BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

The systematic review process
Quick intro

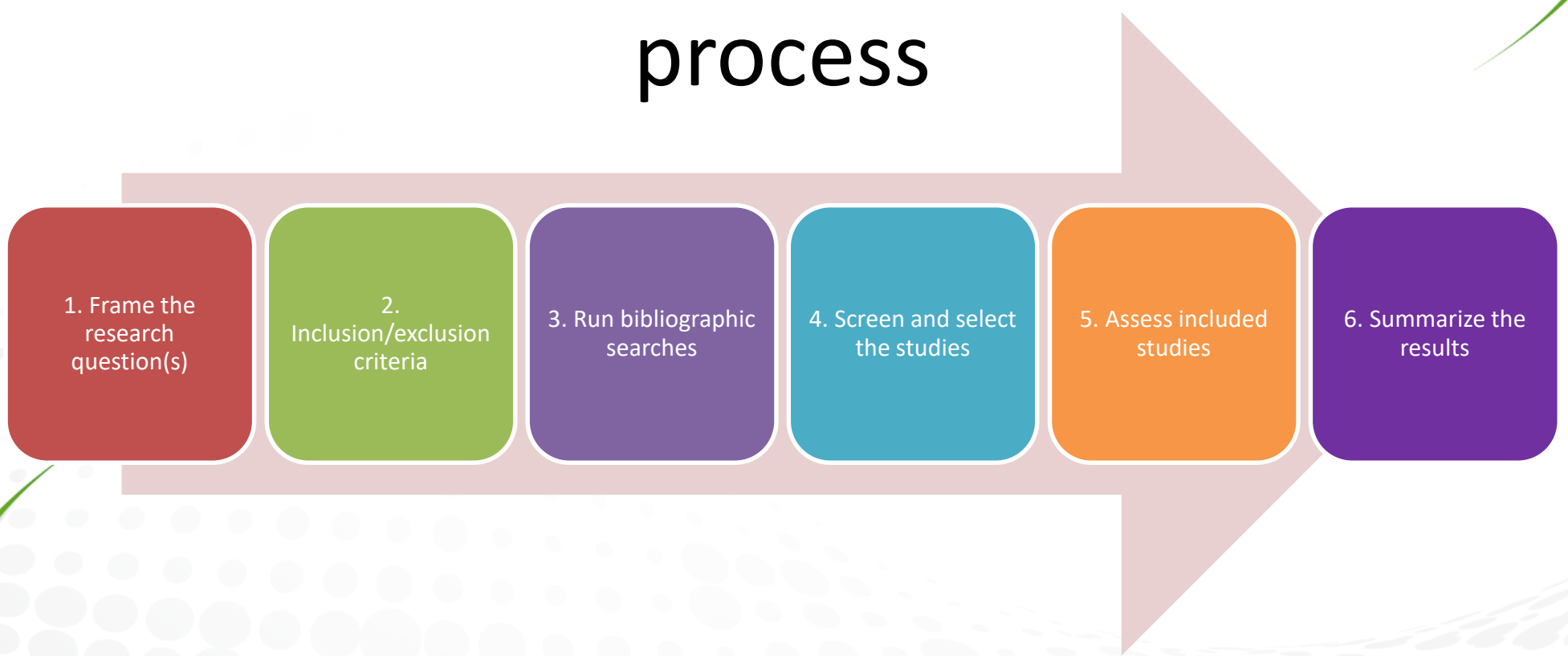
Paolo Gardois.
Biblioteca Federata di Medicina «Ferdinando Rossi». Università di Torino

Meta-analysis Course – Turin, May 18, 2022

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

Overview of the systematic review process



0. Definitions and taxonomies

- «A systematic review attempts to collate all the empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing more reliable findings from which conclusions can be drawn and decisions made»

(Cochrane Handbook for Systematic Reviews of Interventions, 2019)

