively synthesise these results?
By averaging them? (RR=0.72)

But are all primary studies equal?

- Meta-analysis encompasses a set of statistical methods that allow the calculation of weighted measures based on quantitative data from multiple primary studies.

How to weight the studies?

~~By their methodological quality?~~

There is no consensus on how to do it

By their statistical quality (precision)

Meta-analysis

- To perform meta-analysis, one needs to obtain both the **effect size** measure and the **standard error** from each study.

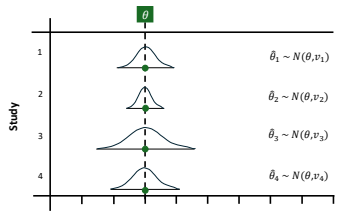
1 group

Continuous dependente variable	Dichotomous dependente variable
• Mean	• Proportion (e.g. prevalence, incidence) • Incidence rate

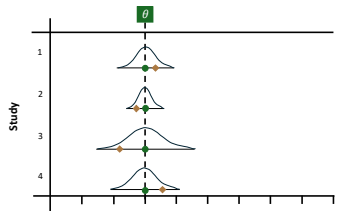
2 groups

Continuous dependente variable	Dichotomous dependente variable
• Mean difference • Standardised mean difference (e.g. Cohen's *d*, Hedges' *g*)	• Risk ratio • *Odds* ratio • Incidence rate ratio

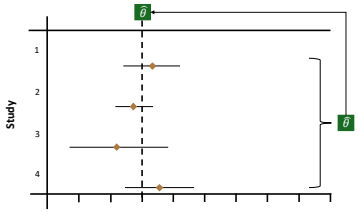
Fixed-effects model



Fixed-effects model



Fixed-effects model



Fixed-effects model

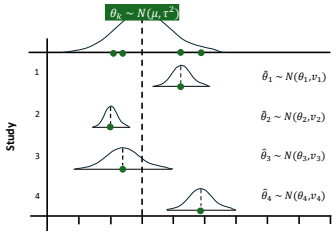
- In meta-analysis, greater weight is given to the more precise studies (those with a smaller standard error).
- To calculate the weight of each study, we use the variance of each study, which corresponds to the square of the standard error. Thus, the weighting is calculated using the inverse variance method:

$$w_k = \frac{1}{SE_k^2} = \frac{1}{v_k}$$

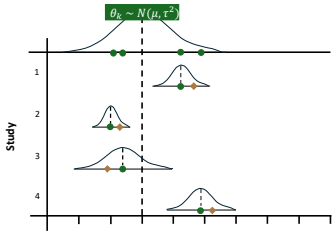
Random-effects model

- In most cases, it is not realistic to assume that all studies measure the effect size in the same population (studies may differ in terms of participant selection criteria, mode of intervention delivery, etc.).
- The random-effects model assumes that the studies come from different populations, estimating different effect sizes. The aim is therefore to estimate the mean of the distribution of effect sizes.

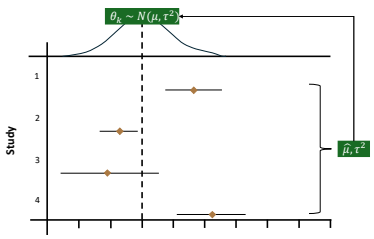
Random-effects model



Random-effects model



Random-effects model



Random-effects model

- In the random-effects model, it is necessary to account for the error, ζ_k. To do this, the variance of the distribution of the true effect sizes, τ^2, must be estimated.
- The weighting is therefore given by:

$$w_k = \frac{1}{v_k + \tau^2}$$

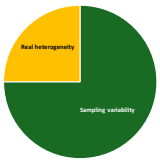
- There are several methods available to estimate τ^2 in the context of the random-effects model.

Heterogeneity

- Differences between studies beyond those that can be explained solely by sampling variability.
- These may arise from clinical and/or methodological differences between the primary studies.
- Ways of assessment:
  - **Cochran's Q statistic:** significant heterogeneity if $p < 0.10$.
  - **I^2 statistic**

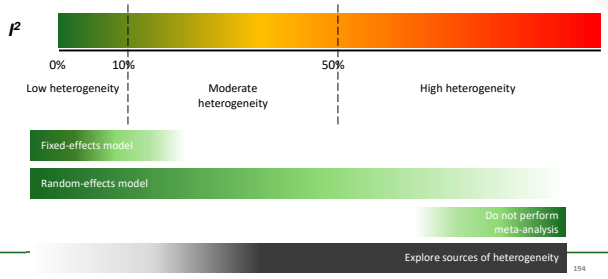
I^2 statistic

- Quantifies the proportion of variability beyond that which would be expected by chance.



