QUEEN'S UNIVERSITY BELFAST

Cochrane Methodology Reviews

Introduction to systematic reviews

20 October 2026

Professor Mike Clarke
m.clarke@QUB.ac.uk

Learning objectives

To help you:

- understand the importance of systematic reviews in evidence-informed decision making
- formulate the research question for a systematic review, to cover key aspects of the population, interventions or exposures and outcome measures
- set the eligibility criteria for a systematic review
- plan a search for studies for a systematic review, including the choice of sources to search and the words or phrases to use in the search
- extract data from the eligible studies
- understand the basic principles of meta-analysis techniques used for systematic reviews, including subgroups and heterogeneity

What's a systematic review?

What do you think it is?

Why do we bother doing reviews?

We need to identify and resolve uncertainties

We need to know

What works

for

whom

by how much

We need to know

What does not work

We need to know

What might be harmful

We need to know

What needs more research

Good research provides needs to minimise chance and bias

Good research provides good evidence

Good evidence
used properly
leads to good practice

Good practice leads to better outcomes

Why do we need systematic reviews?

- Overwhelming amount of healthcare literature.
- New research is rarely reported in context.
- Reliable knowledge is essential for better health care.
- Better health care is essential for better health.

Systematic reviews …

- seek to minimise bias.
- allow the methods to be shown in a transparent way.
- avoid undue emphasis on a single piece of research or a single opinion.
- maximise the power of existing research.
- identify the need for future research.
- are scientific research.

Systematic reviews are scientific research

- State objectives and eligibility criteria.
- Identify (all) potentially eligible studies.
- Apply eligibility criteria.
- Assemble most complete dataset feasible.
- Analyse the dataset, using statistical synthesis and sensitivity analyses, *if appropriate and possible*.
- Prepare a structured report.
- Keep the review up to date.

What is a good systematic review?

Criteria for a ‘good’ systematic review (1)

- Does it address a focused question?
- Are the inclusion criteria appropriate?
- Is it unlikely that important, relevant studies were missed?
- Was the validity of the included studies appraised?
- Were assessments of studies reproducible?
- Are the results similar from study to study?

Does the review address a focused question?

What is the question for the review?

Question formulation

A clear research question is key to a systematic review

- it should provide the **PI(E)CO** for the review: population, interventions or exposures and comparison, outcomes and study design
- and might include the types of study that will be used to answer the question

Are the inclusion criteria appropriate?

What are the eligible participants, interventions/exposures, outcome measures and study designs?

Eligibility (inclusion) criteria

Clear eligibility criteria are key to a systematic review

- they determine what belongs in the review

Is it likely that important, relevant studies were missed?

Did the search look in the right places, in the right way, at the right time?

Searching

Comprehensive searching is key to a systematic review

The search strategy should have considered

- where to search (e.g., types of database and other resources)
- what terms to search for
- languages that might be used to report the studies
- time period to cover

Was the validity of included studies appraised?

Was it possible to identify which studies were “good”?

Were assessments of studies reproducible?

Have the "good" studies been identified in a reliable way?

Quality assessment

Appraising the quality of the included studies is a key step in a systematic review

- it determines whether the answer is correct
- it minimises the impact of “Garbage In, Garbage Out”
- it increases confidence in the findings of the included studies and of the review

Are the results similar from study to study?

Do the results vary? Should they vary? Does it matter?

Heterogeneity

Carefully considering heterogeneity is a key step in a systematic review

- it can occur in all domains in the review: *Participants, Interventions* or *Exposures, Outcome measures, Study designs*
- heterogeneity in the way that an outcome was measured or reported might prevent a meta-analysis

Criteria for a ‘good’ systematic review (2)

- What are the overall results of the review?
- How precise are the results?
- Can the results be applied to my care, or that of my patients?
- Were all the important outcomes considered?
- Are the benefits worth the harms and costs?

Two examples for today's group work

- Does "five a day" improve people's oral health?
- How long do dental implants survive?

Formulating the question for a review

Planning and doing a review depends on the question to be answered

Does "five a day" improve personal oral health?

compared to what?

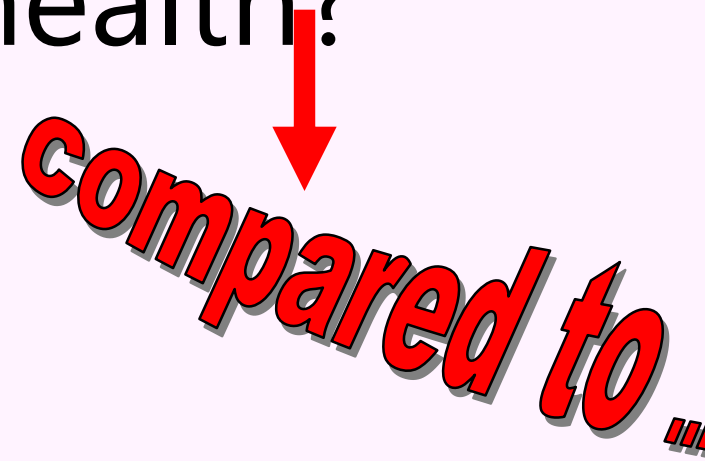
compared to ….?

