University of Dundee

Scottish Graduate School of Social Science
Sgoil Cheumnaichean Saidheans Sòisealta na h-Alba

# An Overview of Systematic Reviews

Calum MacGillivray, PhD Student (Education), University of Dundee

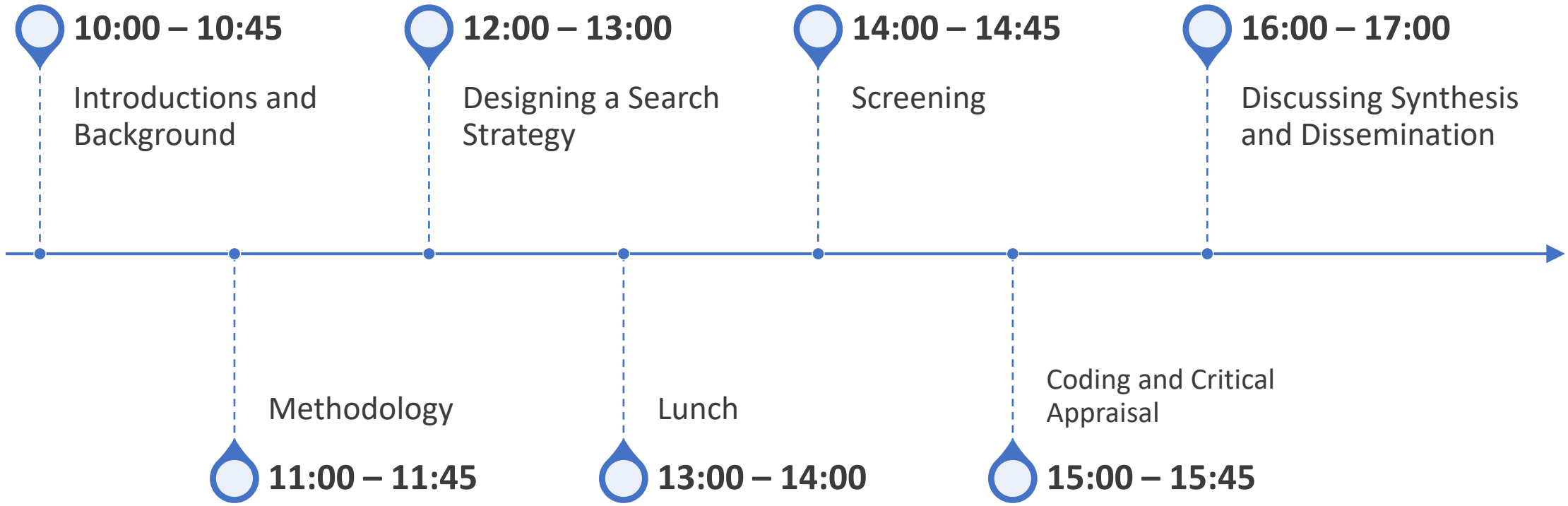
c.y.macgillivray@dundee.ac.uk | University Profile | ORCiD

# Housekeeping

- Timings: planned conservatively – we may finish early!
- Breaks: a short break every hour and an hour for lunch
- Questions: there will be time for questions - but jump in whenever
- Afterwards: feel free to email me with any questions
- Resources for this workshop can be found here:

  https://calummacgillivray.github.io/Training/

# Plan for Today



# Session 1: Introductions and Background

# Introductions

Who am I?

**Name**: Calum MacGillivray

**University**: University of Dundee, School of Humanities, Social Science and Law

**Discipline**: Education (with a developmental psychology background)

**PhD Project**: Primary-secondary transitions experiences and associated educational outcomes

**Experience with systematic reviews**: Have read plenty! Have conducted systematic reviews with meta-analysis, narrative synthesis as well as a review of reviews.

Now who are you?

What's your name? | Your discipline/background? | What are you working on? | Do you have any experience with systematic reviews?

# What do you know already?

- What is a systematic review?
- Why might you conduct a systematic review?
- Are you aware of any methods for conducting systematic reviews?
- What do you hope to get out of today?

# Learning Outcomes

Feel more confident in reading, interpreting and conducting systematic reviews by:

- Learning how to conduct high quality systematic reviews
- Building knowledge of various methods and tools available to support the process
- Gaining practical skills in building a search, screening, coding and quality appraisal
- Being aware of different methods of synthesising findings

# A Note on Terminology

- Literature Review
- Systematic Review
- Rapid Review
- Review of Reviews/Umbrella Review
- Mapping review
- Scoping Review
- Meta-synthesis
  - Meta analysis
  - Narrative synthesis

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

# Systematic Reviews

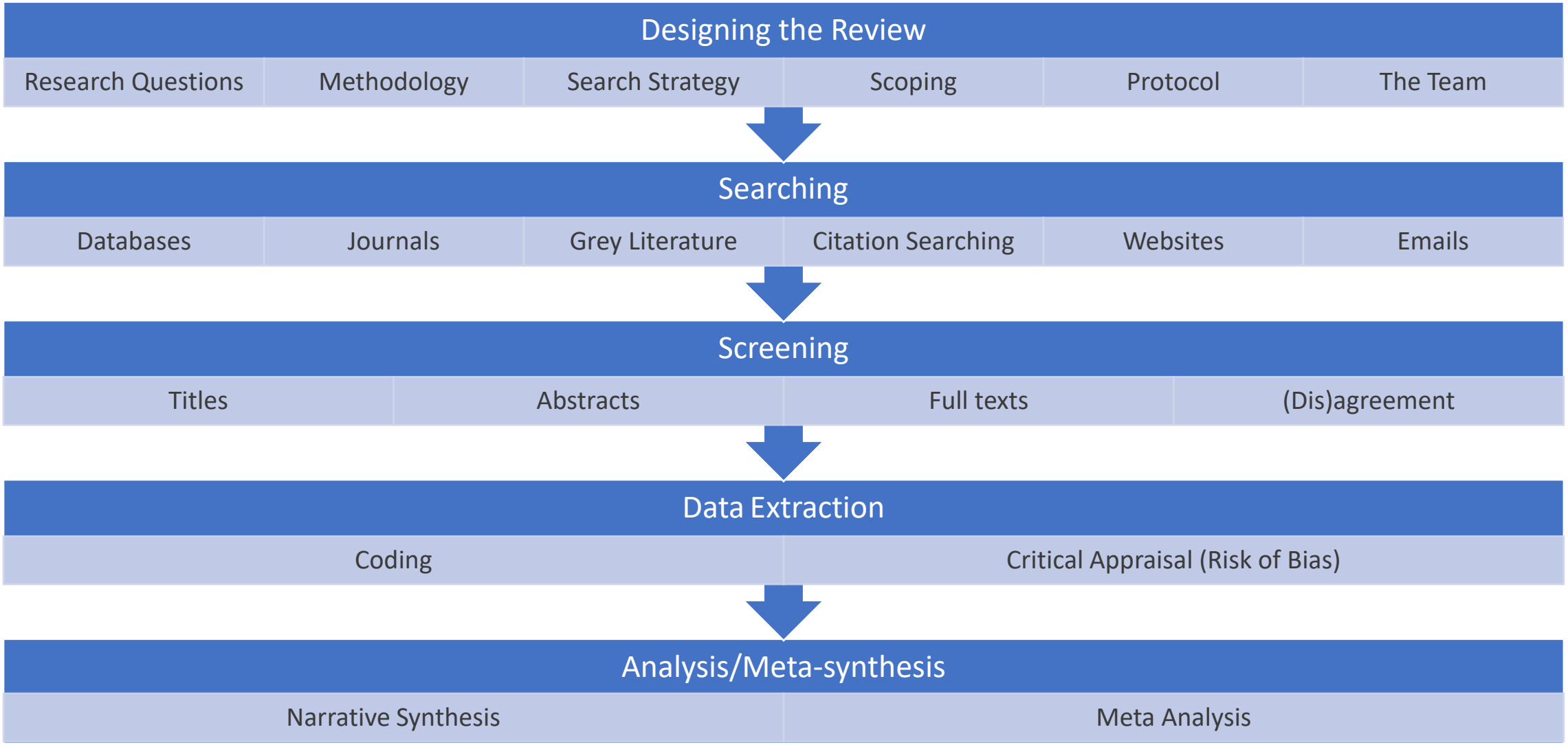
Aim to use systematic methods to search, appraise and synthesise evidence from research findings.

Systematically gathering sources and formally evaluating them can help to address the increased subjectivity involved in other literature reviews.

The usefulness of a systematic review is related to how transparent and therefore replicable the methods are.