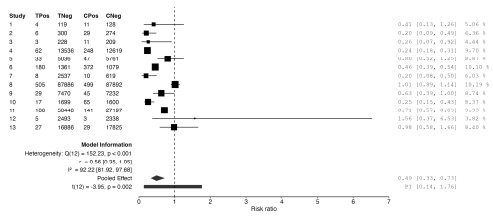
$I^2 =$ 

I^2 statistic



194

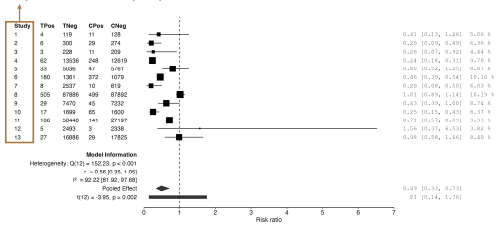
Forest plots



195

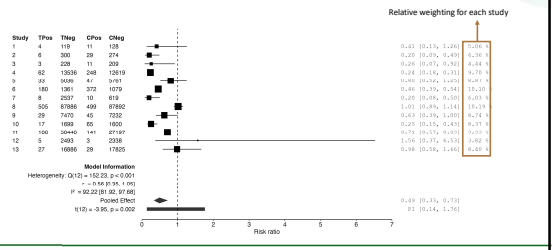
Forest plots

Study list, often identified by first author's last name and year of publication

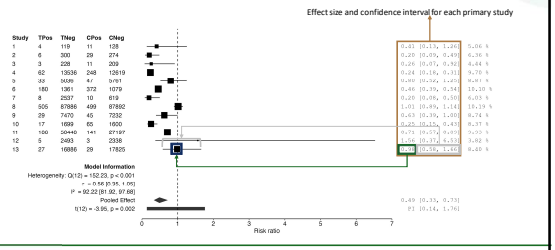


196

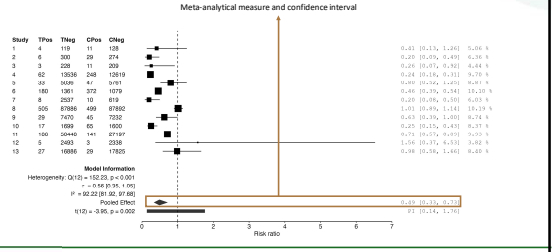
Forest plots



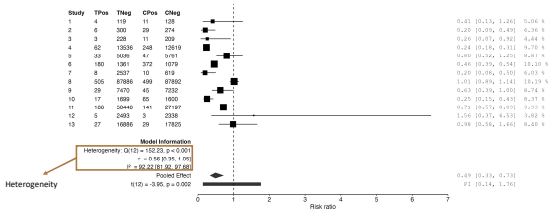
Forest plots



Forest plots

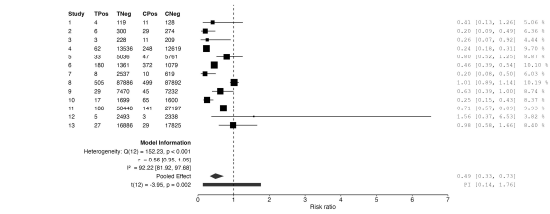


Forest plots



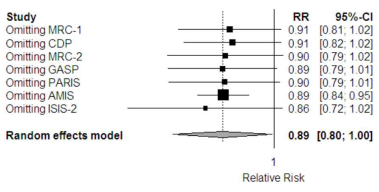
Forest plots

How to interpret?



Exploring sources of heterogeneity

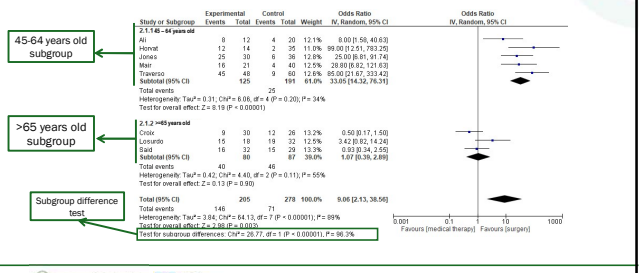
- **Sensitivity analysis**: Modification of the initial assumptions of the study (e.g. removal of each individual study or of specific sets of studies)



