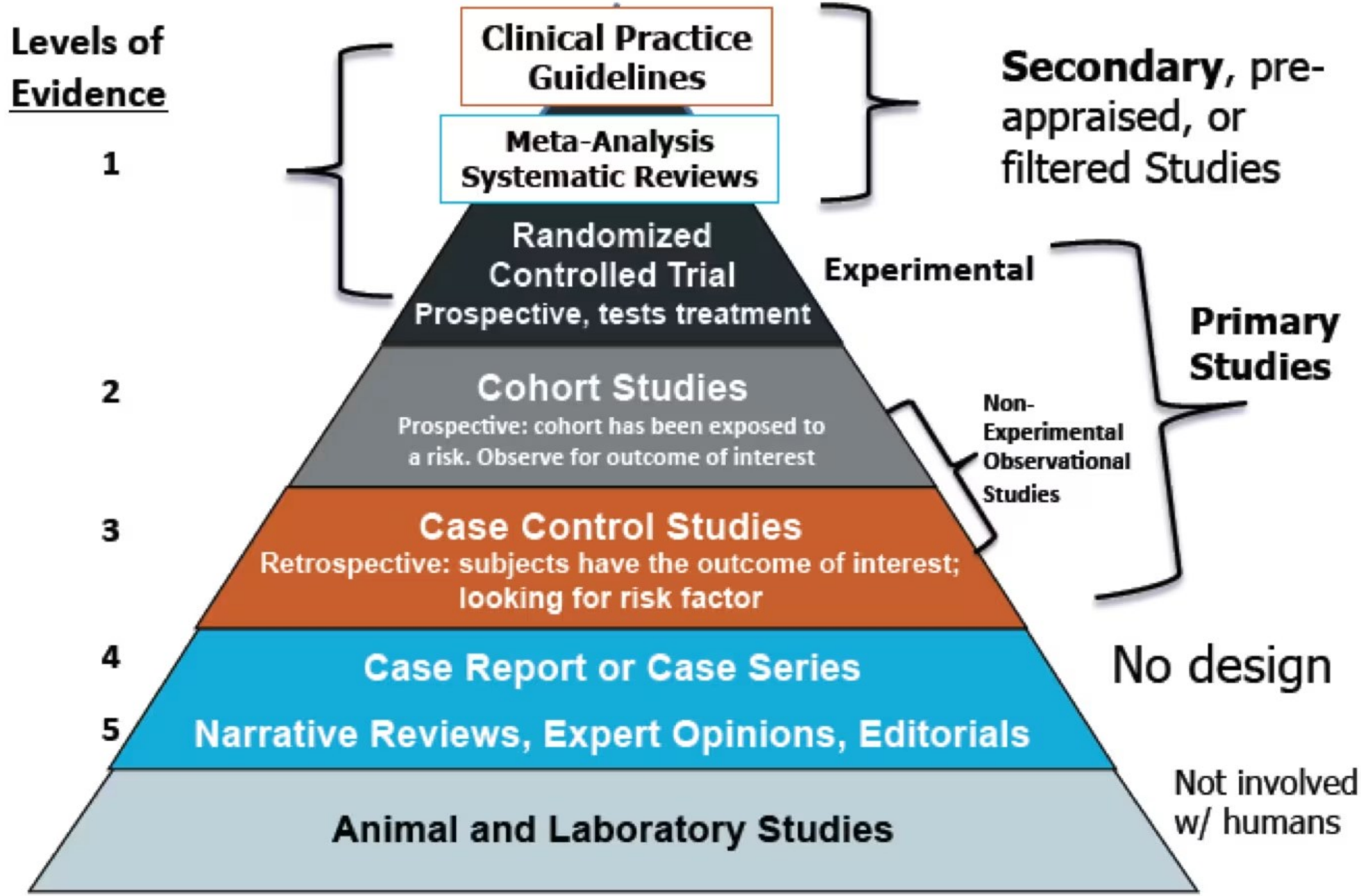
# Systematic Review & Meta-analysis (1)

Konstantinos I. Bougioukas, MSc, PhD

*School of Psychology*
*Faculty of Philosophy*
*Aristotle University of Thessaloniki*


Levels of Evidence
1
2
3
4
5
Clinical Practice Guidelines
Meta-Analysis Systematic Reviews
Randomized Controlled Trial
Prospective, tests treatment
Cohort Studies
Prospective: cohort has been exposed to a risk. Observe for outcome of interest
Case Control Studies
Retrospective: subjects have the outcome of interest; looking for risk factor
Case Report or Case Series
Narrative Reviews, Expert Opinions, Editorials
Animal and Laboratory Studies
Secondary, pre-appraised, or filtered Studies
Experimental
Non-Experimental Observational Studies
Primary Studies
No design
Not involved w/ humans

Health Information and Libraries Journal

DOI: 10.1111/hir.12276

Review Article

# Meeting the review family: exploring review types and associated information retrieval requirements

Anthea Sutton*, Mark Clowes, Louise Preston & Andrew Booth
School of Health and Related Research (ScHARR), The University of Sheffield, Sheffield, UK