0. Definitions and taxonomies

- Systematic review = meta-analysis of the effect sizes of experimental studies? No!
  - «Systematic» refers to the methods used to carry out the review:
    - Systematic reviews (as opposed to narrative reviews) use explicit and reproducible methods in all the steps of the review process
  - A «systematic» review may or may not include a meta-analysis to summarize the results:
    - Different techniques exist to summarize the results (quantitative and qualitative)
  - Systematic reviews may be about all sorts of studies (observational, qualitative, …), not just experimental studies of interventions

Review Type	Aim	Question Format	Question Example
Effectiveness	To evaluate the effectiveness of a certain treatment/practice in terms of its impact on outcomes	Population, Intervention, Comparator/s, Outcomes (PICO) [23]	What is the effectiveness of exercise for treating depression in adults compared to no treatment or a comparison treatment? [69]
Experiential (Qualitative)	To investigate the experience or meaningfulness of a particular phenomenon	Population, Phenomena of Interest, Context (PICo) [13]	What is the experience of undergoing high technology medical imaging (such as Magnetic Resonance Imaging) in adult patients in high income countries? [70]
Costs/Economic Evaluation	To determine the costs associated with a particular approach/treatment strategy, particularly in terms of cost effectiveness or benefit	Population, Intervention, Comparator/s, Outcomes, Context (PICOC) [14]	What is the cost effectiveness of self-monitoring of blood glucose in type 2 diabetes mellitus in high income countries? [71]
Prevalence and/or Incidence	To determine the prevalence and/or incidence of a certain condition	Condition, Context, Population (CoCoPop) [15]	What is the prevalence/incidence of claustrophobia and claustrophobic reactions in adult patients undergoing MRI? [72]
Diagnostic Test Accuracy	To determine how well a diagnostic test works in terms of its sensitivity and specificity for a particular diagnosis	Population, Index Test, Reference Test, Diagnosis of Interest (PIRD) [16]	What is the diagnostic test accuracy of nutritional tools (such as the Malnutrition Screening Tool) compared to the Patient Generated Subjective Global Assessment amongst patients with colorectal cancer to identify undernutrition? [73]
Etiology and/or Risk	To determine the association between particular exposures/risk factors and outcomes	Population, Exposure, Outcome (PEO) [17]	Are adults exposed to radon at risk for developing lung cancer? [74]
Expert opinion/policy	To review and synthesize current expert opinion, text or policy on a certain phenomena	Population, Intervention or Phenomena of Interest, Context (PICo) [18]	What are the policy strategies to reduce maternal mortality in pregnant and birthing women in Cambodia, Thailand, Malaysia and Sri Lanka? [75]
Psychometric	To evaluate the psychometric properties of a certain test, normally to determine how the reliability and validity of a particular test or assessment.	Construct of interest or the name of the measurement instrument(s), Population, Type of measurement instrument, Measurement properties [31, 32]	What is the reliability, validity, responsiveness and interpretability of methods (manual muscle testing, isokinetic dynamometry, hand held dynamometry) to assess muscle strength in adults? [76]
Prognostic	To determine the overall prognosis for a condition, the link between specific prognostic factors and an outcome and/or prognostic/prediction models and prognostic tests.	Population, Prognostic Factors (or models of interest), Outcome (PFO) [20, 34,35,36]	In adults with low back pain, what is the association between individual recovery expectations and disability outcomes? [77]
Methodology	To examine and investigate current research methods and potentially their impact on research quality.	Types of Studies, Types of Data, Types of Methods, Outcomes [39] (SDMO)	What is the effect of masked (blind) peer review for quantitative studies in terms of the study quality as reported in published reports? (question modified from Jefferson 2007) [40]

Munn et al., 2018

1. Frame the research question(s)

- We will now focus on systematic reviews of the effectiveness of interventions →
  - Choose relevant questions to support clinical practice. Such questions should be FINER (**F**easible, **I**nteresting, **N**ovel, **E**thical, and **R**elevant)
  - Take into account potential benefits and harms
  - Use the PICO framework (**P**opulation, **I**ntervention, **C**omparison, **O**utcome)
  - Broad vs narrow questions