Exploring sources of heterogeneity

- **Subgroup analysis**: Presentation of meta-analytic results separately by categories of a given categorical variable
  - Easy to interpret
  - Dependent on the number of primary studies
  - Difficulty in dealing with continuous variables (need for categorisation)
  - Should be prespecified

Exploring sources of heterogeneity



Exploring sources of heterogeneity

- **Meta-regression**: Application of regression models (univariable or multivariable) in the meta-analytic context to identify moderator variables of heterogeneity
  - Models can handle both categorical and continuous variables
  - Require a large number of primary studies (at least 10 for a univariable model)
  - More difficult to interpret

Exploring sources of heterogeneity

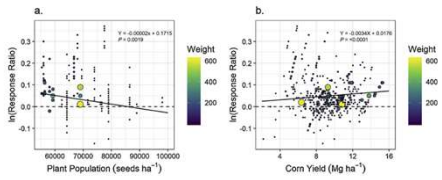


Fig. 5. Relationship between the natural log of the response ratio (LRR; yield of corn containing a starter fertilizer/yield of corn not containing a starter fertilizer) and a) corn plant population (seeds ha^{-1}) and b) corn yield level (Mg ha^{-1}). Horizontal dashed lines indicate where the LRR equals zero. The size and color of individual data points are proportional to the calculated weight of the observation.

Dissemination biases

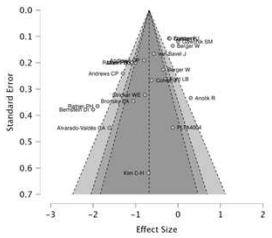
- Studies are often rejected not necessarily because of their methodological shortcomings, but rather because of the results they obtain (they are not considered “sufficiently interesting” or aligned with expectations).
- Searching the “grey literature” and not excluding studies based on language reduces the risk of dissemination bias.
- Funnel plot analysis can help to infer the presence of dissemination bias.

Dissemination biases

Table 10.1.a Definitions of some types of reporting biases

Type of reporting bias	Definition
Publication bias	The *publication* or *non-publication* of research findings, depending on the nature and direction of the results
Time lag bias	The *rapid* or *delayed* publication of research findings, depending on the nature and direction of the results
Multiple (duplicate) publication bias	The *multiple* or *singular* publication of research findings, depending on the nature and direction of the results
Location bias	The publication of research findings in journals with different *ease of access* or *levels of indexing* in standard databases, depending on the nature and direction of results.
Citation bias	The *citation* or *non-citation* of research findings, depending on the nature and direction of the results
Language bias	The publication of research findings *in a particular language*, depending on the nature and direction of the results
Outcome reporting bias	The *selective reporting* of some outcomes but not others, depending on the nature and direction of the results

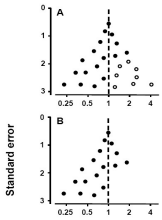
Dissemination biases



In a **funnel plot**, the horizontal axis represents the **effect size**, while the vertical axis shows the **standard error**. Larger and more precise studies are usually located at the top of the plot and close to the overall meta-analytic estimate, whereas smaller and less precise studies appear at the bottom, with more scattered results.

When the funnel is **symmetrical**, this suggests the absence of relevant bias; however, **marked asymmetry** (for example, studies concentrated only on one side) may indicate dissemination bias.

Dissemination biases



Panel A: Symmetrical plot in the absence of bias.

Panel B: Asymmetrical plot in the presence of dissemination bias.

Systematic maps, rapid reviews and alternative approaches to evidence syntheses

Nicola Randall
nrandall@harper-adams.ac.uk

Systematic mapping

Kiely et al. CABI Agriculture and Bioscience (2023) 4:42
https://doi.org/10.1186/s43170-023-00183-1

CABI Agriculture and Bioscience

SYSTEMATIC MAP Open Access

The application of allelopathy in integrated pest management systems to control temperate European crop pests: a systematic map

Charlotte Kiely*, Nicola Randall and Magda Kaczorowska-Dolowy

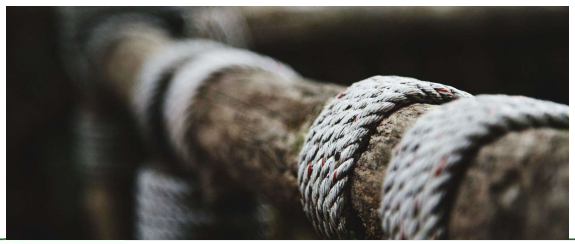
Abstract

Background Pesticides perform vital roles within agriculture but growing concern for their impact on the environment and non-target organisms has created a market for biopesticides with fewer ecological impacts. One source of biopesticides is allelochemicals, here defined as compounds released by an organism that have an inhibitory or stimulatory effect on neighbouring organisms. The focus of this study is allelopathic plants and their inhibitory effects on invertebrate herbivorous agricultural pests of temperate Europe. A systematic map is required to describe the current state of research and collate evidence.