Does “five a day” improve people’s oral health?

Does “five a day” improve people’s oral health?



Does “five a day” improve people’s oral health?



Does “five a day” improve people’s oral health?

Does “five a day” improve people’s oral health?

Does “five a day” improve people’s oral health?

Does “five a day” improve people’s oral health?

Does “five a day” improve people’s oral health?

Question formulation

- You need a clear question. It might change during the planning, but you hope it won't change during conduct.
- Each word in the question is important. Does it narrow or broaden your review? Will other people think each word or phrase means the same thing as you think it means (e.g. "fruit"; "people", "oral health")?
- Each word in the question can be expanded upon in the eligibility criteria for the review.

Setting the eligibility criteria

Key elements in defining the question and the eligibility criteria

Participants (or Population)
Interventions (or Exposures)
Comparators
Outcome measures
Study designs

What eligibility criteria would you use for

Does “five a day” improve people’s oral health?

Eligibility criteria

- The eligibility criteria provide the rules for what should and should not be included in **your** review. They are the **inclusion** and **exclusion** criteria.
- Use a framework for setting the criteria, such as the types of **participants**, **interventions** **(or exposures)**, **comparators**, **outcome measures** and **types of study**.
- Think about the types of study that might have been done and decide if you would want them in your review.
- A study's results must not influence your decision to include it.
- Clear eligibility criteria help you avoid having to make decisions about studies, after you know its results.

How long do dental implants survive and what impacts on their survival?

Searching for studies

What to consider when searching

- What terms to search for?
- What types of database and other resources to search?
- What countries?
- What languages?
- What time period?

Planning the search terms

Divide the review into the eligibility components

- Participants
- Interventions (including comparator)
- Outcomes
- Study designs

The database you are searching will try to help you by only showing the things that you ask for. It might hide millions of other things.

Choosing the search terms

In listing terms for each component, consider:

- Synonyms (from different times and places) *(e.g. other words for implants)*
- Other words and phrases that are related to what you're interested in *(e.g. reasons for implants)*
- Words that are broader *(e.g. head and neck or oral instead of dental)*
- Words that are more focused *(e.g. specific implants)*
- Index terms or keywords (which may differ across databases)

Combining search terms

Two main ways to link terms

- AND (decreases the number of hits, requires every item to be present)
- OR (increases the number of hits, requires any of the items to be present)
- Try to avoid NOT (it might remove records that are eligible)

Combining the components

Within the component, use OR to combine the terms and then use AND to combine the components. But ...

- Do you need all the components?
- Are you confident that you have all the terms within each component?
- Which component is least likely to be relevant?
- Which component is most likely to be relevant?
- The ideal might be the component that is most likely to be relevant and has the highest proportion of good things amongst its hits.

Where to search

- Bibliographic databases (e.g. PubMed, Embase)
- Journal websites (e.g. Cochrane Library)
- Trial registries (e.g. ClinicalTrials.gov, WHO portal)
- Handsearching of journals and conference proceedings
- Grey literature (e.g. government reports)
- Reference lists in articles
- Correspondence with researchers

What search terms would you use for

How long do dental implants survive?

Extracting the data

Data extraction - why

Reasons for doing data extraction:

- Remembering the information
- Organizing information into a particular structure
- Summarising the content of the reports
- Ensuring that you look for the key things
- Making it easier to compare different studies
- Making it easier to do meta-analyses and subgroup analyses
- Assessing quality

Data extraction - how

- Make a list of the items you need to summarise each include study. This will become your data extraction form
- Do you need all the data?
- How much detail do you need?
- Plan what you will use the data for. For example, do you want to know the mean age, the age range, the number of people in different age groups, or the results for different age groups separately?
- Might you just record that an article has some information (and where it is) so that you can extract it later if you wish?

What data would you extract for

What factors impact on the survival of dental implants?