Relevant References: Grant & Booth, (2009); Hodgkinson and Ford, (2014); Linnenluecke et al., (2020); Page et al., (2021)

# An Anatomy of a Systematic review



# A brief (and incomplete) history

1975: Meta-Analysis (psychotherapy)

Archie Cochrane: evidence based medicine

Cochrane Systematic Reviews: health care intervention efficacy

Cambell Collaboration: public policy scope

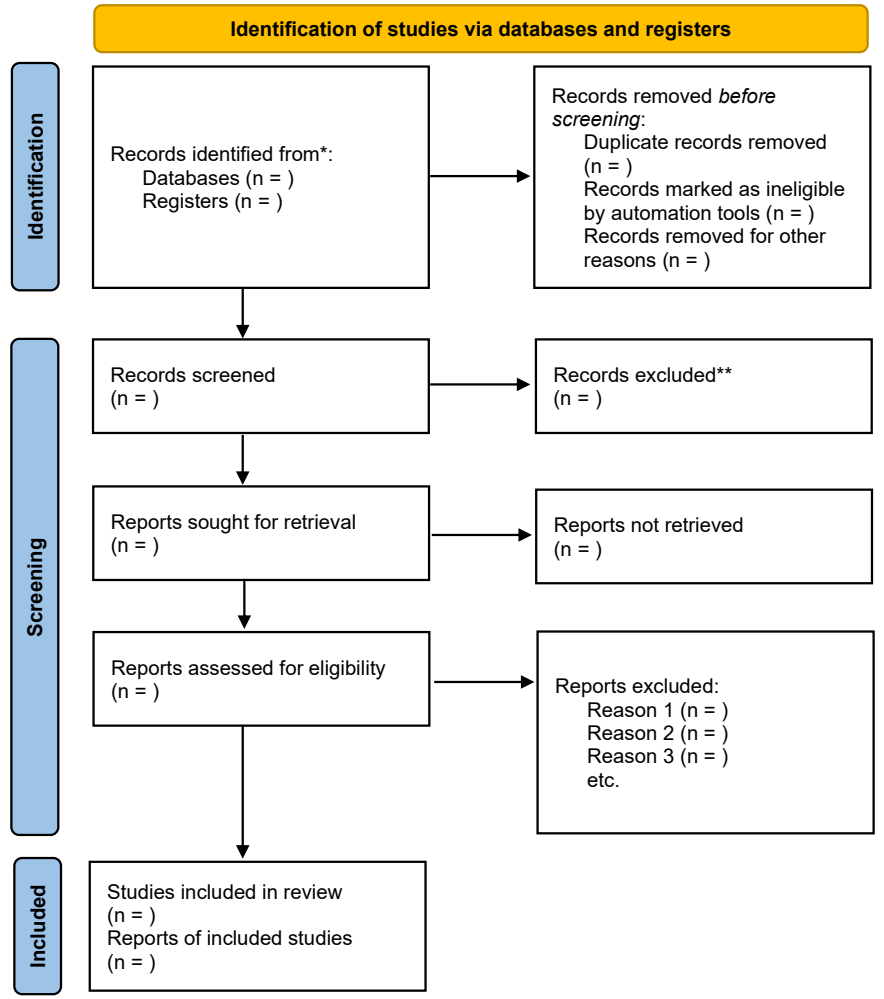
EPPI Centre: education and welfare

QUOROM, then PRISMA: reporting guidelines

https://eppi.ioe.ac.uk/cms/Resources/EvidenceInformedPolicyandPractice/HistoryofSystematicReviews/tabid/68/Default.aspx

# PRISMA Flow Diagram

- Transparent Method of Reporting
- Records identified
- Records excluded at each stage
- Reasons for exclusion
- Studies finally included




*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

https://www.prisma-statement.org/prisma-2020-flow-diagram

Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

# Why conduct a systematic review?

- Following the ‘recipe.’
  - A structured approach
  - Reproducible (when transparent)
- Reduce subjectivity – although not always the goal
- Get a good view of the available literature
  - Find gaps
- Assess the quality of methodology
- Assess the strength of evidence across the literature

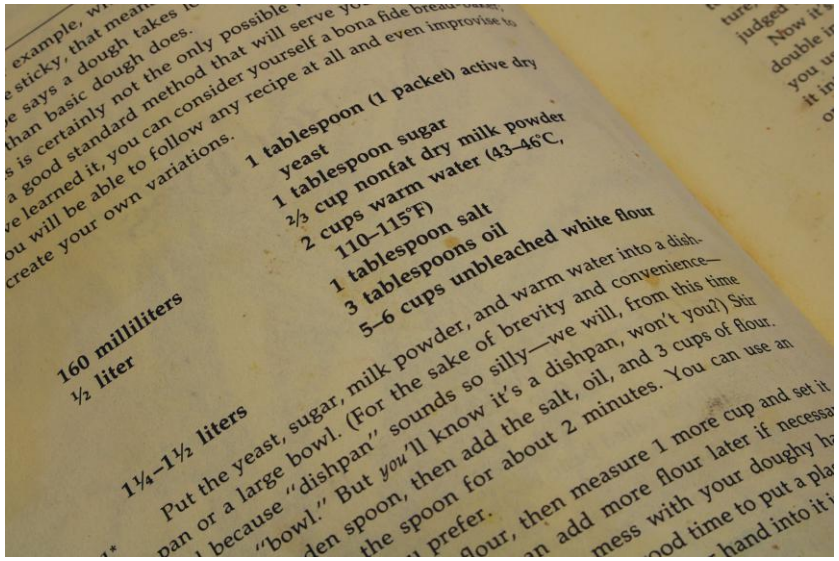



Photo by James Lee on Unsplash

# Addressing the Replication Crisis

2011 – a bad year for psychology

- Precognition – dubiously published
- Diedrik Stapel – fabricating data
- Using common methods to argue that listening to a song could reduce a participants age – to prove a point
- The name-letter effect – failed to replicate
- Many studies suffer from large amounts of unrecognised bias

The result: “Crisis of Confidence”

But this is not limited to psychology

- Dubious practice exists across scientific disciplines

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, *39*(4), 202–217. https://doi.org/10.1037/teo0000137

Pashler, H., & Wagenmakers, E. (2012). Editors’ Introduction to the Special Section on Replicability in Psychological Science: A Crisis of Confidence? *Perspectives on Psychological Science*, *7*(6), 528-530. https://doi.org/10.1177/1745691612465253

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*, 452–454. https://doi.org/10.1038/533452a

# Bias

- Bias from incorrectly applied (or understood) methodology
  - Failure to address confounding factors
  - Faulty metrics
  - Participant selection
  - Missing data
  - Misinterpretation
  - Congruity between philosophical, theoretical, and methodological approach
- Publication bias (the file drawer problem)

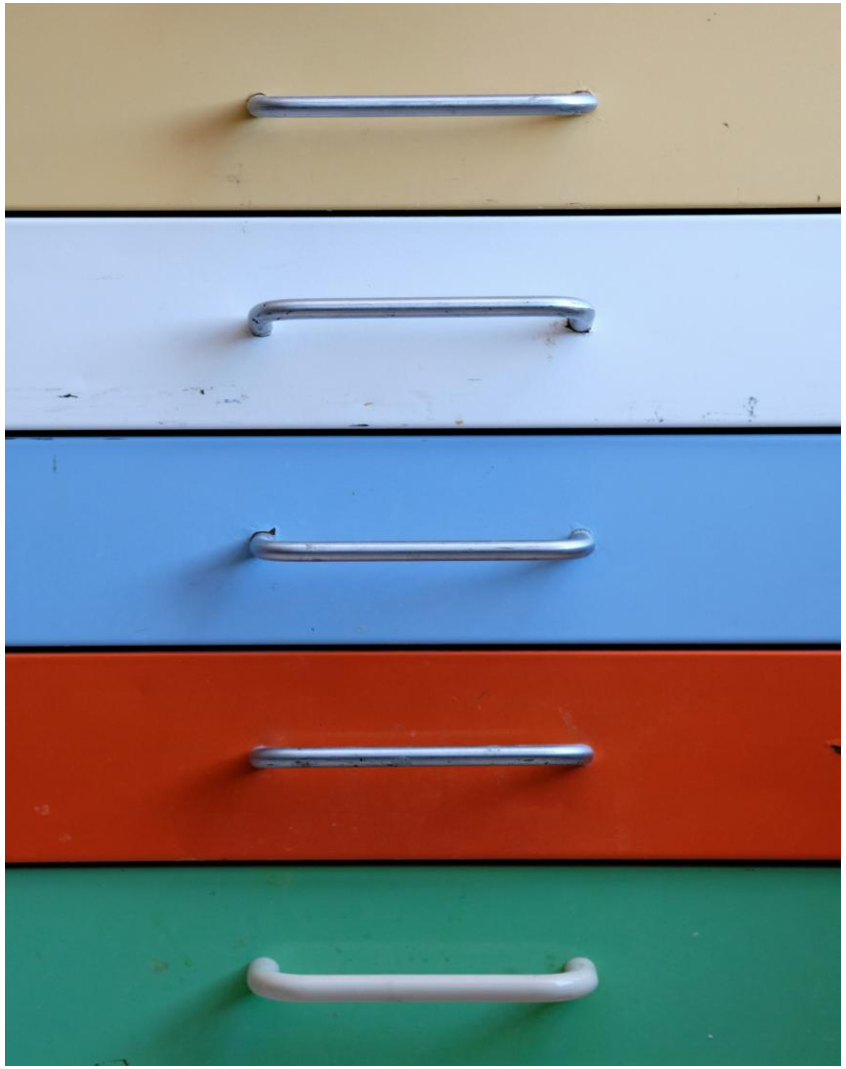


Photo by Drew Beamer on Unsplash

# Addressing the Replication Crisis

- Systematic reviews pre-date the replication crisis
  - Rubbish in > Rubbish out
- But can begin to address the file drawer problem
- Can also make informed decisions on what to include
  - Exclusion
  - Weighting
- Can be a tool for systematic, rigorous critique
- Can highlight the methodological shortcomings/strengths in the literature
- Can assess risk of bias