Methods Two academic databases were searched for relevant studies in temperate climates. The results were imported into EPPI-Reviewer, duplicates removed, studies screened and data extracted into a searchable database following the inclusion criteria and coding tool set out in the protocol. Screening consistency was checked at each stage using 5% of the studies. Critical appraisal was not conducted. Each unique combination of key variables (pest, plant, allelochemical, application method, intervention form) was treated as a separate datapoint or experiment. The data was then analysed and cross-tabulated to produce descriptive statistics and heatmaps.

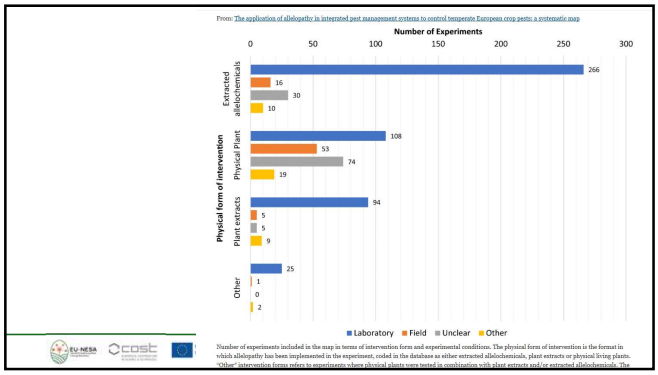
Results This systematic map produced a database which included 243 studies containing 717 experiments from 5550 initial results. Research was unevenly distributed among all key variables with a distinct bias towards extracted allelochemical experiments under laboratory conditions. Allyl isothiocyanate was the most studied

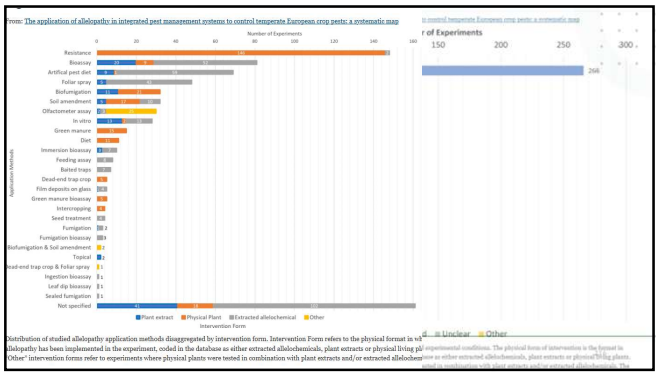
What kinds of detail might you want to represent?



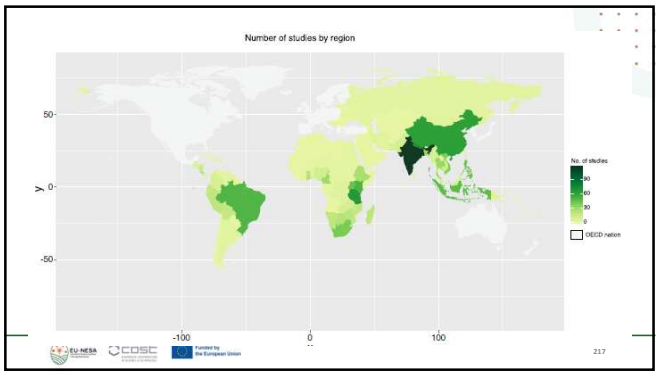
Use a variety of approaches to organise and present evidence

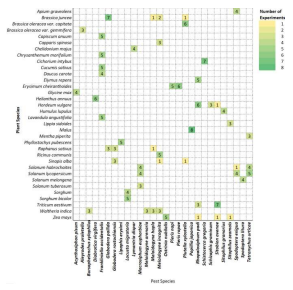
TEXTUAL DESCRIPTION | GRAPHS | MAPS | TABLES | DIAGRAMS