Statistics of reviews: meta-analysis

Key statistical principles of meta-analysis

- Patients in one study are not directly compared with those in another study
- Each study is analysed separately
- Summary statistics are calculated for each study
- The summary statistics are combined in the meta-analysis

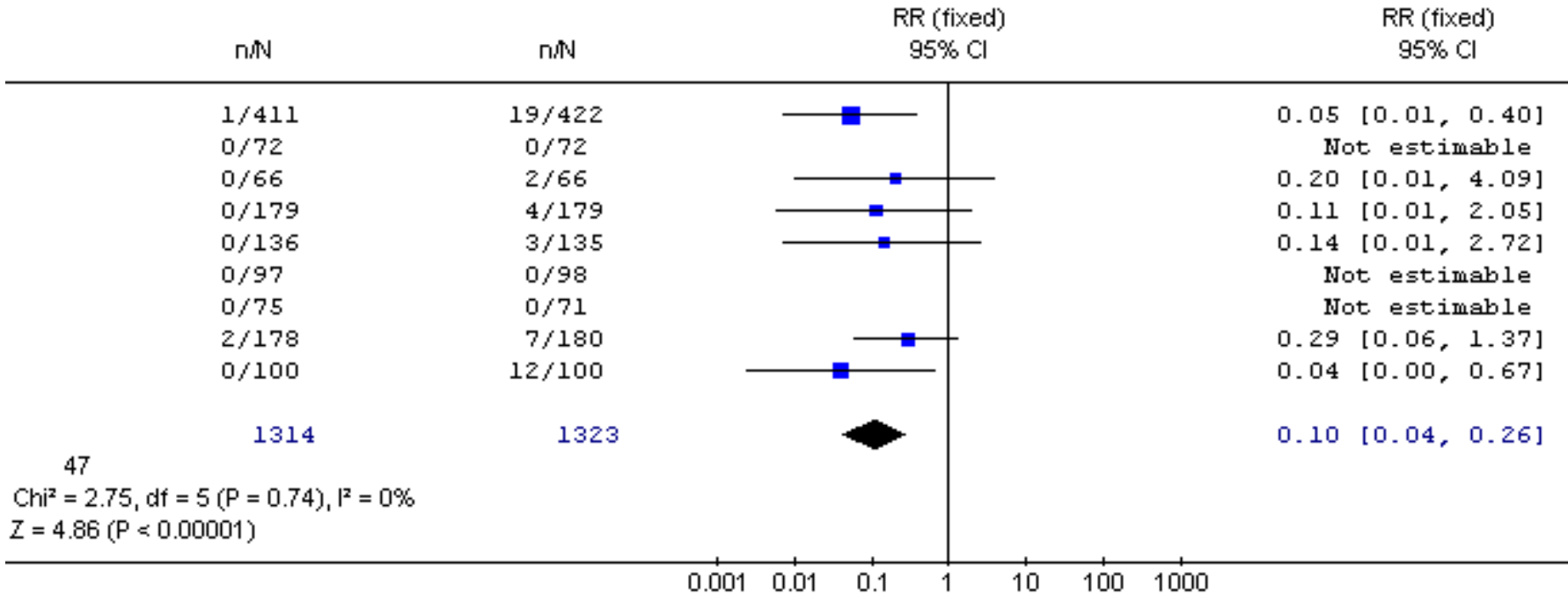
Reasons for statistical analysis

Compare and contrast the findings of each study

Combine the findings of each study

- Increase power and precision (through pooling)
- Test statistical significance
- Provide a single “average” value
- Conduct subgroup analyses
- Conduct sensitivity analyses
- Assess for study-level bias (e.g. funnel plots)

Example of a meta-analysis



3

Subgroups and heterogeneity

Subgroup analyses - planning

- Keep the number of subgroup analyses as low as possible, to reduce the possibility of false positive results (a video of Mike Clarke's lecture on this is available at https://training.cochrane.org/resource/mulitiplicity-and-subgroup-analysis-beware)
- Make sure that the ones you do are important
- Make sure you extract the relevant data
- Will you have enough power in a specific subgroup? Is it too uncommon or unlikely to have been studied?
- Will subgroups be reported by the original researchers, and in the way you need them? What will you do if they are not reported in the way you want?
- Do you have independent evidence for your predictions on what the subgroups will show (rather than evidence from the studies that would be included in the analysis)?

Subgroup analyses - interpretation

- The most statistically powerful estimate for an individual patient is likely to come from the full meta-analysis. But, this will not be meaningful if there is too much clinical heterogeneity in the meta-analysis.
- Will you do subgroup analyses to show that the results are similar for different people, or to identify people, settings or interventions where the results are different?
- Subgroup results might provide a more specific answer for an individual patient, but they suffer from being
  - Underpowered
  - More subject to the effects of chance

What subgroups would you use for

Does "five a day" improve people's oral health?

Heterogeneity – planning for it

- Do you want heterogeneity in your review?
- Heterogeneity can occur throughout the review (interventions, participants, outcome measures and study designs)
- In your review, might heterogeneity lead to
  - Variation
  - Diversity
  - Complications
  - Messiness
  - Wide coverage
  - Ability to make comparisons
  - Contradictory findings?
- Which of these would you like to have in your review?