Photo by Evan Demicoli on Unsplash

# Identifying Gaps in the Literature

- The aim: capture the majority of the literature on the research topic
    - related to inclusion/exclusion criteria
- Overview – can see gaps for future studies
- Weaknesses worth addressing
- The place for replication studies
- What methods are used?
    - Is there one dominant strand?
    - Is that telling the full story?

# Assessing Interventions

- A classic – with a basis in healthcare
  - But applicable to other fields
- Gathering all known studies on an intervention
- Assessing the quality of those studies
  - And risk of bias
- Bringing together their findings (synthesis)
  - Potentially weighting or excluding some
    - Based on predefined criteria
- But beware the file drawer problem!

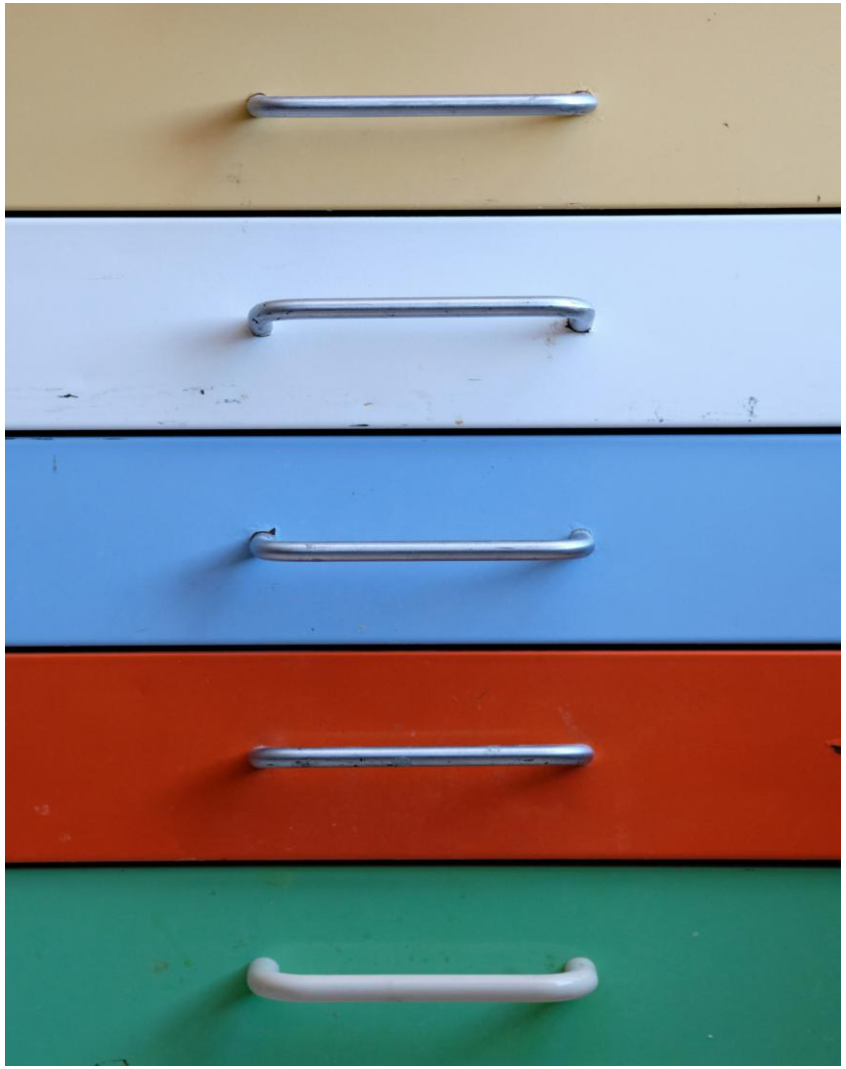


Photo by Drew Beamer on Unsplash

# Synthesis of the literature

- Bringing together findings
- Creating a new perspective
- Can cover any methodology
  - Quantitative
  - Qualitative
  - Mixed Methods
- Narrative Synthesis
- Meta Analysis
- Carefully consider what you are bringing together
- Don’t accidentally equate apples with oranges!

Photo by Gowtham AGM on Unsplash

# Building Evidence for Policy and Practice

- The evidence is brought together
  - Then assessed
  - Then synthesised
  - Rigorously and systematically
- This can provide credibility
  - Transparency/reproducibility helps with this
- Can make policy/practice recommendations (when appropriate/congruous)
  - However, do consider the views of the people that are the topic of your review
  - Some reviews are codesigned
    - Many involve checking with those affected
    - This can be ongoing

# A note on AI

- Some researchers are now using AI tools for every step of the review.
  - There must be caution here.
- A key strength of the systematic review comes from its transparency and replicability
  - AI can obfuscate these processes, especially when based on probability.
- If you do use these methods you must be transparent about when and how you used them.
- Be aware of the suitability and what is being lost.
- Your judgement is an important tool.

Responsible AI in Evidence Synthesis (RAISE): guidance and recommendations: https://osf.io/fwaud/

# Any questions?

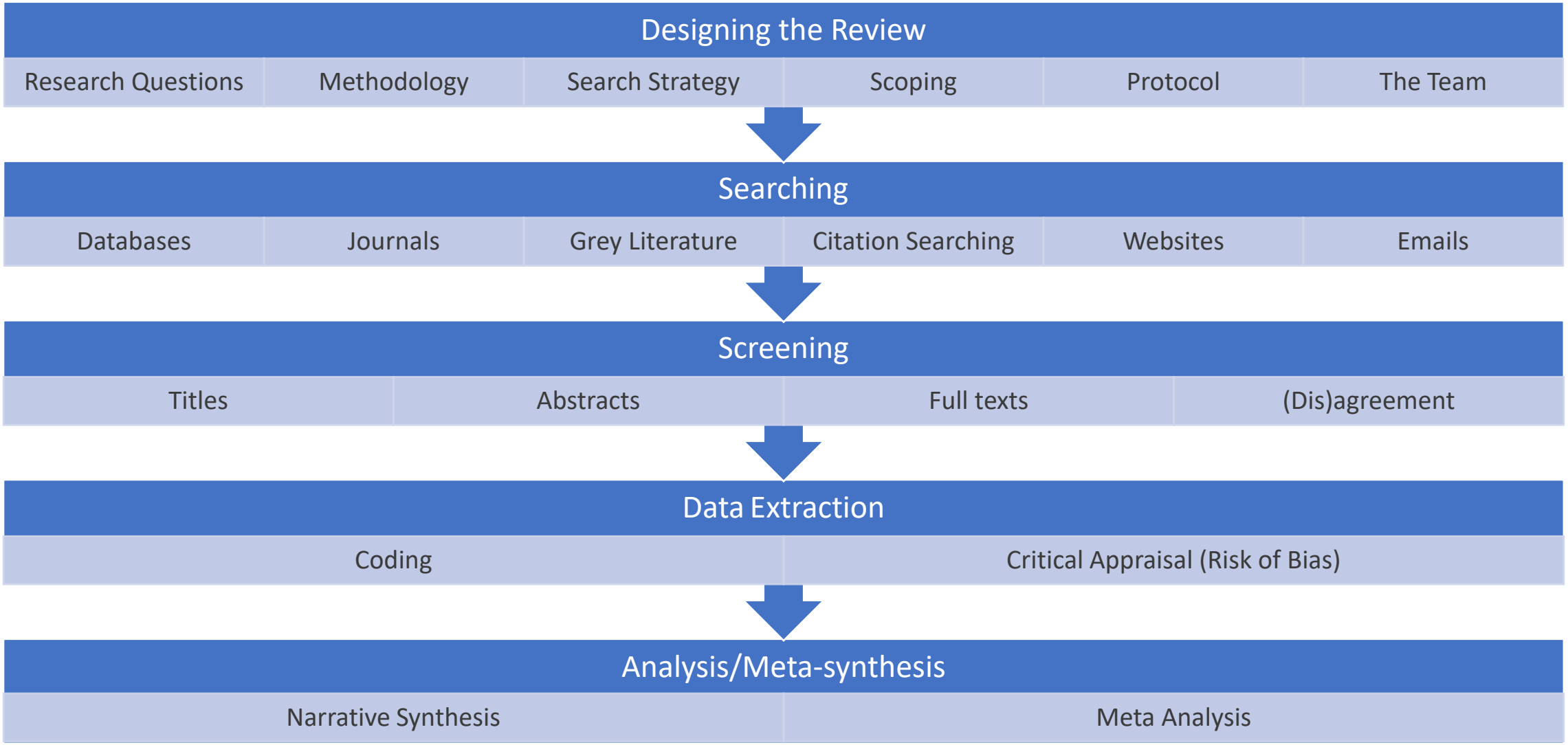
# 10-15 minute break

# Up Next: Methodology

# Session 2: Methodology

Is anyone aware of any methods used in systematic reviews already?