https://doi.org/10.1111/hir.12276

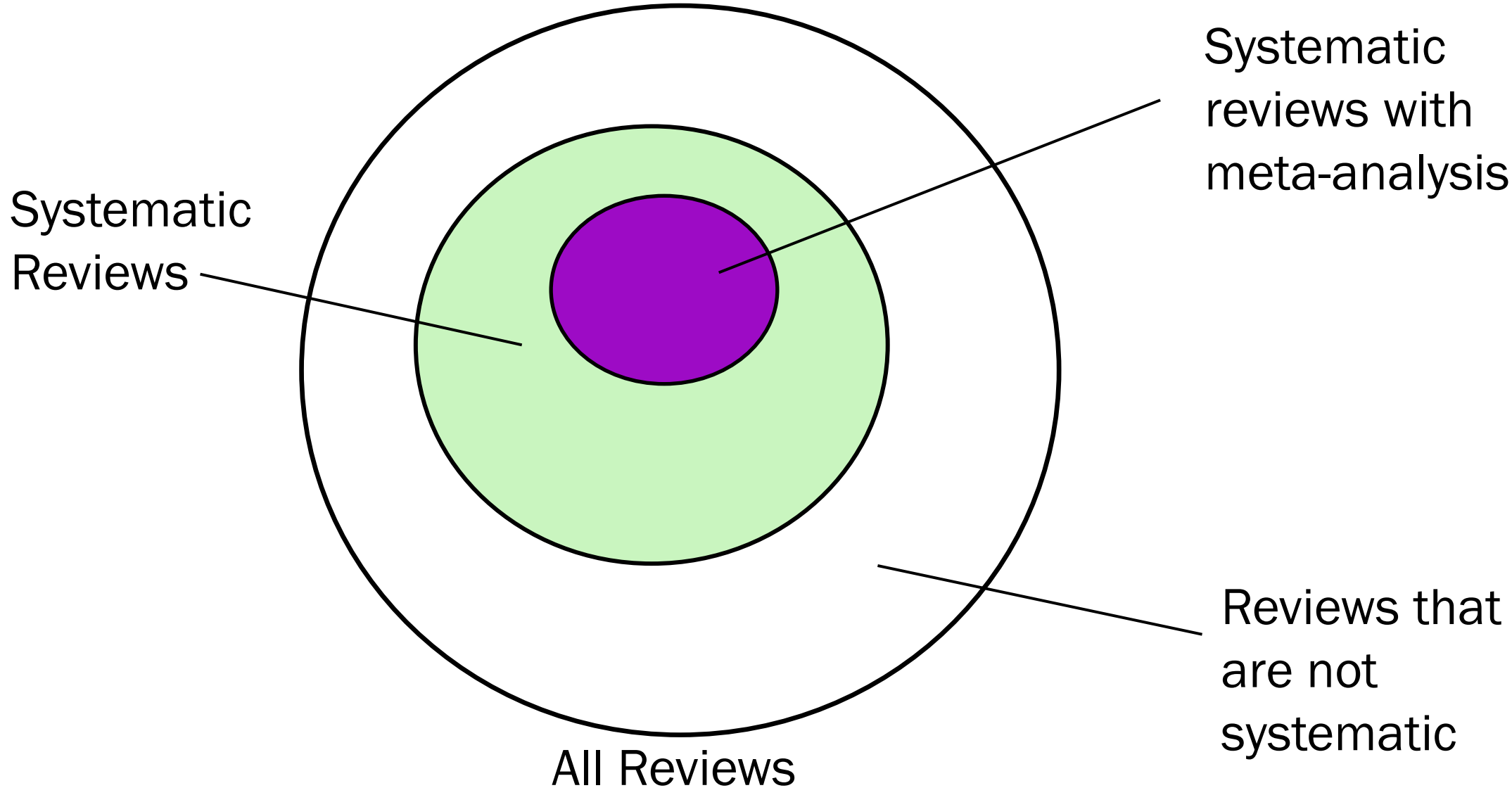
# Meeting the review family

Forty-eight review types were identified and categorised into **seven** families.

- **Traditional review**
- **Systematic review**
- Rapid review
- **Qualitative review**
- **Mixed-methods review**
- Purpose-specific review (HTA)
- **Review of reviews**

But there are much more:

- **Scoping review**
- Living review
- Economic evaluation review
- **Measurement properties review**


Systematic
Reviews
Systematic
reviews with
meta-analysis
Reviews that
are not
systematic
All Reviews

# Systematic Review (SR)

Systematic review is a review of clearly formulated question that uses explicit, pre-planned scientific methods to:

- Identify
- Select
- Appraise
- Synthesize results

from similar but separate primary studies.

**NOTE:** Meta-analysis (MA), the statistical analysis of a large collection of results from individual studies is an optional component of a systematic review

# Who Is Doing Systematic Reviews?

- Independent researchers
- Universities, Institutes, and Organizations: Cochrane Collaboration, Joanna Briggs Institute (JBI), University of York—Centre for Reviews and Dissemination.
- Groups interested in policy (professional societies, governments, payers)

# Who Is Using Systematic Reviews? Decision Makers

- Individual doctors, psychologists and researchers
- Patients
- Guidelines producers
- Policy makers
- Purchasers
- Regulatory authorities

# Classifying Question Types

| Question | Classification/type |
|---|---|
| What proportion of the population is newly diagnosed with this problem each year? | Incidence |
| What proportion of the population is currently living with this problem? | Prevalence |
| What should be done to treat this problem? | Therapy |
| Will detecting this problem early, before symptoms, make a difference in my health? | Screening |
| How good is this test at detecting this problem? | Diagnostic accuracy |
| What is the likely outcome of this problem? | Prognosis |
| Will there be any negative effects (of an intervention)? | Harm |
| What causes this problem? | Etiology |
| How can this problem be prevented? | Prevention |

# Examples of Types of Questions

| Type question | Example |
|---|---|
| Incidence, prevalence | What is the incidence of low birth weight in minority populations compared to the white population? |
| Therapy | Is exercise effective in improving quality of life in persons with COPD? |
| Screening | Is PSA to detect prostate cancer effective in reducing mortality? |
| Diagnostic accuracy | How effective is an MRI at detecting new breast cancers in follow-up of women with breast cancer having lumpectomy? |
| Prognosis | What is the effect of pregnancy on exacerbating the symptoms of MS |
| Harm | What proportion of postmenopausal women receiving Ca++/vita D can expect to have kidney stones? |
| Etiology | Is coffee consumption causally associated with developing pancreatic cancer? |