1. Frame the research question(s)

- Examples from Cochrane Reviews:
  - «To assess the effectiveness of international travel-related control measures during the COVID-19 pandemic on infectious disease transmission and screening-related outcomes.» (Burns et al., 2021)
  - «To assess the effects of wearing compression stockings versus not wearing them for preventing deep venous thrombosis in people travelling on flights lasting at least four hours» (Clarke et al., 2021).
  - «To assess the effects of education and training interventions compared to no intervention or alternative interventions for preventing sharps injuries and splash exposures in health care workers» (Cheetham et al., 2021)

2. Define the protocol: inclusion/exclusion criteria

- Inclusion/exclusion criteria (IEC) are based on the P-I-C of the PICO framework
- Population: “The criteria for considering types of people included in studies in a review should be sufficiently broad to encompass the likely diversity of studies and the likely scenarios in which the interventions will be used, but sufficiently narrow to ensure that a meaningful answer can be obtained when studies are considered together” *(Cochrane Handbook for Systematic Reviews of Interventions, 2019*)

2. Define the protocol: inclusion/exclusion criteria

- How is the disease/condition defined?
- What are the most important characteristics that describe these people (participants)?
- Are there any relevant demographic factors (e.g. age, sex, ethnicity)?
- What is the setting (e.g. hospital, community, etc)?
- Who should make the diagnosis?
- Are there other types of people who should be excluded from the review (because they are likely to react to the intervention in a different way)?
- How will studies involving only a subset of relevant participants be handled?

(Cochrane Handbook for Systematic Reviews of Interventions, 2019)

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

2. Define the protocol: inclusion/exclusion criteria

Box 3.2.b Factors to consider when developing criteria for 'Types of interventions'

- What are the experimental and control (comparator) interventions of interest?
- Does the intervention have variations (e.g. dosage/intensity, mode of delivery, personnel who deliver it, frequency, duration or timing of delivery)?
- Are all variations to be included (for example, is there a dose below which the intervention may not be clinically appropriate, will all providers be included)?
- Will studies including only part of the intervention be included?
- Will studies including the intervention of interest combined with another intervention (co-intervention) be included?
- Have the different meanings of phrases such as 'control', 'placebo', 'no intervention' or 'usual care' been considered?

(Cochrane Handbook for Systematic Reviews of Interventions, 2019)

2. Define the protocol: inclusion/exclusion criteria

- "Outcomes usually are not part of the criteria for including studies, and a Cochrane Review would typically seek all sufficiently rigorous studies (most commonly randomized trials) of a particular comparison of interventions in a particular population of participants, irrespective of the outcomes measured or reported. It should be noted that some reviews do legitimately restrict eligibility to specific outcomes." *(Cochrane Handbook for Systematic Reviews of Interventions, 2019*)

2. Define the protocol: inclusion/exclusion criteria

- Further criteria include:
  - Specific settings
  - Study designs (experimental > observational > anecdotal > opinions)
  - Publication status (e.g.: published, partially published, unpublished, grey literature, etc.)
  - Publication dates
  - Languages
  - Geographical locations

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

2. Define the protocol: inclusion/exclusion criteria

Cochrane Database of Systematic Reviews | Review - Intervention

Hematopoietic stem cell transplantation for people with β-thalassaemia

Akshay Sharma, Vanitha A Jagannath, Latika Puri Authors' declarations of interest

Criteria for considering studies for this review

Types of studies

Randomised controlled trials (RCTs) and quasi-randomised trials.

Types of participants

People with a diagnosis of transfusion-dependent thalassaemia.