# The Anatomy of a Systematic review



# Frameworks/approaches

- Useful in guiding the development of the protocol
  - Methods often designed for specific purposes
- Examples
  - Eppi-Centre
  - JBI
  - Campbell
- Worth looking into several to make a decision on best fit
  - Don’t let the methodological tail wag the dog!
- Approaches differ but most systematic reviews share some elements
  - Searching
  - Screening
  - Extraction
  - Appraisal
  - Synthesis

Photo by Kostiantyn Li on Unsplash

# Protocols

- Protocols are a useful way to enhance transparency
  - They also help you
  - Like a recipe you follow
  - This can help with planning/organisation
- Protocols do not have to be restrictive
  - Rather they help to show when you changed an approach
  - It is important to communicate this – why?
- It is helpful if there is a record of your protocol – aids transparency
  - Journals
  - Repositories – e.g. PROSPERO and OSF
  - Can be embargoed
- Guidelines exist which can help – e.g. PRISMA-P

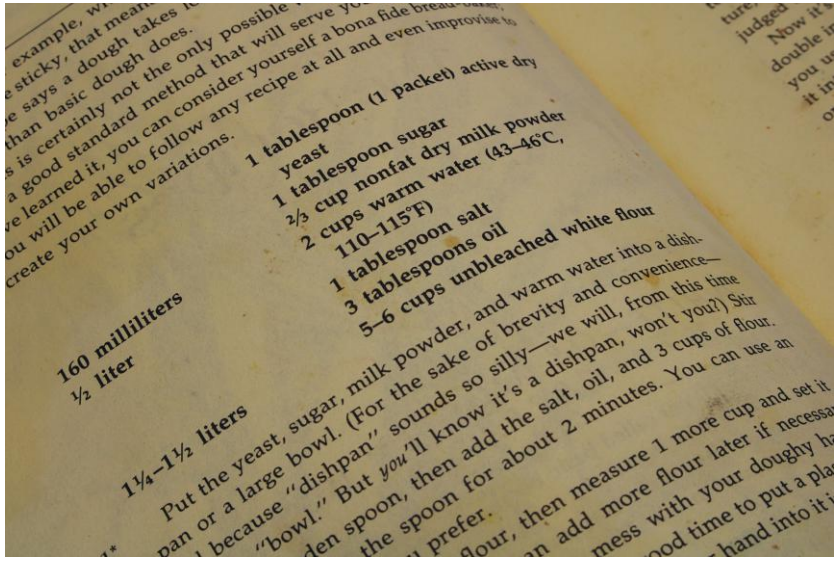



Photo by James Lee on Unsplash

https://www.prisma-statement.org/protocols

# Designing the Review

- It is usually helpful to design tools in advance
  - Can use specialist tools – will cover some later
  - Often simple spreadsheets will suffice to keep records
    - But design these mindfully
      - Will others be able to use these independently?
  - Comprehensive and precise record keeping is highly important
    - This is part of being transparent/replicable
- Consider carefully your review questions
  - Then decide in advance what you will keep and what you will not
    - Inclusion/exclusion criteria

# Searching

- Choose multiple databases – one is rarely enough
  - There are studies on optimising this – but use your best judgement
    - E.g. Bramer et al., (2017)
      - From a medical perspective (recall percentage based on search using all databases)
        - Embase (85.9%), Embase and MEDLINE (92.8%), Embase, MEDLINE and Web of Science (95.9%), Embase, MEDLINE, Web of Science, and Google Scholar (98.3%).
          - But in specific circumstances PSYCINFO was important
      - Choose multiple databases to cover blindspots
        - Think about your specific topic
      - Be careful with search engines like google scholar as algorithms/obfuscation may make searches difficult to replicate
  - Targeted journal searching and citation searching can be useful

Bramer, W. M., Rethlefsen, M. L., Kleijnen, J., & Franco, O. H. (2017). Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. *Systematic reviews*, *6*, 1-12. https://doi.org/10.1186/s13643-017-0644-y

# Searching

- Most reviews look at grey literature – at least to investigate the problem of publication bias
  - Grey literature databases
  - Websites
  - Contacting researchers
  - Registrations/protocols
  - Conference abstracts
- Can also include grey literature in synthesis

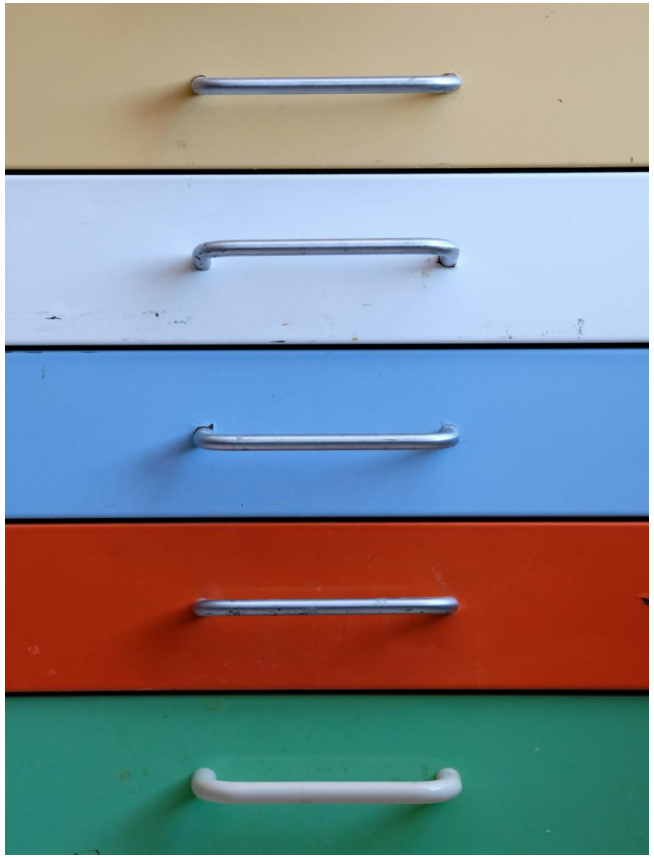


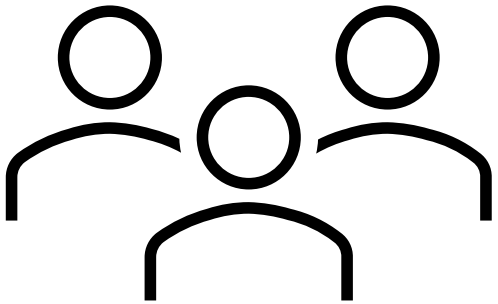
Photo by Drew Beamer on Unsplash

# Searching

- Design your search strategy to maximise relevant records while minimising irrelevant records
  - This is tricky and can take some time to get right
  - Scoping searches
    - Iterative design
    - Eventually you will have your final search terms
- We will discuss Boolean operators for search terms later
- Keep record of numbers – to complete a PRISMA diagram
- Referencing software is well suited to this task
  - Sub folders can be very helpful

# Working with others

- The process can be made more rigorous by ensuring no one researcher is solely responsible
  - Reduces risk of bias – from an individual
- Ideally – 2 or more researchers undertake processes
  - Disputes are settled by another researcher
  - This can be done blindly
- This can be limited by practical limitations – be transparent
- This is helpful at various stages:
  - Screening
  - Coding
  - Appraisal



# Screening

- De-duplication often happens first
  - Software can help to flag potential duplicates – but be careful!
- Can then screen by title/abstract
  - Sometimes just titles first, then abstracts
  - At this stage only records that are obviously not suitable is done
    - Erring on the conservative side
- Then full texts are screened
  - Reasons are provided for exclusions at this stage
- Referencing software can be used – tools like Covidence and Eppi-Reviewer also facilitate screening

# Data Extraction

- Data extraction usually uses predesigned coding spreadsheets
  - But tools like Covidence and Eppi-reviewer are designed for this too
- Usually key information
  - E.g. bibliographic information, population information, sample size, research design, sampling, analysis methods, items relevant to research questions, etc.
  - Effect sizes, Standard errors, or information to calculate effect sizes if conducting meta-analysis

# Critical Appraisal

- Critical appraisal, quality appraisal, risk of bias assessment
  - Various methods are used to assess quality of included studies
  - Important to address rubbish in > rubbish out
    - Although poor quality research is a finding itself
    - Cut offs are often decided in advance – although judgement will be required
- We will cover various tools later to do this
  - It is worth carefully considering which tool best suits your review
- Sometimes there is a reason to include low quality research
  - E.g. you are interested in methods being used
  - But if findings are synthesised you may be placing equal weight on two unequal studies

# Analysis/Synthesis

- We will spend some time discussing analysis/synthesis at the end
- This tends to take a qualitative approach, quantitative approach, or mixed methods approach
- Evidence synthesis is the combination of information from across studies.
    - Meta-synthesis, narrative synthesis, meta-analysis, etc.