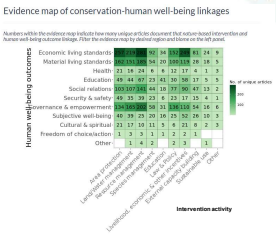




Distribution of studied allelopathy application methods disaggregated by intervention form. Intervention Form refers to the physical format in wh
Allelopathy has been implemented in the experiment, coded in the database as either extracted allelochemicals, plant extracts or physical living pl
Other' intervention forms refer to experiments where physical plants were tested in combination with plant extracts and/or extracted allelochem







Nature and people portal – interactive mapping

- Evidence Based Map Portal (natureandpeopleevidence.org)

We can also rate weight of evidence e.g. where systematic review evidence exists

- http://gapmaps.3ieimpact.org/evidence-maps/forest-conservation-gap-map

Review of reviews

- Similar process to review of primary research, but instead using existing reviews
- Less common in environmental evidence fields as there are less existing systematic reviews than there are in fields such as medicine.

Review of reviews

- Similar process to review of primary research, but instead using existing reviews
- Less common in environmental evidence fields as there are less existing systematic reviews than there are in fields such as medicine.

Abstract

Relevant systematic reviews and impact evaluations were systematically sought and described in order to assess the size and nature of the evidence-base about the effectiveness of interventions for smallholder farmers in Africa. A total of 21 relevant systematic reviews and 415 reports of impact evaluations were identified. This paper describes this African evidence-base in terms of the interventions and outcomes assessed, and the geographical spread of the primary research across the continent. Gaps in the evidence-base are identified and recommendations made for future research.

Stewart, R., Erasmus, Y., Zaranyika, H., Rebelo Da Silva, N., Korth, M., Langer, L., Randall, N., Madinga, N. and de Wet, T. 2014. The size and [illegible] map. *[illegible]* 58-68.

https://gapmaps.3ieimpact.org/evidence-maps/map-evidence-maps-relating-sustainable-development-lmics

Multi- stage approaches to reviews

- Sometimes you may want to carry out multiple reviews or maps/reviews at different stages.
- For example funder-set question: What are the effects of interventions such as training, innovation, new technology, grants/subsidies, microfinance, extension, policy etc. for smallholder farms in Africa (eg ag investment & innovation, yields/productivity, income, food security, nutrition).
- Very broad question so a three stage approach enabled:
  - funding and time resources to be maximised,
  - prioritisation of work to a topic area of interest to funders
  - An in-depth background to the existing evidence base
  - Avoid wasting time on an 'empty' systematic review

  This approach also encouraged and supported substantive interaction between our review team and the key policy stakeholders involved in this review.

Multi- stage approaches to

- Stage 1: an overview of systematic reviews for these topics
  - avoid duplication, and begin to identify primary research for the next stage.
- Stage 2: A systematic map on relevant interventions and outcomes for smallholder farmers in Africa.
  - Thematic framework to consult with advisory group and funders about where best to focus
- Stage 3: A systematic review of one focused area of interest
  - What are the effects of training and new technology / innovation interventions on the food security and income of smallholder farmers in Africa?

Stewart R, Erasmus Y, Zaranyika H, DaSilva N , Korth M, Langer L, **Randall NP**, Madinga N & de Wet T (2014) The size & nature of the evidence-base for smallholder farming in Africa *Journal of Development Effectiveness*

Natalie Rebelo Da Silva, Hazel Zaranyikal, Laurenz Langer, Nicola Randall, Evans Muchiri, and Ruth Stewart (2016) Making the Most of What We Already Know: A Three-Stage Approach to Systematic Reviewing Evaluation Review

Stewart, R., Langer, L., Da Silva, N.R., Muchiri, E., Zaranyika, H., Erasmus, Y., **Randall, N.,** Rafferty, S., Korth, M., Madinga, N. and de Wet, T., 2015. The effects of training, innovation and new technology on African smallholder farmers' wealth and food security: A systematic review. *Campbell Systematic Reviews*, 11(16). (*Available in English and in Spanish*)

Rapid reviews

Rapid Review

Criteria

Rapid Reviews are evidence syntheses that would ideally be conducted as a Systematic Reviews, but where methodology needs to be accelerated and potentially compromised to meet the demand for evidence on timescales that preclude Systematic Review conducted to full CEE or equivalent standards. As with a Systematic Review, Rapid Reviews are reviews of evidence relevant to a clearly formulated question that use systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included within the

Official journal of

Collaboration for Environmental Evidence

Submit manuscript

Editorial Board

Rapid reviews

Discuss:

- Why would we consider these?
- With reference to your own questions, at which stages do you think you could cut time?
- Where do you think issues might arise with rapid reviews?
- What can you do to overcome these issues?