Types of interventions

Any type of HSCT, including bone marrow (bone marrow), peripheral blood (peripheral blood derived stem cells), or umbilical cord blood; any donor (an HLA-identical related donor, HLA-matched unrelated donor, or an HLA -mismatched donor) with any type of conditioning regimen will be included. We will also include trials that have autologous HSCT with genetically modified hematopoietic stem cells (gene therapy) as one of the interventions.

Trials comparing these interventions with each other or with standard therapy (regular transfusion and chelation regimen) were eligible for inclusion.

3. Run bibliographic searches

- Define bibliographic sources
- Devise systematic search strategies (reducing bias)

3. Bibliographic searches: define sources

- «The search for studies (…) should be as extensive as possible in order to reduce the risk of reporting bias and to identify as much relevant evidence as possible (…). Searches of health-related bibliographic databases are generally the most efficient way to identify an initial set of relevant reports of studies (…). Database selection should be guided by the review topic» *(Cochrane Handbook for Systematic Reviews of Interventions, 2019*)

3. Bibliographic searches: define sources

- Most frequently used databases:
  - Pubmed / MEDLINE
  - EMBASE
  - Cochrane CENTRAL
  - CINAHL
  - PsycInfo
- National and regional databases
- «Citation indexes are bibliographic databases that record instances where a particular reference is cited, in addition to the standard bibliographic content. Citation indexes can be used to identify studies that are similar to a study report of interest, as it is probable that other reports citing or cited by a study will contain similar or related content» *(Cochrane Handbook 2019)*
  - Examples: Web of Science, Scopus

3. Bibliographic searches: define sources

- It is important to identify **ongoing** and **unpublished** studies, by using:
  - contacts with relevant authors in the field;
  - trials and trial results registers (e.g.: clinicaltrials.gov, WHO International Clinical Trials Registry Platform (ICTRP) portal);
  - regulatory documents from agencies (EMA, FDA, …), with clinical studies reports (CSR)
  - "grey literature", dissertations, conference abstracts, technical reports…
- Moreover, it can be useful to search:
  - other reviews, HTA reports, guidelines
  - Handsearching fulltext electronic journals, etc.

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

3. Bibliographic searches: define sources

Cochrane Database of Systematic Reviews | Review - Intervention

New search | Conclusions changed

Topical emollient for preventing infection in preterm infants

Jemma Cleminson, William McGuire Authors' declarations of interest

Search methods for identification of studies

Electronic searches

We conducted a comprehensive search in January 2021 including: Cochrane Central Register of Controlled Trials (CENTRAL 2021, Issue 1, 01 January 2015 to 08 January 2021) in the Cochrane Library and Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) (01 January 2015 to 08 January 2021). We have included the search strategies for each database in Appendix 1. We did not apply language restrictions.

We searched clinical trial registries for ongoing or recently completed trials. We searched the World Health Organization's International Clinical Trials Registry Platform (ICTRP) (who.int/ictrp/search/en/), and the United States' National Library of Medicine's ClinicalTrials.gov (clinicaltrials.gov) via Cochrane CENTRAL. Additionally, we searched the ISRCTN Registry (http://www.isrctn.com/) for any unique trials not found through the Cochrane CENTRAL search.

Previous search details are listed in Appendix 2.

Searching other resources

We searched the reference lists of any articles selected for inclusion in this review.

3. Bibliographic searches: search strategies

- Identifying relevant terms (semantic analysis):
  - manual analysis using sample articles (documents) of interest
  - text mining techniques
  - it is important to consider synonyms, related terms, variant spellings
- Build search strings, made by terms + operators/other syntax elements:
  - Boolean operators (e.g. AND, OR, NOT, ADJ, NEAR, …)
  - Parentheses for term grouping
  - Truncation
  - Wildcards
  - …

3. Bibliographic searches: search strategies

- It is advisable to combine:
  - Controlled vocubulary searches – such as thesauri (e.g. MeSH, EMTree)
  - Text word searches (including terms in the title, abstract, etc.)
- Predefined (and possibly validated) search filters may be considered