# What Next?

- Many studies check with the impacted population (or have codesigned the review from the start)
    - Patient and Public Involvement (PPI)
    - Participatory approaches
    - Realist reviews
        - Iterative negotiation
- Consider how best to communicate findings to target audience

# Any questions?

# 10-15 minute break

# Up Next: Search Strategies

# Session 3: Search Strategy

# Considerations when designing a search strategy

- Try to capture as many relevant records as possible, while minimising irrelevant records
- Screening is a labour-intensive process
- The less hits, the quicker it will be
  - But the more that could be potentially missed

| PICOS | SPIDER |
| --- | --- |
| • Population | • Sample |
| • Intervention | • Phenomenon of Interest |
| • Comparison | • Design |
| • Outcome | • Evaluation |
| • Study Design | • Research type |

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC health services research*, *14*, 579. https://doi.org/10.1186/s12913-014-0579-0

# Search Strategy Group Task – Part 1

In groups of 4-5 – choose 1 person to keep notes

1. Pick a review topic: from the list on the website
2. Pick key terms: to maximise relevant results while minimising irrelevant results – consider PICOS or SPIDER etc.
3. Find relevant databases: some discipline specific suggestion can be found on the website to get you started
4. Discuss relevant journals & other places where you may find relevant results

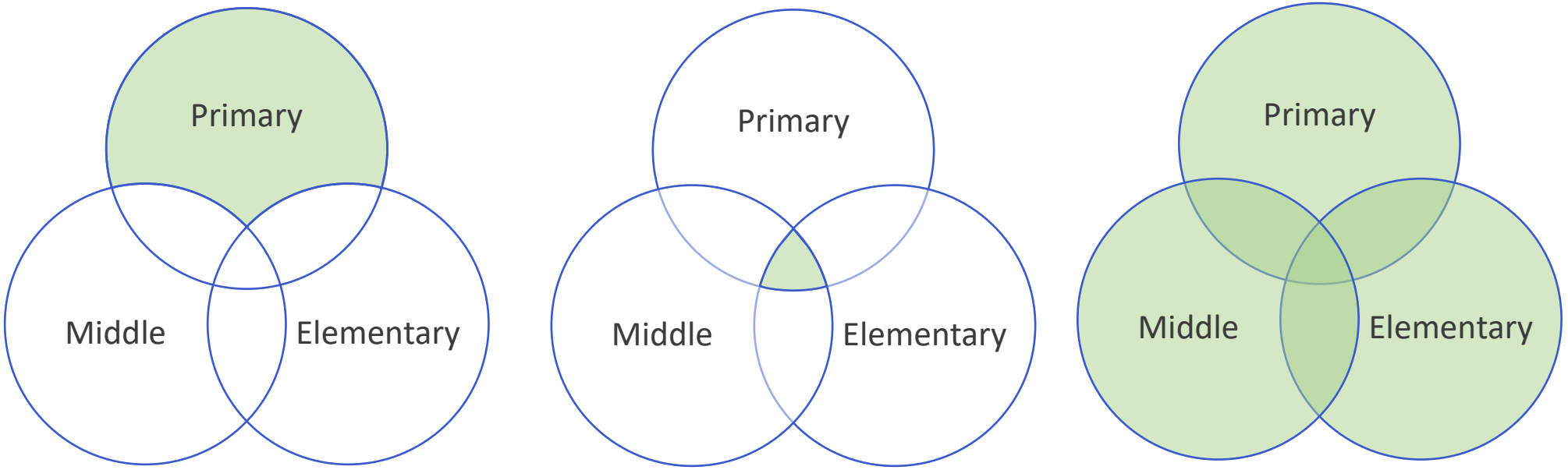
Website address: https://calummacgillivray.github.io/Training/

# Search Strategy Group Task – Part 2

Still in groups

- Discuss how to combine your search terms with Boolean operators
- Brackets can be used to group terms
- Asterisks can be used for wildcards e.g. Mov* = move, moving, movers, movers, moved, etc.

Primary NOT Middle NOT Elementary | Primary AND Middle AND Elementary | Primary OR Middle OR Elementary




Website address: https://calummacgillivray.github.io/Training/

# Any questions?

# Lunch

# Up Next: Screening

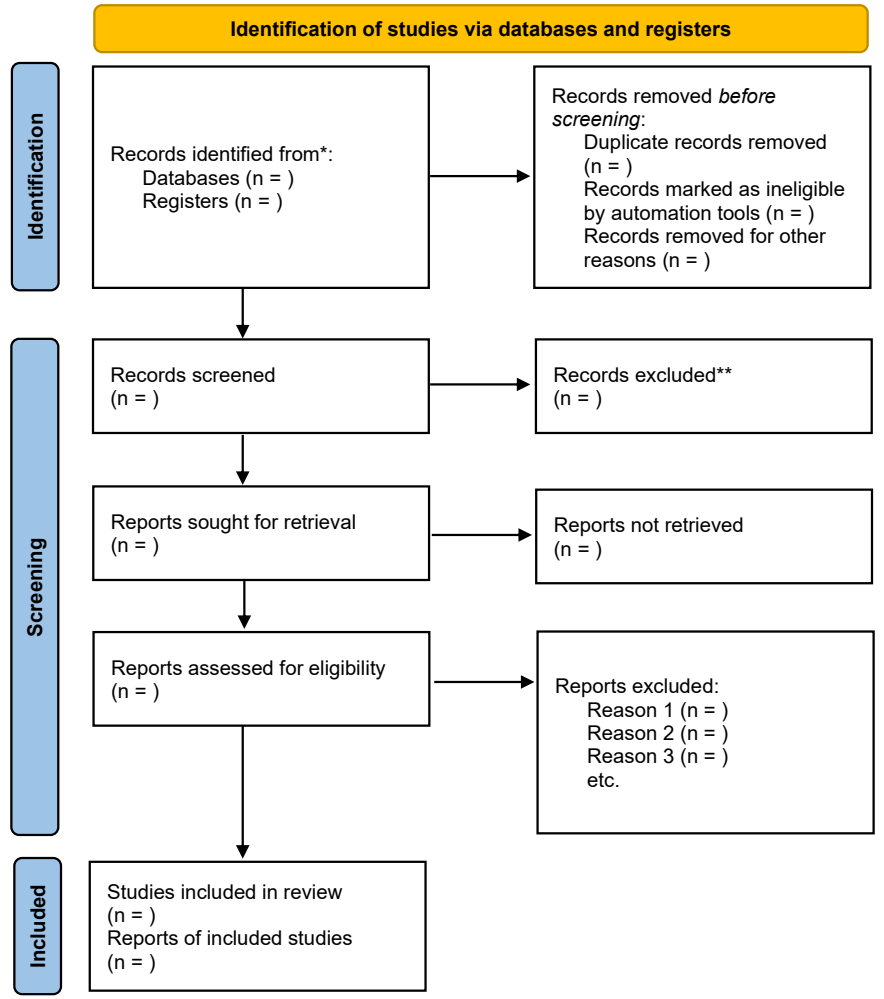
# Session 4: Screening

# Considerations For Screening

- Screening is often done with titles and abstracts together
  - In this case we will look at each level individually
- Screening is usually done with 2 or more screeners
  - They look individually and then disagreements are resolved by another
  - This can be blinded
    - The goal is to reduce the risk of bias from one screener
- Spreadsheets are handy for keeping track of this process
  - Other tools available include, but are not limited to:
    - Covidence
    - Eppi-reviewer
    - And reference management software
- Keep meticulous notes!

# PRISMA Flow Diagram

- Transparent Method of Reporting
- Records identified
- Records excluded at each stage
- Reasons for exclusion
- Studies finally included




*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

https://www.prisma-statement.org/prisma-2020-flow-diagram

Source: Page MJ, et al. BMJ 2021;372:n71. doi: 10.1136/bmj.n71.