# One Size Does NOT Fit All! Use Your Question Classification to Seek the Appropriate Type of Evidence

| Question | Look for evidence from: |
|---|---|
| Incidence, prevalence | Surveys, cohort studies |
| Therapy | Clinical trials |
| Screening | Clinical trials |
| Diagnostic accuracy | Clinical trials, cross sectional studies |
| Prognosis | Clinical trials, cohort studies |
| Harm | Clinical trials, cohort studies, case control studies |
| Etiology | Cohort studies, case control |

# How Do You Do a Systematic Review?

Step 1: Gather together your team and develop a focused research question

Step 2: Develop your research plan i.e., protocol (pre-registration, registration)

Step 3: Screen and collect studies

Step 4: Abstract data, appraise risk of bias in the individual studies and heterogeneity

Step 5: Synthesize findings, interpret, and assess overall body of evidence

Step 6: Write report

Step 7: Update

# Define Research Question

Establish in detail the primary and secondary aims of the study (focused or broad, scanning the literature to identify gaps in the field)

- **P**opulation
- **I**ntervention(s)
- **C**omparison(s)
- **O**utcome

EXAMPLE

# Psychological therapies for anxiety and depression in children and adolescents with long-term physical conditions

**P** children and adolescents aged up to 18 years with depression or anxiety or both with 'long-term conditions'

**I** individual or group-based psychological or psychologically-oriented therapy excluding e-health therapies

**C** placebo

**O** changes in severity of anxiety and depression symptoms measured separately using validated scales

Thabrew H, et a. https://doi.org/10.1002/14651858.CD012488.pub2

# Search Methods for Identification of Studies

**Major bibliographic databases for RCTs and observational studies:**

- MEDLINE/PubMED (www.ncbi.nlm.nih.gov/sites/entrez?)
- EMBASE (www.embase.com)
- Cochrane Central Register of Controlled Trials (CENTRAL) (www.thecochranelibrary.com)
- National and regional databases (often local language)
  - LILACS (http://lilacs.bvsalud.org)
- Subject-specific databases
  - CINAHL (http://www.ebscohost.com/academic/cinahl-plus-with-full-text/ )
  - PsychINFO (http://www.apa.org/pubs/databases/psycinfo/index.aspx)
  - OTSeeker (www.otseeker.com)

# Other Bibliographic Databases to Consider

- Citation databases
  - Web of Science (www.thomsonreuters.com)
  - Scopus (www.scopus.com)
    - ✓ Dissertations, thesis databases
      - ProQuest (www.Proquest.com)

- Gray literature databases
  - Opengrey (www.opengrey.eu): Industry files or internal reports
  - National Guidelines Clearing House (www.guideline.gov/)

# Development of Search Strategy

1. Start with simple search strategy
2. Run search, and retrieve reports
3. Analyze controlled vocabulary terms and keywords of studies fitting your criteria, and revise strategy
4. Re-run search with revised strategy
5. Repeat steps 2 through 4 if necessary
6. Run optimal search strategy
7. Retrieve reports identified with optimal search strategy

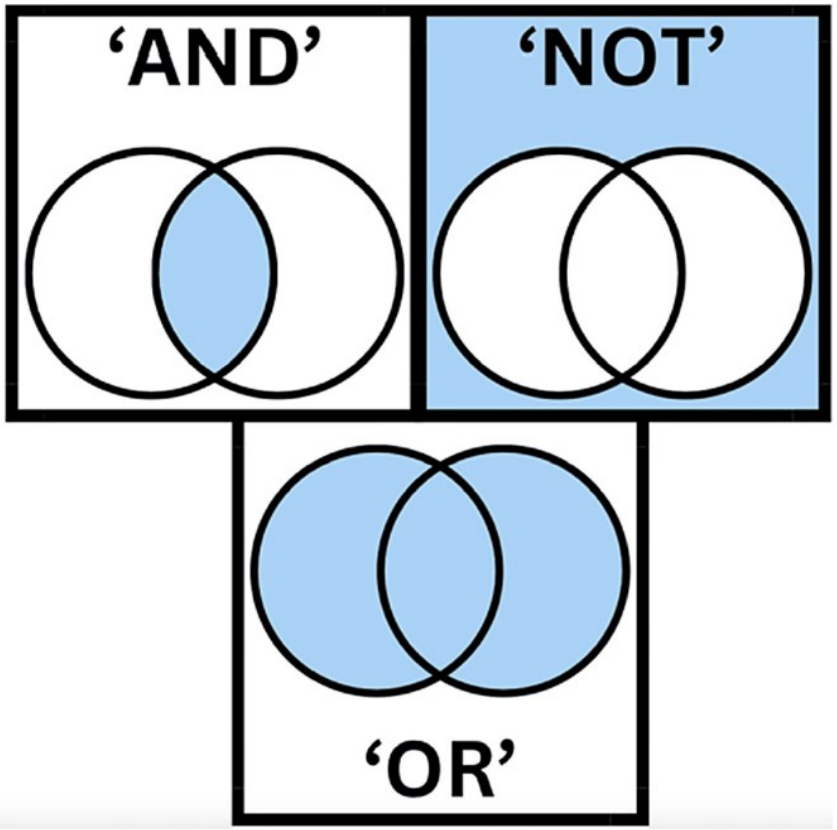
# Controlled Vocabulary

e.g., MeSH terms, Emtree terms

- Consistency
  - Alternative spellings
  - Synonyms
  - Plurals
  - Related terms

- For systematic review searches, use controlled vocabulary ***and*** keywords
  - If you use only keywords, you could miss articles that don't use your precise terms
  - If you use only controlled vocabulary, you could miss articles that have not been indexed yet or that have older indexing

# Boolean operators




Start with the **general format**:

- (**P**opulation OR synonym #1 OR synonym #2) AND
- (**I**ntervention OR synonym #1 OR synonym #2) AND
- (**C**omparator OR synonym #1 OR synonym #2) AND
- (**O**utcome OR synonym #1 OR synonym #2) AND