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

3. Bibliographic searches: search strategies

Box 3.h Demonstration search strategy for MEDLINE (Ovid format), for the topic 'treating breast cancer with tamoxifen'

1	randomized controlled trial.pt.
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	drug therapy.fs.
6	randomly.ab.
7	trial.ab.
8	groups.ab.
9	or 1-8
10	exp animals/ not humans/
11	9 not 10
12	exp Breast Neoplasms/
13	(breast adj6 cancer$).mp.
14	(breast adj6 neoplasm$).mp.
15	(breast adj6 carcinoma$).mp.
16	(breast adj6 tumour$).mp.
17	(breast adj6 tumor$).mp.
18	or 12-17
19	exp Tamoxifen/
20	tamoxifen.mp.
21	19 or 20
22	11 and 18 and 21

(Cochrane Handbook for Systematic Reviews of Interventions, 2019)

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

4. Screen and select the studies

- Managing references
- Deduplication
- Importing results into specific apps for selection
- The selection process
- Output of the selection process

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

4. Screen and select the studies

- Once the searches on specific sources have been completed, results are imported in a reference management app (e.g.: Endnote) and deduplicated
- The final set of results is the exported to a (web-based) app used to streamline the selection process (e.g. Rayyan, Colandr, or simply an online shared spreadsheet- e.g. Google Sheets)
- As a good practice, every record is assigned a random number, then the dataset is sorted according to the numbers, and a small part (e.g. 5%) is (blindly) screened by all reviewers, in order to test the decision making process according to the inclusion criteria

4. Screen and select the studies

- The level of agreement is then assessed. A discussion follows, to solve the disagreements, and more specific guidelines are adopted.
- The remaining 95% of the dataset is then divided between all reviewers, and for each item a decision is made (inclusion vs. exclusion)
- A first pool of potentially relevant items is produced. Full texts are retrieved, to confirm the inclusion decision
- The final dataset is then ready for data collection

Gardois *et al. BMC Public Health* 2014, **14**:409
http://www.biomedcentral.com/1471-2458/14/409

BMC
Public Health

RESEARCH ARTICLE Open Access

Health promotion interventions for increasing stroke awareness in ethnic minorities: a systematic review of the literature

Paolo Gardois[1*†], Andrew Booth[1†], Elizabeth Goyder[1†] and Tony Ryan[2†]