# Screening Task – Part 1, Titles

- Individually
- From the website – download the documents called Titles and Search Criteria
- Based on this remove duplicates
- Then decide which are irrelevant to the search criteria
- Consider:
  - What is easy to remove?
  - What are you unsure about?
- Compare your decisions with others in your group
  - Do you all agree?
  - What do you disagree with, and why?

Website address: https://calummacgillivray.github.io/Training/

# Screening Task – Part 2, Abstracts

- Individually
- From the website – download the documents called Abstracts and Search Criteria
- Decide which are irrelevant to the search criteria
- Consider:
  - What is easy to remove?
  - What are you unsure about?
  - Was there anything that could have been removed previously?

Website address: https://calummacgillivray.github.io/Training/

# Screening Task – Part 3, Full Texts

- Individually
- From the website – download the documents called Full Texts and Search Criteria
- Then decide which are irrelevant to the search criteria
- Consider:
  - What is easy to remove?
  - What are you unsure about?
  - How much longer does this stage take?
  - Was there anything that could have been removed before?

Website address: https://calummacgillivray.github.io/Training/

# Any questions?

# 10-15 minute break

# Up Next: Coding and Critical Appraisal

# Section 5: Coding and Critical Appraisal

# Considerations for Data Extraction

- Data extraction can be done via spreadsheet or specialist tools.
  - E.g. Covidence; EPPI-Reviewer; JBI Sumari
- This is where you procedurally extract the information of interest from each study – to form the basis of synthesis.
  - That is not to divorce that information from context
  - But provides a basis from which to build on
- This can be done with multiple reviewers and cross-checking as in other stages
- Consider carefully what you will capture here
  - If conducting meta-analysis, consider what happens when information you want is not available

# Coding task

- Individually
  - From the website, find the coding sheet
  - Use the 3 full texts from session 4
  - Try and fill it in as best you can
- Consider
  - Is it capturing everything? What is lost?
  - What is easy to find?
  - What is hard to find?
  - Do you agree with others in your group with how to go about this?

Website address: https://calummacgillivray.github.io/Training/

# Considerations for Critical Appraisal

- Critical appraisal involves evaluating the strengths and weaknesses of studies
    - Often this is about considering the risk of bias
- Here it is very useful to use a predesigned tool
    - Some tools are made for different purposes
    - We will all try to achieve the same task with a different tool to find the best fit
- This is often done by more than one reviewer
- Decisions can be made to exclude studies based on this

# Critical Appraisal Group Task

- In groups
  - You will each take a different critical appraisal tool
  - Follow the links online
  - Choose one person to fill in the form
  - Get it in an editable state
- Consider
  - What do you agree/disagree on?
  - What about the tool works well?
  - Is the tool suitable for this task?
- We will all use this paper: https://doi.org/10.1016/j.lindif.2020.101854

# Any questions?

## 10-15 minute break

## Up Next: Synthesis and Dissemination

# Section 6: Synthesis and Dissemination

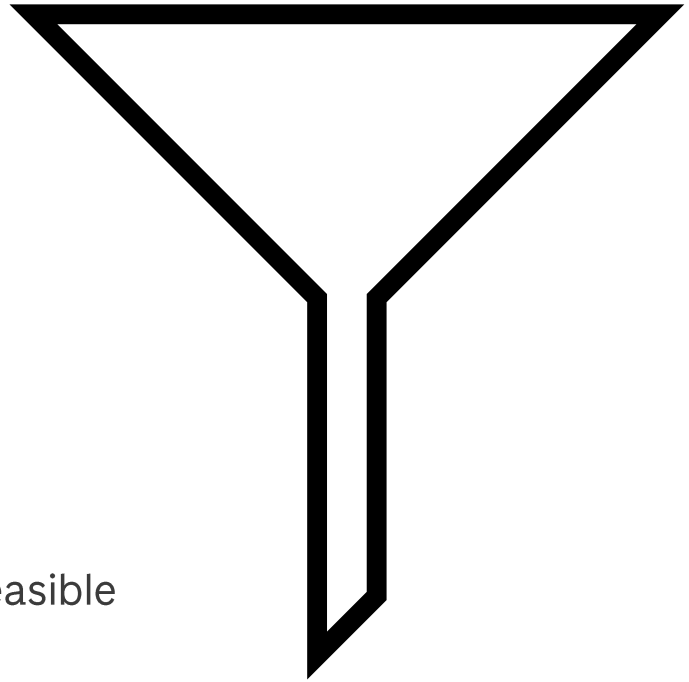
# Types of synthesis

- Meta-synthesis
  - Quantitative Synthesis
    - Meta Analysis
  - Qualitative Synthesis
    - Meta-aggregation
    - Narrative Synthesis
  - Mixed Methods (e.g. JBI Mixed Method Systematic Reviews)
    - Sequential – one after the other
    - Convergent - together
      - Integrated – quantitative and qualitative
        - Data is transformed – usually quant is qualitised
          - E.g. quantities could become declarative statements
          - It is more difficult the other way!
        - These are then synthesised in groups to address the review questions
      - Segregated – different parts of a wider topic
        - Separate qualitative and quantitative synthesis is conducted where possible
        - Then strands are compared and brought together

https://jbi-global-wiki.refined.site/space/MANUAL/355829259/8.3+The+JBI+approach+to+mixed+method+systematic+reviews

# Qualitative Synthesis

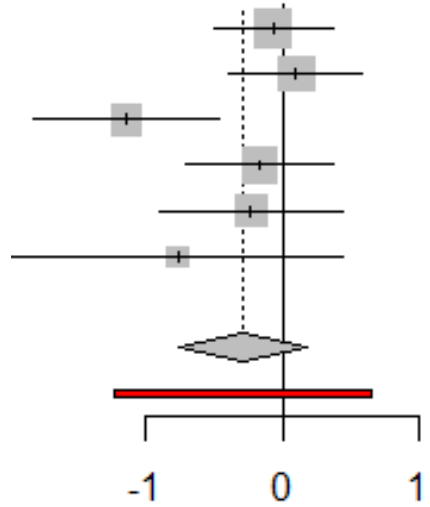
- Meta-aggregation
  - Specifically qualitative studies
  - Extracted findings are analytical statements
    - The plausibility is assessed
  - Findings are grouped – creating categories
  - Categories are grouped into synthesised findings
- Narrative Synthesis
  - Qualitative or quantitative studies
  - Summarising, and grouping findings in a narrative way
  - Often enhanced by use of tabular data
  - Drawing out themes and patterns
  - Can be the planned synthesis
    - Can also be an option when planned meta-analysis is no longer feasible
- Other options
  - Meta-ethnography
  - Realist synthesis
  - Critical interpretive synthesis



Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation: http://doi.org/10.1097/XEB.0000000000000062

# Meta Analysis

- Random effects or fixed effects
  - Fixed effects – assume a common underlying effect (higher homogeneity)
  - Random effects – assume similar but different effects, drawn from a distribution (higher heterogeneity)
- Assessing heterogeneity
  - Chi$^2$ (for independent groups) or Cochran's *Q* (repeated measures with binary outcomes)
    - to detect heterogeneity – on significant p-value test
  - Higgins and Thompson's $I^2$ – to quantify heterogeneity
  - Tau-squared – forms the basis of random effects analyses
- Effect sizes
- Pooling effect sizes
- Publication bias
- Other options:
  - Meta-regression Multilevel, SEM, Network, Bayesian




https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

# A little more on heterogeneity

- According to the Cochrane handbook:
- Chi$^2$ is prone to bias at small study numbers, or sample sizes
  - Therefore a non-significant finding does not mean hetergoeniety can be ruled out
  - At the extreme end – high numbers of studies can lead to unimportant heterogeneity being detected
  - 0.1 (rather than 0.05) is often used as the threshold for this test
- For $I^2$
  - 0% - 40%: possibly less important
  - 30% - 60%: possibly moderate
  - 50% - 90%: possibly substantial
  - 75% - 100%: possibly considerable

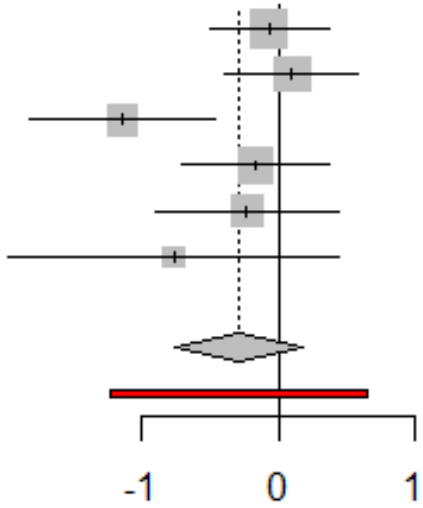
https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-10