Heterogeneity – measuring and explaining it

- How sure are you about the factors that might drive heterogeneity in the results of the studies?
- The statistical tests measure heterogeneity in the *results* of the study, not in their *content* or *conduct*.
- Might statistical heterogeneity be due to more than one factor or due to a different factor to the one you are thinking of?

Dealing with heterogeneity

- Ignore
- Eligibility
- Descriptive
- Subgroup analyses (perhaps using individual participant data)
- Sensitivity analysis
- Separate meta-analyses
- Assess it statistically (e.g. I^2 or Chi-squared)

What heterogeneity would you expect for

How long do dental implants survive?

Criteria for a ‘good’ systematic review (1)

- Does it address a focused question?
- Are the inclusion criteria appropriate?
- Is it unlikely that important, relevant studies were missed?
- Was the validity of the included studies appraised?
- Were assessments of studies reproducible?
- Are the results similar from study to study?

Criteria for a ‘good’ systematic review (2)

- What are the overall results of the review?
- How precise are the results?
- Can the results be applied to my care, or that of my patients?
- Were all the important outcomes considered?
- Are the benefits worth the harms and costs?

Selected reading

General textbooks / guidance on doing reviews

- Higgins JPT, et al (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.5 (updated August 2024). (https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current)
- Centre for Reviews and Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. (www.york.ac.uk/inst/crd)

Systematic reviews (history)

- Clarke M. History of evidence synthesis to assess treatment effects: personal reflections on something that is very much alive. J R Soc Med 2016;109:154-63
- Clarke M. Partially systematic thoughts on the history of systematic reviews. Syst Rev 2018; 7: 176
- Page MJ, et al. Epidemiology and reporting characteristics of systematic reviews of biomedical research: a cross-sectional study. PLoS Med 2016;13(5):e1002028.
- Gurevitch J, et al. Meta-analysis and the science of research synthesis. Nature 2018;555:175-82.
- Andersen MZ, et al. Comparative analysis of Cochrane and non-Cochrane reviews over three decades. Syst Rev 2024;13(1):120.

Searching

- Booth A. The number needed to retrieve: a practically useful measure of information retrieval. Health Info Lib J 2006;23:229-32.
- Edwards P, et al. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. Stat Med 2002;21:1635-40.

Assessing risk of bias

- Higgins JP, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928
- Sterne JA, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
- Wang Y, et al. Development of ROBUST-RCT: Risk Of Bias instrument for Use in SysTematic reviews-for Randomised Controlled Trials. BMJ 2025;388:e081199.

Forest plots

- Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. BMJ 2001;322:1479-80.

Heterogeneity

- Higgins JPT, et al. Measuring inconsistency in meta-analyses. BMJ 2003;327:557-60.

Chance effects

- Clarke M. The true meaning of DICE: don't ignore chance effects. J R Soc Med 2021;114:575-7.

Reporting of systematic reviews

- Page MJ, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews BMJ 2021;372:n71.
- Shamseer L, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;349:g7647.

Quality assessment of systematic reviews

- Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44:1271-8.
- Shea BJ, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.

Artificial intelligence

- Thomas J, et al. Responsible use of AI in evidence SynthEsis (RAISE): recommendations and guidance. OSF. https://osf.io/fwaud/overview

Selected software

Systematic review software
Comprehensive meta-analysis: www.meta-analysis.com
Covidence: www.covidence.org
DistillerSR: www.evidencepartners.com
EPPI reviewer: eppi.ioe.ac.uk/CMS/Default.aspx?alias=eppi.ioe.ac.uk/cms/er4&
Rayyan: rayyan.qcri.org
RevMan: revman.cochrane.org/info

Evidence Review Accelerator (TERA)
bond.edu.au/iebh/systematic-review-accelerator-sra

Website with a general search for relevant tools
systematicreviewtools.com

List of bibliographic databases
training.cochrane.org/handbook/current/chapter-04-appendix-resources

Critical appraisal tools
CASP: casp-uk.net/casp-tools-checklists/
JBI: jbi.global/critical-appraisal-tools

Right Review website
rightreview.knowledgetranslation.net

Selected AI and automation tools

Searching: WordFreq, Search refiner, Polyglot (from TERA), Elicit, ResearchRabbit

Screening: Covidence, DistillerSR, EPPI-Reviewer, Rayyan, SyRF

Assessing eligibility: ASReview, EPPI-Reviewer, DistillerSR, Rayyan, Research Screener, SWIFT-Active Screener, SWIFT-Review

Quality assessment: RobotReviewer

Data extraction: RobotReviewer