Interpreting synthesis data, and critically appraising reviews

Nicola Randall
nrandall@harper-adams.ac.uk

- If you are going to make decisions based on review findings, it is as important to critically appraise a review as it is to critically appraise primary research.
- We often do this informally
- Or we could use a tool to help us.

CEEDER and CEESAT

- CEESAT is the CEE Synthesis Assessment Tool that can be used to critically appraise a review.
- Reviews that have been evaluated using this tool, are presented in the CEEDER database of reviews.

- https://environmentalevidence.org/ceeder/about-ceesat/
- https://environmentalevidence.org/ceeder-search/

Strengths and limitations of different approaches

Nicola Randall
nrandall@harper-adams.ac.uk



- There are a number of different approaches to synthesising information in an evidence synthesis.
- Typically we see
  - Meta-analysis/pooling
  - Narrative synthesis
  - Evidence mapping

- Meta-analysis

Strengths	Limitations
• Provides a systematic approach to make sense of evidence base • Increased power to give a more accurate indication of effect size • May give a simple take home message • Can prevent research waste • Can prevent activity waste	• Only as reliable as the studies that inform it (rubbish in, rubbish out) • May give false sense of precision/ accuracy /meaningfulness • Teams working on the same question have come to different conclusions • Doesn't overcome problems of publication (and other) bias • Small data sets may be less reliable

- Narrative synthesis

Strengths	Limitations
• Provides a systematic approach to make sense of evidence base • Can accommodate variety on intervention, outcomes, methods *etc.* • May be the only sensible approach to some questions/ evidence bases	• Potential for bias in presentation & interpretation • Lack of a simple "take home" message

- Evidence mapping

Strengths	Limitations
• Provides a systematic approach to make sense of evidence base • Can accommodate variety on intervention, outcomes, methods etc • May provide clear guidance for future research agenda	• Potential for misinterpretation • Doesn't tell you what works

Tools and resources to support evidence syntheses

Nicola Randall
nrandall@harper-adams.ac.uk

- Methodology help: CEE Guidelines
- Open/public access to protocols (and feedback): PROCEED
- CEE critical appraisal tool (prototype)
- Reporting standards/checking: ROSES
- Softwares during the process: various!

- Reliability of reviews: CEEDER (using CEESAT)

236

CEE Guidelines and Standards

Guidelines and Standards for Evidence Synthesis in Environmental Management

VERSION 5.1

2022

Editors: Andrew S Pullin, Geoff K Frampton,

Barbara Livoreil & Gillian Petrokofsky

237

Tools and resources for Protocols e.g. PROCEED

- https://www.proceedevidence.info/

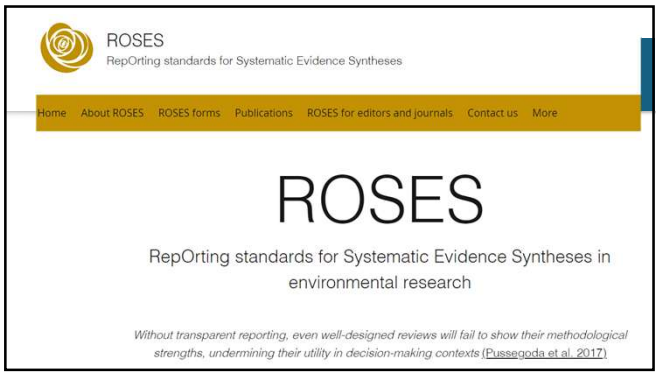


ROSES reporting forms

- ROSES forms help review authors to ensure that all relevant methodological information is reported in their review, and help editors and peer-reviewers to critique the reliability and validity of a review.

https://www.roses-reporting.com/

ROSES
RepOrting standards for Systematic Evidence Syntheses

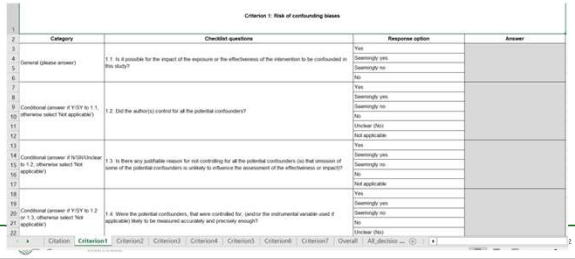


Checklists for coding/data extraction creation

- https://casp-uk.net/casp-tools-checklists/

CEE Critical appraisal tool (CEECAT)