Add study type filter terms

Condition = (*anxiety OR depressi* OR mood OR mutism or neuroses OR neurotic OR "obsessive compulsive" OR panic OR *phobi* OR psychoneuroses OR "stress disorder*" OR"psychological stress" OR "school refusal")*

AND Comorbidity = *not empty*

AND Age Group = *(child OR adolescent)*

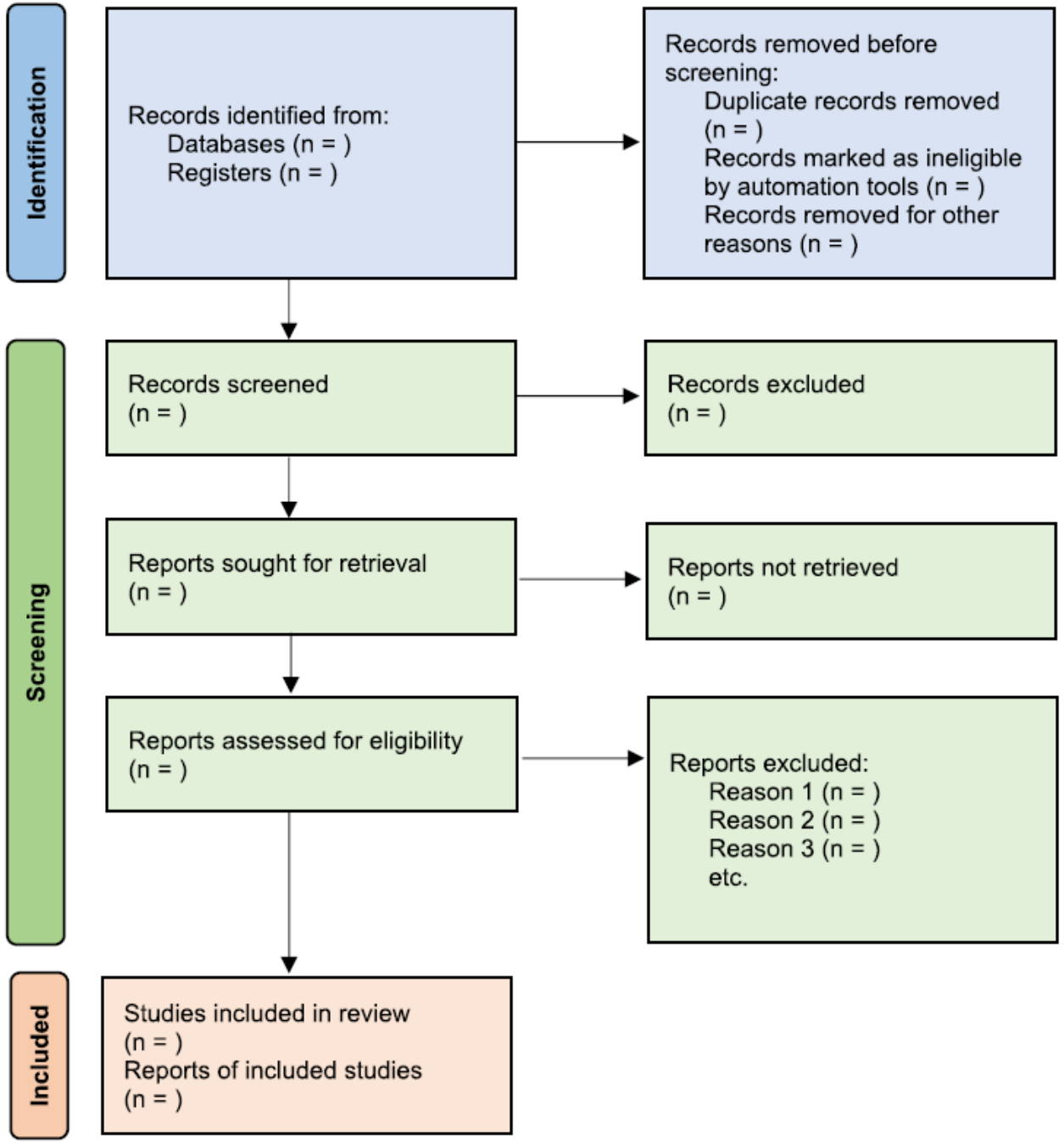
# Screening and selection of studies

It is necessary to upload all our reference files (RIS, BibText, CSV, bib, or txt file) into a predetermined tool for the collection and management of records.

- Rayyan (https://www.rayyan.ai/)
- Covidence (https://www.covidence.org/)
- EPPI (https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2914)
- CADIMA (https://www.cadima.info/)
- DistillerSR (https://www.distillersr.com/products/distillersr-systematic-review-software )

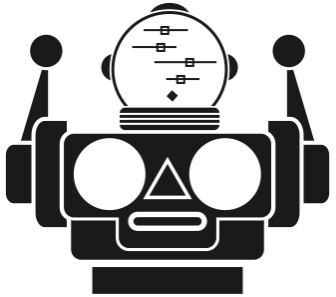
Screening titles and abstracts of studies that are identified by the search (2 reviewers, 1 reviewer + AI). Eligible or potentially-eligible articles are retrieved for full-text inspection by two authors independently.

# PRISMA flow chart



# Data extraction

At least two authors independently extract data on trial characteristics, the methodology, participant characteristics, intervention characteristics, outcome measures, and outcome data.