# Forest Plots - Example Data

Simulating 6 studies comparing an outcome at T2 after an intervention for an experimental group vs a control group

| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

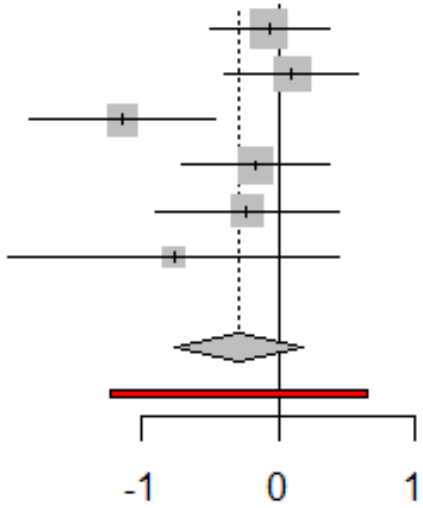
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Measures

| Study | Experimental Total | Experimental Mean | Experimental SD | Control Total | Control Mean | Control SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

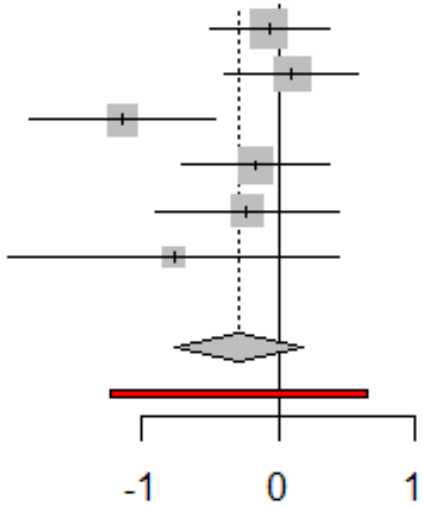
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Totals

| Study | Experimental Total | Experimental Mean | Experimental SD | Control Total | Control Mean | Control SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

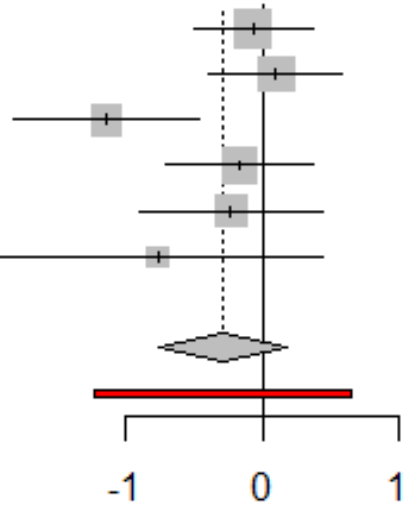
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Means and Standard Deviations

| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

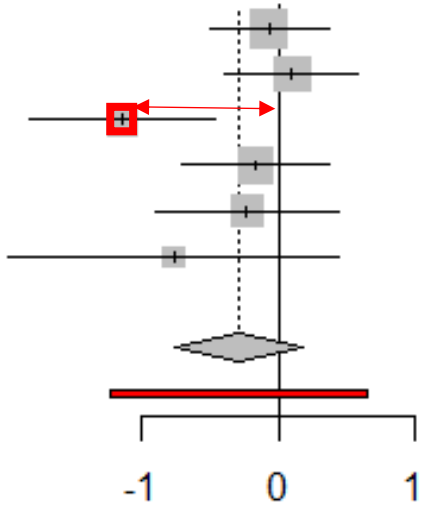
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Standardised Mean Difference

| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

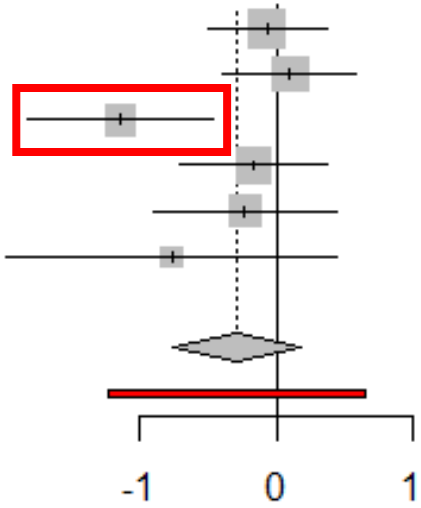
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Confidence Intervals

| Study | Experimental Total | Experimental Mean | Experimental SD | Control Total | Control Mean | Control SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

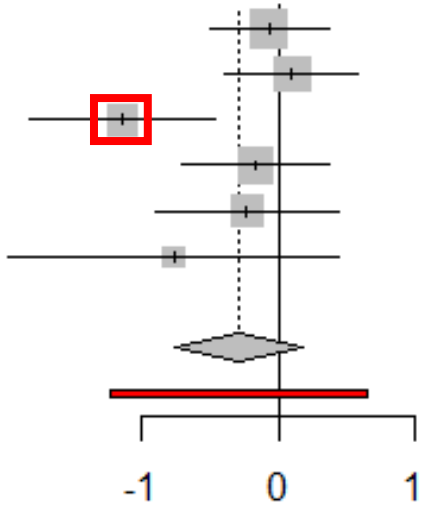
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Weighting

| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

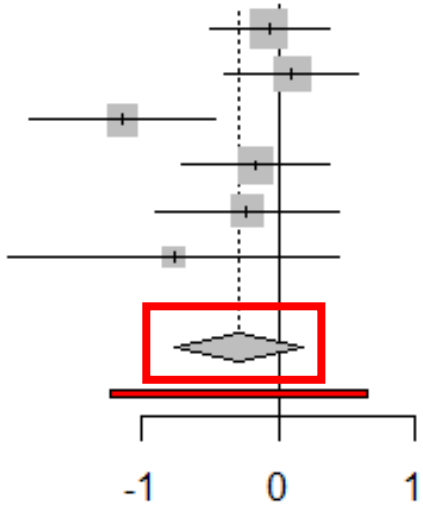
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



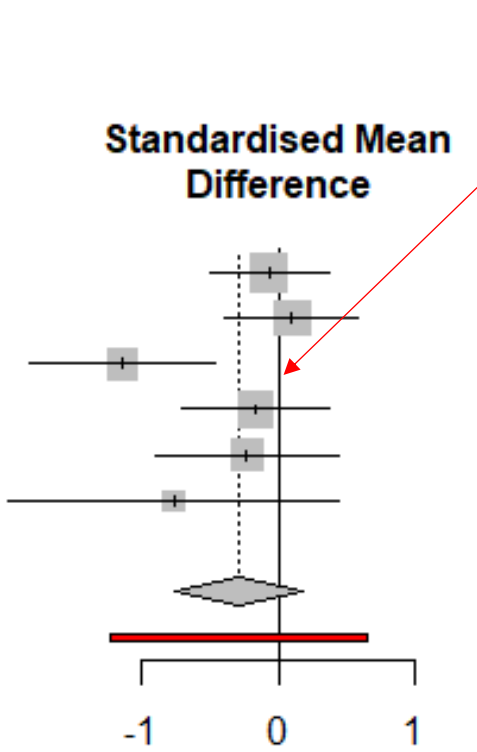
# Forest Plots – Pooled Statistics

| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764



# Forest Plots – Interpreting the Plot




| Study | Experimental Total | Mean | SD | Control Total | Mean | SD | Standardised Mean Difference | SMD | 95%-CI | Weight |
|---|---|---|---|---|---|---|---|---|---|---|
| Study 1 | 41 | 26.60 | 10.1000 | 39 | 27.30 | 10.6000 | | -0.07 | [-0.51; 0.37] | 22.4% |
| Study 2 | 34 | 4.00 | 1.9000 | 31 | 3.80 | 2.0000 | | 0.10 | [-0.39; 0.59] | 20.7% |
| Study 3 | 20 | 13.25 | 2.6300 | 19 | 16.26 | 2.5400 | | -1.14 | [-1.82; -0.46] | 15.2% |
| Study 4 | 27 | 1.95 | 1.5700 | 26 | 2.16 | 0.7600 | | -0.17 | [-0.71; 0.37] | 19.1% |
| Study 5 | 18 | 6.30 | 1.2600 | 17 | 6.58 | 1.1200 | | -0.23 | [-0.89; 0.44] | 15.6% |
| Study 6 | 7 | 3.30 | 3.4000 | 5 | 6.30 | 3.9000 | | -0.77 | [-1.97; 0.44] | 7.0% |
| **Random effects model** | **147** | | | **137** | | | | **-0.29** | **[-0.77; 0.19]** | **100.0%** |
| **Prediction interval** | | | | | | | | | **[-1.23; 0.65]** | |