https://environmentalevidence.org/cee-critical-appraisal-tool/



Criterion 1: Risk of confounding biases

Category	Checklist questions	Response option	Answer
General (please answer)	1.1. Is it possible for the impact of the exposure or the effectiveness of the intervention to be confounded in this study?	Yes Seemingly yes Seemingly no No	
Conditional (answer if Y/SY to 1.1, otherwise select 'Not applicable')	1.2. Did the author(s) control for all the potential confounders?	Yes Seemingly yes Seemingly no No Unclear (No) Not applicable	
Conditional (answer if N/SN/Unclear to 1.2, otherwise select 'Not applicable')	1.3. Is there any justifiable reason for not controlling for all the potential confounders (so that omission of some of the potential confounders is unlikely to influence the assessment of the effectiveness or impact)?	Yes Seemingly yes Seemingly no No Not applicable	
Conditional (answer if Y/SY to 1.2 or 1.3, otherwise select 'Not applicable')	1.4. Were the potential confounders, that were controlled for, (and/or the instrumental variable used if applicable) likely to be measured accurately and precisely enough?	Yes Seemingly yes Seemingly no No Unclear (No)	

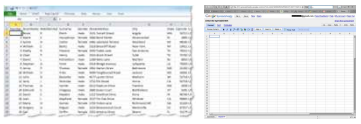
Citation | Criterion1 | Criterion2 | Criterion3 | Criterion4 | Criterion5 | Criterion6 | Criterion7 | Overall | All_decisions ...

Software to support conduct

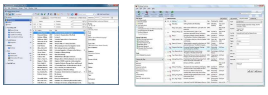
- Softwares exist for support throughout the review, though the most useful is during screening..

Software

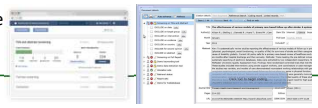
- Spreadsheets



- Reference management software



- Review management software



Excel/Spreadsheet software

- Pros:
  - Easy to access/use
  - Can work 'offline'
  - Easy to share
  - Easy to remove identical duplicates (Data>Remove duplicates)
- Cons:
  - Version control can be a challenge
  - Easy to 'damage' library
  - Difficult to identify partial duplicates

Reference management software

- *e.g. EndNote, Mendeley, Zotero, etc.*
- Pros:
  - Designed for citations
  - Generally easy import/export citations
  - Able to identify partial duplicates
  - Can work 'offline'
- Cons:
  - May be expensive
  - Less familiar
  - Challenging to share
  - Not designed for rapid screening
  - Sometimes easy to damage library

Review management software

- *e.g. EPPI-Reviewer, RevMan, CADIMA, Covidence, etc.*
- Pros:
  - Specifically designed for reviews
  - Designed for rapid screening
  - Designed for sharing
  - Allows full documentation of activities
  - Difficult to damage library
- Cons:
  - May be expensive
  - Not as familiar
  - Some cannot work 'offline'

EPPI-Reviewer

- Institute of Education (EPPI Centre)
- Specifically designed for SRs
- Cheap (£10/month), 1 month trial
- Shareable review £35/month
- Permanent read-only access, PAYG
- Excellent support



EPPI-Reviewer features

- **Reference management**: (Limited) search, reference import (.ris), storage (.pdf, .doc), duplicate removal (manually/ automatically)
- **Study classification and data extraction**: (Multiuser) coding (e.g. inclusion/exclusion; study description), consistency checking (no of disagreements), text mining/automatic text clustering
- **Synthesis:** Meta-analysis, inductive (line-by-line) coding
- **Overall review management:** Task allocation, progress reports, export of coding decisions/review data, review flow charts

➢ More info here: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3396
➢ Manual and instructional videos here: http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2933

Rayyan

- https://www.rayyan.ai/

RevMan



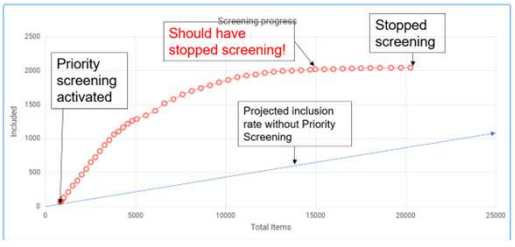
CADIMA

- Free (open access)
- Specifically designed for SRs (fp7 GRACE)
- Uses CEE Guidelines
- Based on EPPI Reviewer