RobotReviewer

It uses AI to summarize uploaded papers into PICO-based research questions.

https://www.robotreviewer.net/

| Challenges in study selection and extraction | Acceptable solution |
| --- | --- |
| Different measurements | • Choose the most common or valid measure<br>• Convert to a common measure<br>• Request more information from author |
| Different statistical analysis | • Set a spreadsheet for data overview before transfer to statistical software |
| Incompatible results | • Report them in the text as a narrative synthesis<br>• Report them in a separate table |
| Unclear data | • Contact authors for clarification |
| Unusable or non-numerical data | • Report in the review to avoid bias |
| No outcome of interest or selective reporting of outcomes | • Request more information from authors<br>• Can exclude the study if specific outcome is not included in the eligibility criteria |

# Assessment of risk of bias in included studies

| Randomized trials |
| --- |
| RoB2 - Revised Cochrane Risk-of-Bias tool for randomized trials |

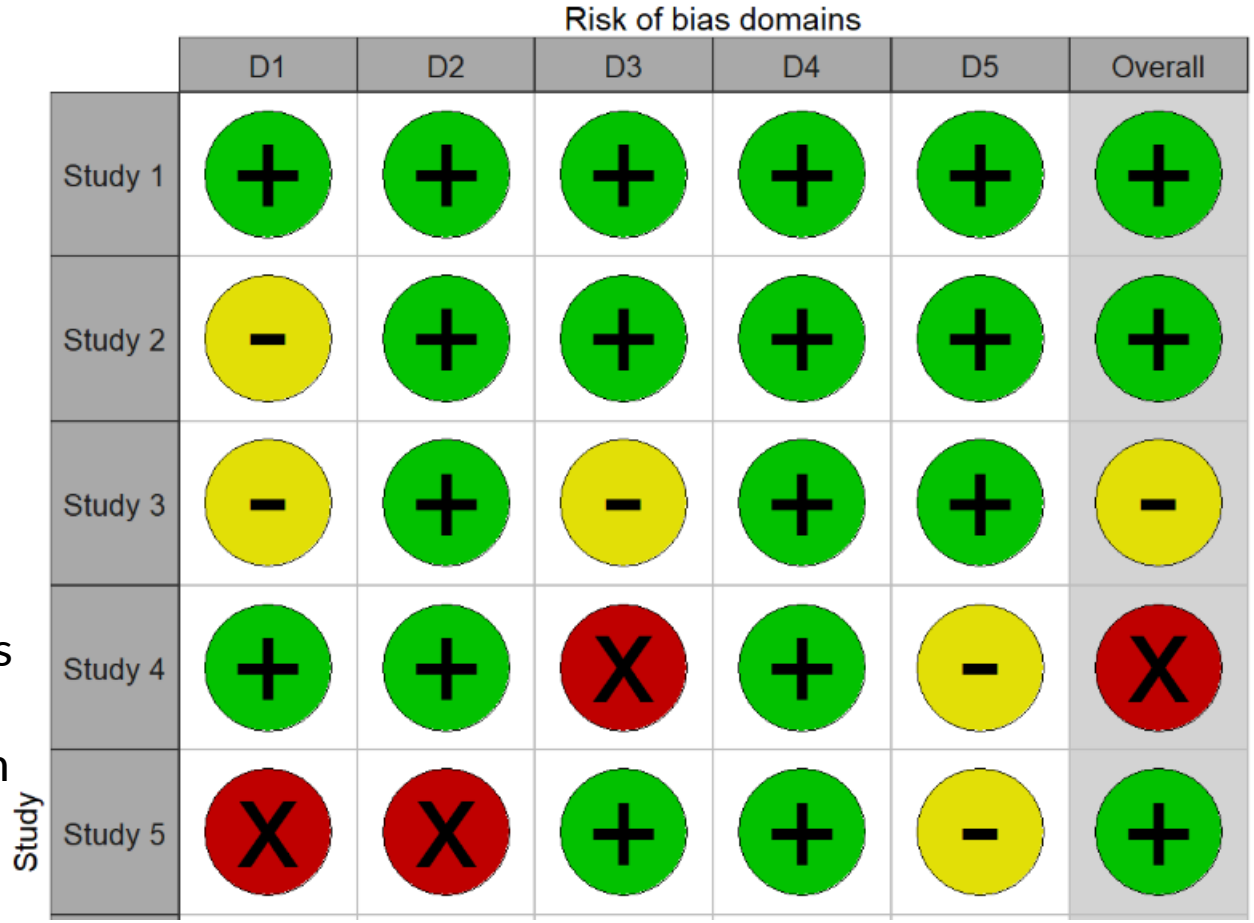
## Non-randomized studies

| |
| --- |
| ROBINS-I - Cochrane's Risk of Bias In Non-randomized Studies – of Interventions |
| ROBINS-E - Cochrane's Risk of Bias In Non-randomized Studies - of Exposure |
| The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomized studies |
| NIH Study Quality Assessment Tools - for Controlled Intervention, Observational Cohort and Cross-Sectional, Case-Control, Case Series Studies |
| QUADAS – Quality Assessment of Diagnostic Accuracy Studies |
| QUIPS – Quality in Prognosis Studies |
| The Joanna Briggs Institute Prevalence Critical Appraisal Tool - for studies reporting prevalence data |
| PROBAST - Prediction model Risk Of Bias ASsessment Tool |

# Traffic light plots

The risk-of-bias judgment in each domain for each study, as well as the overall risk-of-bias judgement for that study.