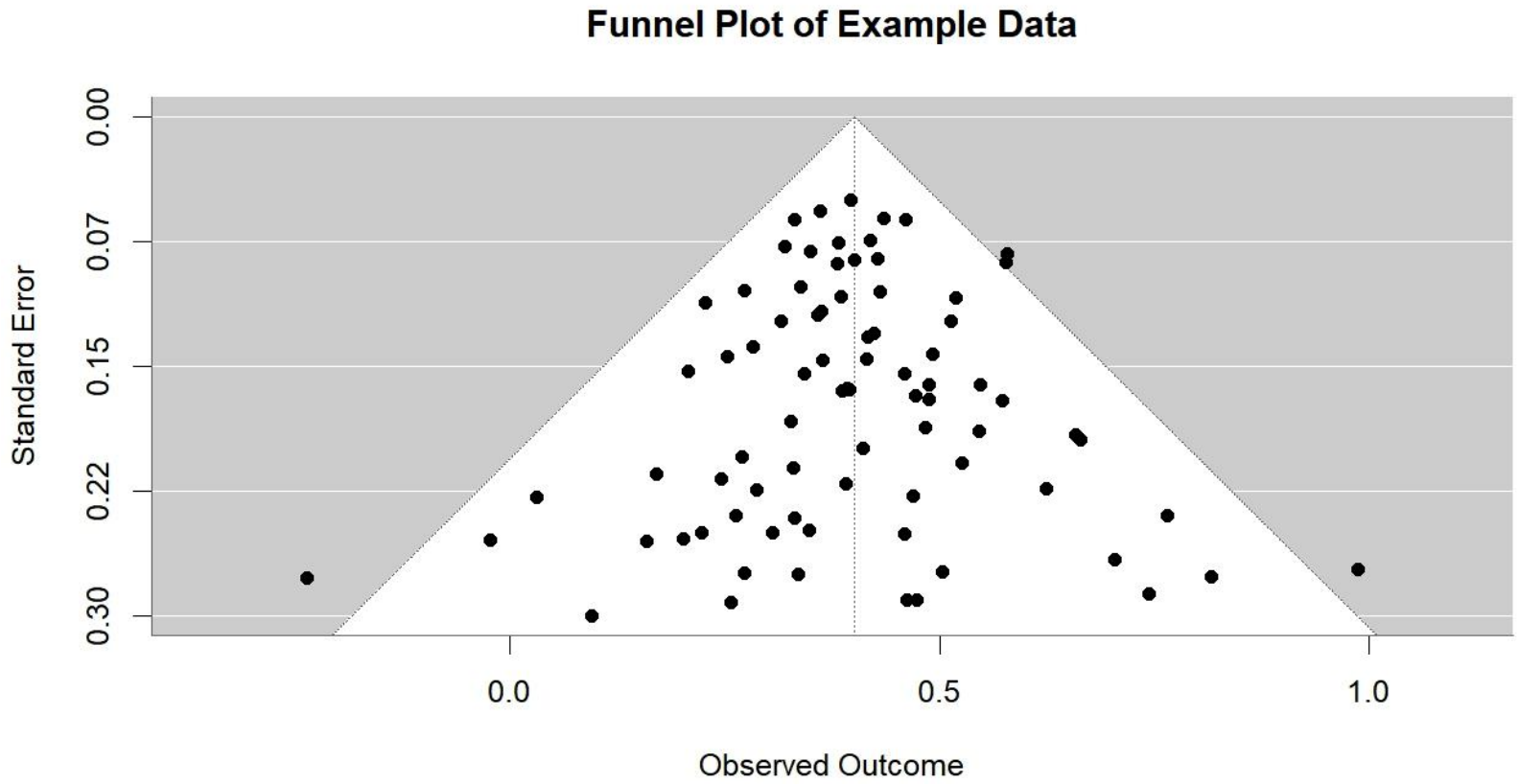
Heterogeneity: $I^2$ = 49.8%, $p$ = 0.0764

# Funnel Plots

- Larger studies are more likely to be published - More resources are used
  - Sample size is related to standard error – higher sample leads to narrower SE and wider confidence intervals
    - Smaller studies less likely to find a significant finding
    - Non-significant studies are less likely to be published
- Funnel Plots – more studies is better
  - Plotting effect sizes by Standardised mean difference and standard error
  - An exemplar expected funnel-shape in dotted lines
  - A middle line showing the average effect size
  - A symmetrical plot suggests publication bias is less likely
    - Can also look at contours related to significance

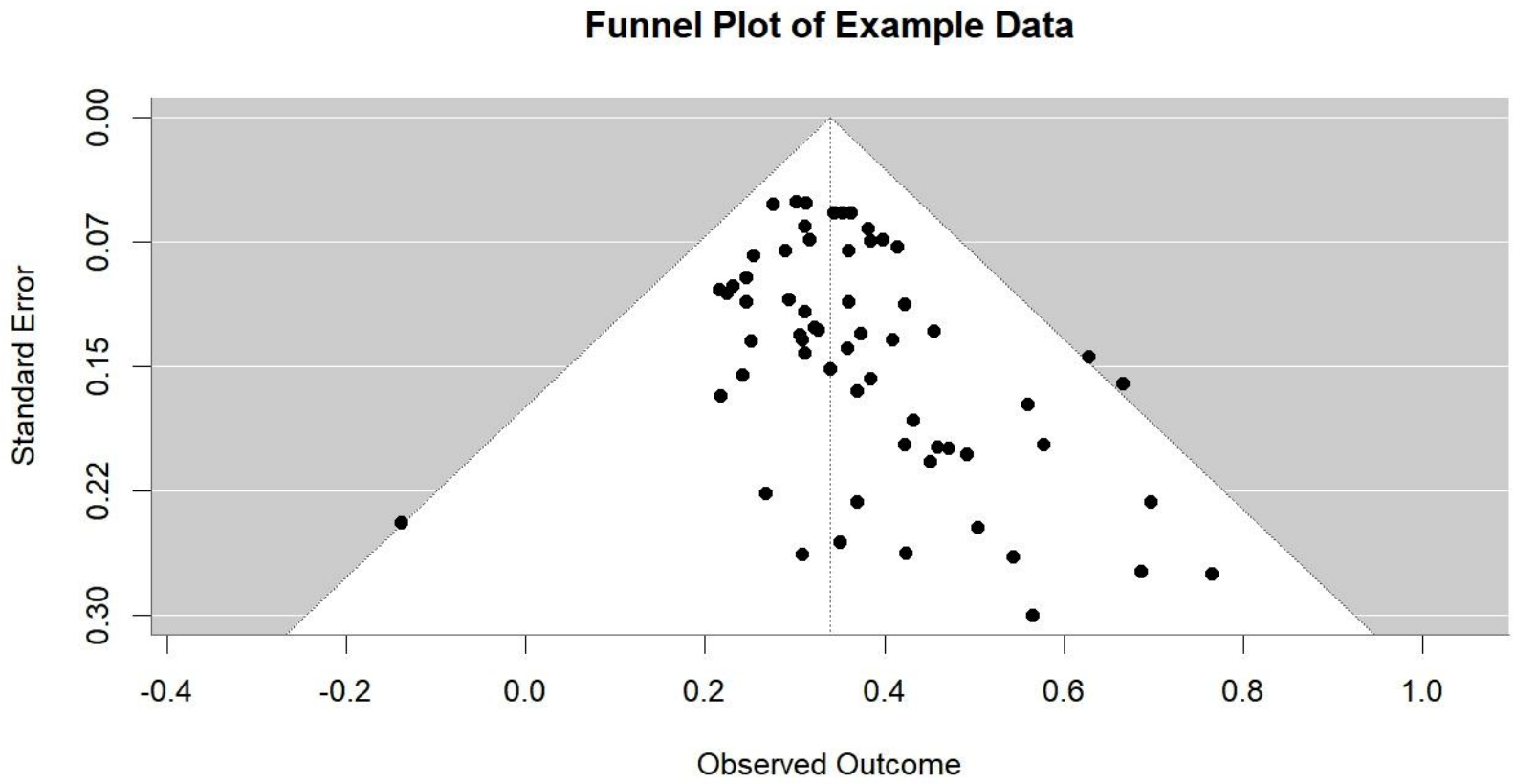
Borenstein, Michael, Larry V Hedges, Julian PT Higgins, and Hannah R Rothstein. 2011. *Introduction to Meta-Analysis*. John Wiley & Sons.

# Funnel Plots




From example data, synthesised to be relatively symmetrical

# Funnel Plots




From example data, synthesised to be less likely to be published if non-significant

# Dissemination

- So you've done the work, now what?
- A paper is good – use PRISMA report guidelines
  - So are conferences!
- But should consider how best to reach the target audience
- E.g. Infographics or comics
- Be creative!
  - And always be transparent
- Involve those that are the subject of your review
  - You could get their feedback

# Any final questions?

# Thank You!

One more thing:

A list of useful resources is available on the website:
https://calummacgillivray.github.io/Training/

And here's a handy resource from my time working in open research:

- Unpaywall, a browser extension for easily getting legal open access when available: https://unpaywall.org
  - It makes a little green padlock appear when an open access file is legally accessible in a repository somewhere

Feel free to get in touch:

c.y.macgillivray@dundee.ac.uk