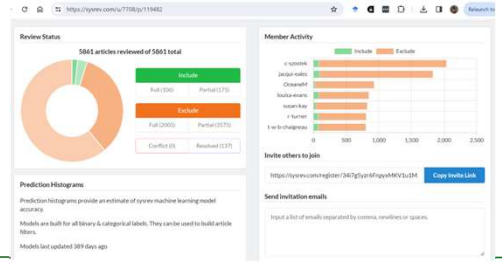
CADIMA

https://www.cadima.info/index.php/area/evidenceSynthesisDatabase

Use of machine learning – priority screening (Eppi Reviewer)



SysRev



ESMARConf
Evidence Synthesis and Meta-Analysis in R Conference

HOME MISSION EVENTS RECOF

OUR MISSION AT ESMARCONF

ESMARConf is a FREE, online annual conference series dedicated to evidence synthesis and meta-analysis in R. Our aim is to raise awareness of the utility of Open Source tools in R for conducting all aspects of evidence syntheses (systematic reviews/maps, meta-analysis, rapid reviews, scoping reviews, etc.), to build capacity for conducting rigorous evidence syntheses, to support the development of novel tools and frameworks for robust evidence synthesis, and to support a community of practice working in evidence synthesis tool development.

Central to our Mission at ESMARConf are the principles of accessibility, equity, fairness and family. We believe that everyone should have unrestricted access to tools necessary to conduct efficient, transaparent, accessible, and rigorous evidence syntheses, as well as resources needed for training and capacity building.

Summary

- Need for transparency and repeatability
- Review management software important: many search results and duplicates
- Various software options: ease of use, price, accessibility, collaboration

Interactive mapping

- See the 3ie examples but you could also use review management software or specific software for mapping e.g. eviatlas if you don't have suitable review software
- https://estech.shinyapps.io/eviatlas/

CEEDER (using CEESAT)


COMMENTARY Open Access

Standards of conduct and reporting in evidence syntheses that could inform environmental policy and management decisions

Abstract

Environmental Evidence

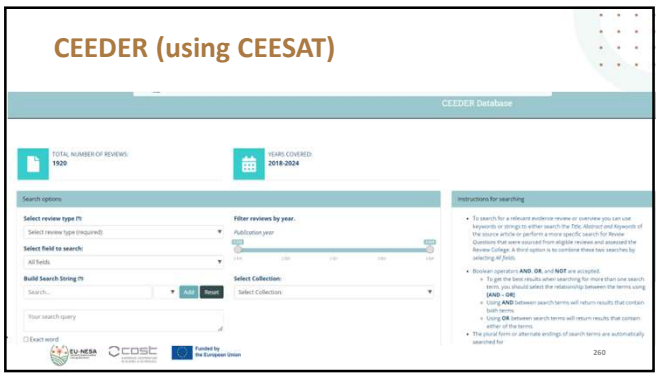
METHODOLOGY Open Access

Evidence maps and evidence gaps: evidence review mapping as a method for collating and appraising evidence reviews to inform research and policy

Abstract

- CEESAT is the CEE Synthesis Assessment Tool that can be used to critically appraise a review.
- We don't have time to use CEESAT to evaluate a review, but instead we will look at how other people have used it.

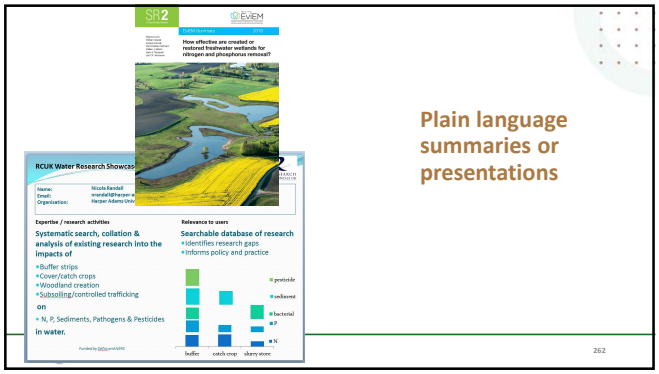
CEEDER (using CEESAT)



Communicating evidence syntheses when finished!

EU-NESA cost Funded by the European Union 261

Plain language summaries or presentations



262

Policy Briefs or even videos



Social Networks

ResearchGate SCIENTIFIC NETWORK · Academia.edu · LinkedIn · facebook

- Links to Protocol and Final Review Documents
- Calls for submission of evidence (questions)
- Polls
- Creation of mailing lists