**NOTE:** RobotReviewer provides a preliminary risk of bias assessment for RCTs, which can streamline users' own assessments

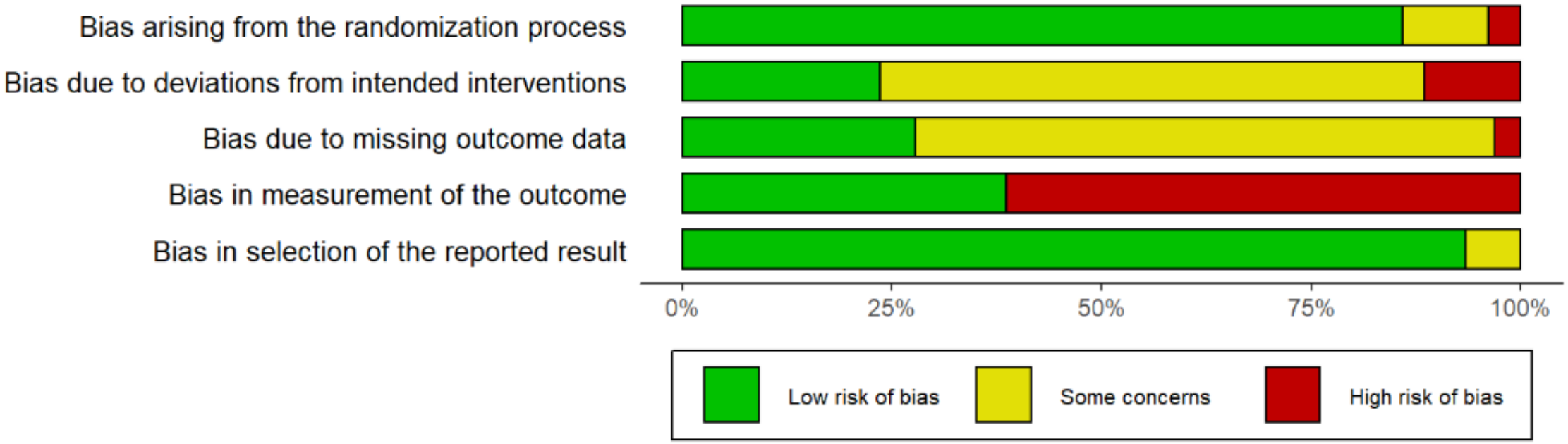



Domains:
D1: Bias due to randomisation.
D2: Bias due to deviations from intended intervention.
D3: Bias due to missing data.
D4: Bias due to outcome measurement.
D5: Bias due to selection of reported result.

Judgement
X High
- Some concerns
+ Low

https://cran.r-project.org/web/packages/robvis/vignettes/Introduction_to_robvis.html

# Overall risk-of-bias judgments



# Data synthesis

Example

- When available and sufficiently clinically and statistically homogenous, we combined final score data from included trials in meta-analyses. ← Type of synthesis
- As we were anticipating heterogeneity of data, we planned on analysing the data using RevMan 5.3 using a random effects model for analysis. ← Type of model
- We presented the 'Risk of bias' assessment in a 'Risk of bias' graph. We presented results for each comparison as forest plots, when appropriate. ← Presentation of results
- We provided narrative summaries for comparisons with fewer than two available studies and those with a moderate or high level of statistical heterogeneity following exploration of heterogeneity. ← Heterogeneity

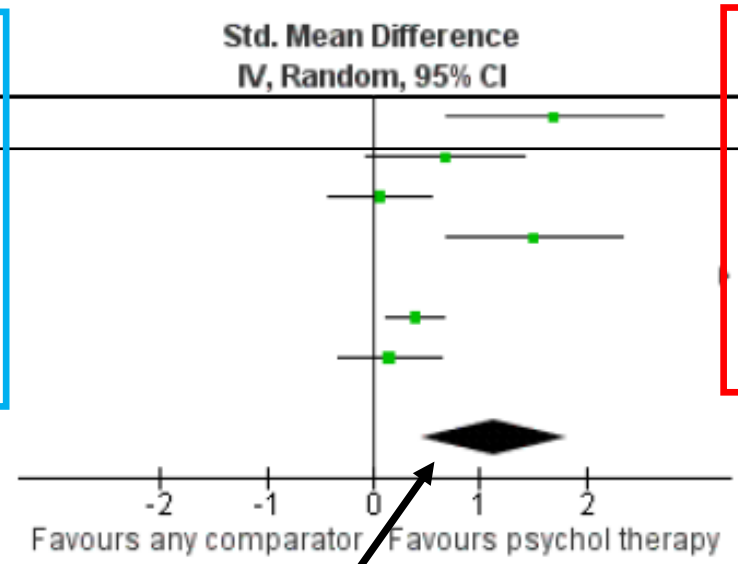
Extracted Data from studies

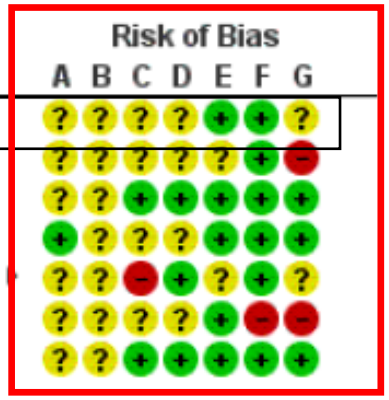
Effect size (95% CI)

| Study or Subgroup | Experimental Mean | Experimental SD | Experimental Total | Control Mean | Control SD | Control Total | Weight | Std. Mean Difference IV, Random, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Beebe 2010 | 9.73 | 13.71 | 11 | -13.36 | 12.48 | 11 | 12.7% | 1.69 [0.69, 2.70] |
| Bignall 2015 | 74.83 | 7.32 | 14 | 67.12 | 13.71 | 16 | 14.5% | 0.67 [-0.07, 1.41] |
| Grey 1998 | 67.8 | 11.3 | 34 | 67 | 13.5 | 31 | 16.0% | 0.06 [-0.42, 0.55] |
| Martinović 2006 | 52.78 | 6.4 | 15 | 41.35 | 8.26 | 15 | 14.0% | 1.51 [0.68, 2.33] |
| Moghanloo 2015 | 24.69 | 2.57 | 17 | 11.87 | 2.23 | 17 | 9.7% | 5.20 [3.73, 6.68] |
| Szigethy 2014 | 142.52 | 20.8 | 110 | 133.85 | 22.9 | 107 | 17.0% | 0.40 [0.13, 0.66] |
| Wei 2017 | 66 | 14.4 | 33 | 63.8 | 14.9 | 33 | 16.1% | 0.15 [-0.33, 0.63] |
| **Total (95% CI)** | | | **234** | | | **230** | **100.0%** | **1.13 [0.44, 1.82]** |

Heterogeneity: Tau² = 0.71; Chi² = 56.00, df = 6 (P < 0.00001); I² = 89%

Test for overall effect: Z = 3.21 (P = 0.001)




Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Weight of each study (1/Variance)

Summary Effect size (95% CI)

# Grading of Recommendations Assessment Development and Evaluation (GRADE) criteria


- **High** ⨁⨁⨁⨁ Very confident that the true effect lies close to the estimate of the effect.
- **Moderate** ⨁⨁⨁◯ Moderately confident in the effect estimate: true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
- **Low** ⨁⨁◯◯ Limited confidence in the effect estimate: true effect may be substantially different from the estimate of effect.
- **Very low** ⨁◯◯◯ Very little confidence in the effect estimate: true effect is likely to be substantially different from the estimate of effect.

https://book.gradepro.org/

https://www.gradeworkinggroup.org/

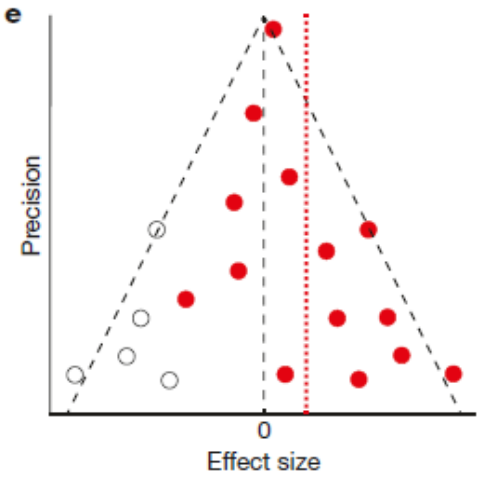
https://gdt.gradepro.org/app/handbook/handbook.html

A summary of GRADE's approach to rating quality of evidence

| Study design | Initial quality of a body of evidence | Lower if | Higher if | Quality of a body of evidence |
|---|---|---|---|---|
| Randomized trials | High | Risk of Bias<br>−1 Serious<br>−2 Very serious<br>Inconsistency<br>−1 Serious<br>−2 Very serious<br>Indirectness<br>−1 Serious<br>−2 Very serious<br>Imprecision<br>−1 Serious<br>−2 Very serious<br>Publication bias<br>−1 Likely<br>−2 Very likely | Large effect<br>+1 Large<br>+2 Very large<br>Dose response<br>+1 Evidence of a gradient<br>All plausible residual confounding<br>+1 Would reduce a demonstrated effect<br>+1 Would suggest a spurious effect if no effect was observed | High (four plus: ⊕⊕⊕⊕)<br>Moderate (three plus: ⊕⊕⊕◯)<br>Low (two plus: ⊕⊕◯◯)<br>Very low (one plus: ⊕◯◯◯) |
| Observational studies | Low | | | |

# More specific

- Study design (RCT vs Observational)
- Risk of bias (RoB tool)
- Inconsistency (heterogeneity, $I^2$)
- Indirectness (differences in PICO questions)
- Imprecision (sample size, few events, 95% CI of the summary effect)
- Publication Bias (Funnel Plot)



**Psychological therapy compared to any comparator for anxiety and depression in children and adolescents with long-term physical conditions**

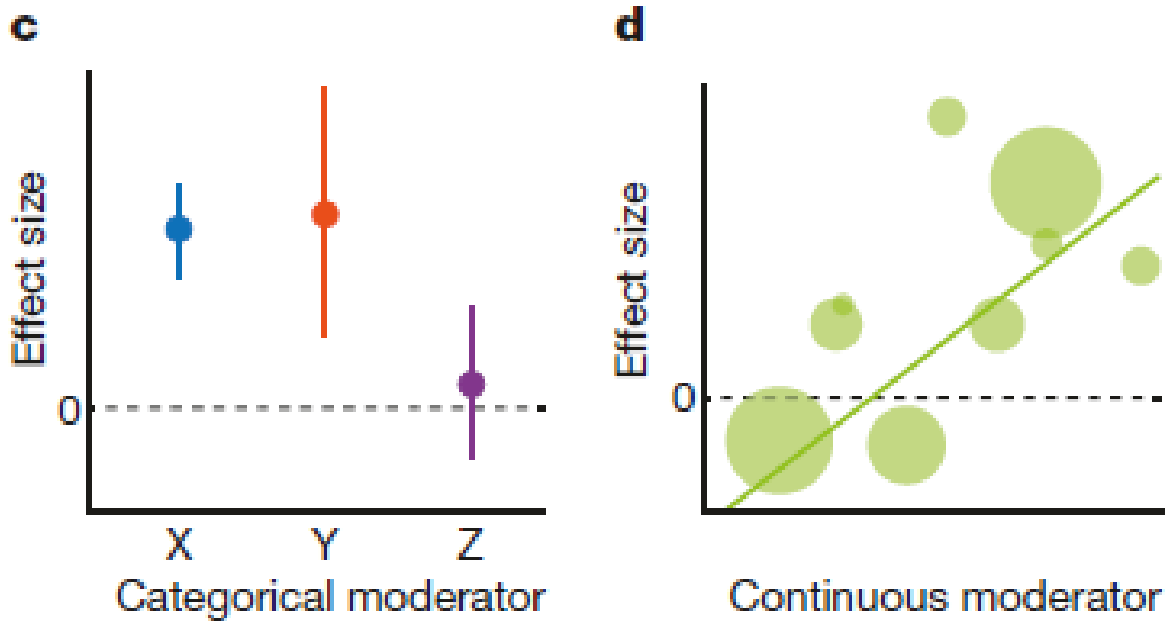
**Patient or population:** anxiety and depression in children and adolescents with long-term physical conditions
**Setting:**
**Intervention:** psychological therapy
**Comparison:** any comparator

| Outcomes | Anticipated absolute effects* (95% CI) | | Relative effect (95% CI) | № of participants (studies) | Certainty of the evidence (GRADE) | comments |
|---|---|---|---|---|---|---|
| | Risk with any comparator | Risk with Psychological therapy | | | | |
| Treatment efficacy: depression short-term | The mean depression short-term was 0 | SMD 0.31 lower (0.59 lower to 0.03 lower) | - | 1121 (16 RCTs) | ⨁⨁◯◯ LOW [1,2,3] | A SMD of 0.31 is a small effect size |
| Treatment efficacy: anxiety short-term | The mean anxiety short-term was 0 | SMD 0.26 lower (0.59 lower to 0.07 higher) | - | 578 (13 RCTs) | ⨁⨁⨁◯ MODERATE [1,2] | The confidence interval crosses the line of no effect |
| Quality of life short-term | The mean quality of life short-term was 0 | SMD 1.13 higher (0.44 higher to 1.82 higher) | - | 464 (7 RCTs) | ⨁⨁◯◯ LOW [1,2,3] | A SMD of 1.13 is a large effect size |
| Functioning short-term | The mean functioning short-term was 0 | SMD 0.49 higher (0.3 lower to 1.29 higher) | - | 483 (7 RCTs) | ⨁⨁◯◯ LOW [1,2,3] | The confidence interval crosses the line of no effect |

[1] High degree of inconsistency between results
[2] Substantial heterogeneity
[3] Upper/lower CI crosses the effect size (SMD) of 0.5 in either direction

# Exploration of data

- Subgroup analysis and investigation of heterogeneity, Meta-regression
- Sensitivity analysis



# Reporting the review

The methods and results of systematic reviews should be reported in sufficient detail to allow users to assess the trustworthiness and applicability of the review findings.


https://www.equator-network.org/

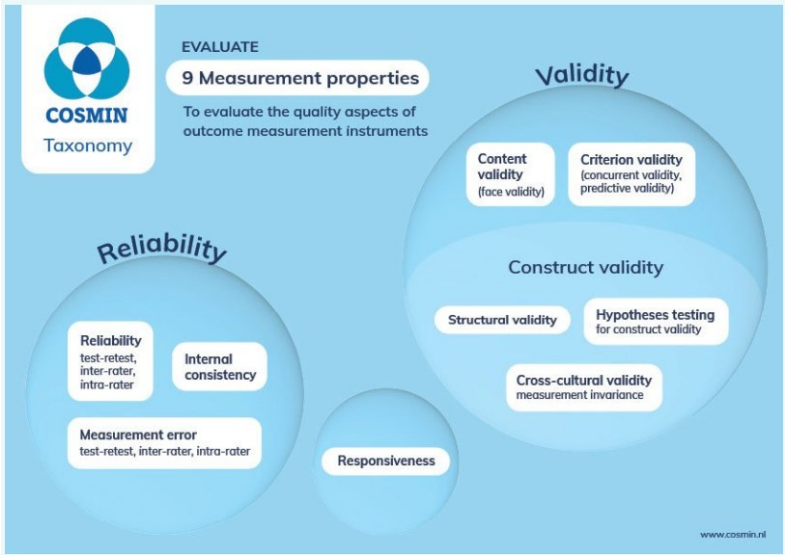

https://lights.science/

# Reporting tools for systematic reviews

- PRISMA 2020 (https://www.bmj.com/content/372/bmj.n160) (https://www.prisma-statement.org/extensions)

- Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline (https://www.bmj.com/content/368/bmj.l6890)

- COSMIN (https://www.cosmin.nl/tools/cosmin-taxonomy-measurement-properties/)



# PRISMA 2020 Checklist

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

# Concise guide to best practices for evidence syntheses

| Resources | Intervention | Diagnostic | Prognostic | Qualitative or mixed methods | Prevalence and incidence | Etiology and risk | Measurement properties | Overviews (umbrella reviews) | Scoping reviews |
|---|---|---|---|---|---|---|---|---|---|
| **Methodological guidance** | Cochrane[b], JBI | Cochrane, JBI | Cochrane | Cochrane, JBI | JBI | JBI | JBI | Cochrane, JBI | JBI |
| **Reporting**[c] | | | | | | | | | |
| Protocol | PRISMA-P[116] | PRISMA-P | PRISMA-P | PRISMA-P | PRISMA-P | PRISMA-P | PRISMA-P | PRISMA-P | PRISMA-P |
| Systematic review | PRISMA 2020[112] | PRISMA-DTA[120] | PRISMA 2020 | eMERGe[213,d] ENTREQ[214,d] | PRISMA 2020 | PRISMA 2020 | PRISMA 2020 | PRIOR[215] | PRISMA-ScR[121] |
| Synthesis without meta-analysis | SWiM[180] | | SWiM[e] | | SWiM[e] | SWiM[e] | SWiM[e] | | |
| **RoB assessment of included studies**[f] | For RCTs: Cochrane RoB2[157] For NRSI: ROBINS-I[158] Other primary research[g] | QUADAS-2[216] | Factor review QUIPS[217] Model review PROBAST[65] | CASP qualitative checklist[218] JBI critical appraisal checklist[219,h] | JBI checklist for studies reporting prevalence data[220] | For NRSI: ROBINS-I[158] Other primary research[g] | COSMIN RoB Checklist[67] | AMSTAR-2[6] or ROBIS[4] | Not required[i] |
| **Overall level of evidence certainty** | GRADE[27] | GRADE adaptation[j] | GRADE adaptation[k] | CERQual[221] ConQual[222,l] | GRADE adaptation[m] | Risk factors[n] | GRADE adaptation[o] | GRADE (for intervention reviews) Risk factors[n] | Not applicable |