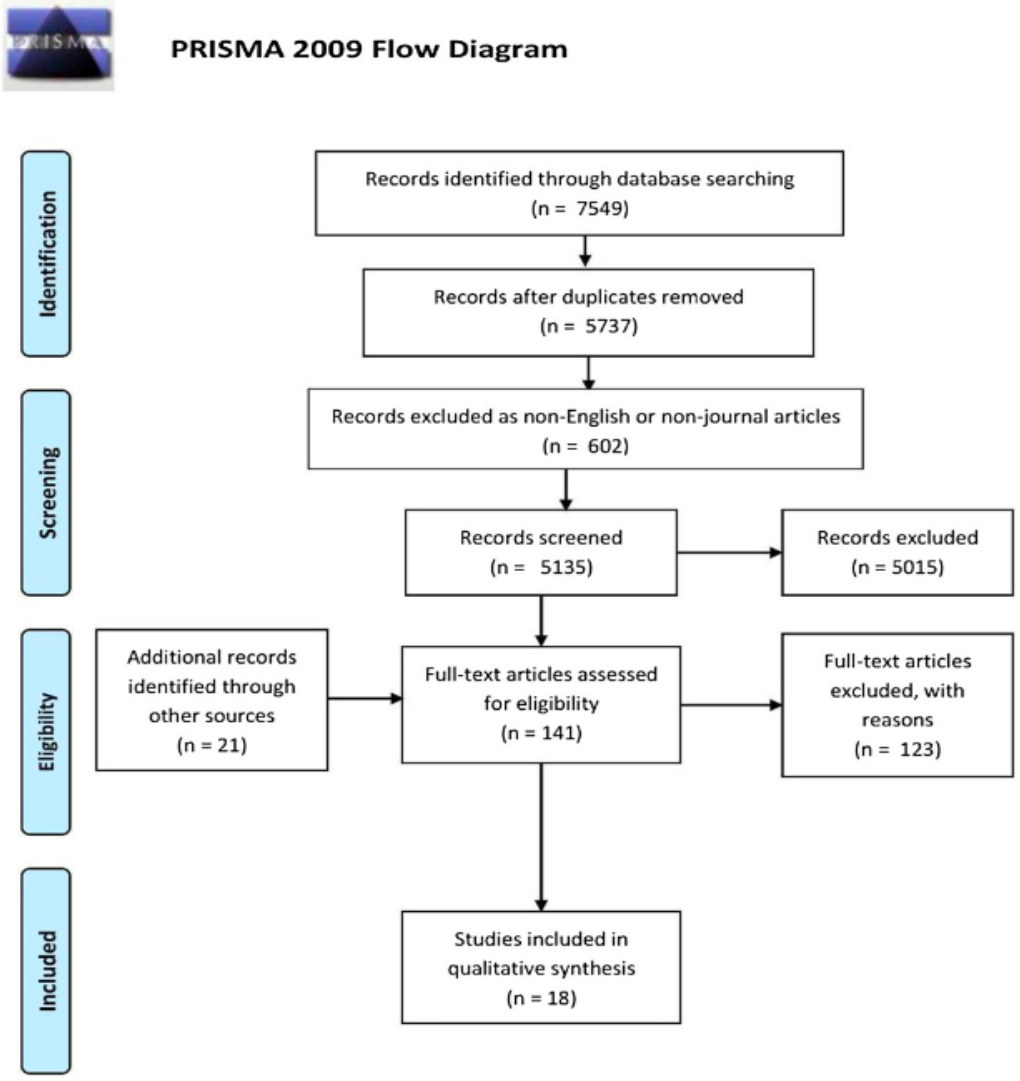


Figure 1 PRISMA 2009 flow diagram representing the selection process.

4. Screen and select the studies

- Data collection is organised using data extraction sheets. Here are some key suggestions from the Cochrane Handbook:
  - «Review authors are encouraged to develop outlines of tables and figures that will appear in the review to facilitate the design of data collection forms. The key to successful data collection is to construct easy-to-use forms and collect sufficient and unambiguous data that faithfully represent the source in a structured and organized manner.»

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

4. Screen and select: data to collect

Study methods

Study design:

- Parallel, factorial, crossover, cluster aspects of design for randomized trials, and/or study design features for non-randomized studies
- Single or multicentre study; if multicentre, number of recruiting centres

Recruitment and sampling procedures used (including at the level of individual participants and clusters/sites if relevant)
Enrolment start and end dates; length of participant follow-up
Details of random sequence generation, allocation sequence concealment, and masking for randomized trials, and methods used to prevent and control for confounding, selection biases, and information biases for non-randomized studies*
Methods used to prevent and address missing data*
Statistical analysis:
Unit of analysis (e.g. individual participant, clinic, village, body part)
Statistical methods used if computed effect estimates are extracted from reports, including any covariates included in the statistical model
Likelihood of reporting and other biases*
Source(s) of funding or other material support for the study
Authors' financial relationship and other potential conflicts of interest

Participants

Setting
Region(s) and country/countries from which study participants were recruited
Study eligibility criteria, including diagnostic criteria
Characteristics of participants at the beginning (or baseline) of the study (e.g. age, sex, comorbidity, socio-economic status)

Intervention

Description of the intervention(s) and comparison intervention(s), ideally with sufficient detail for replication:

- Components, routes of delivery, doses, timing, frequency, intervention protocols, length of intervention
- Factors relevant to implementation (e.g. staff qualifications, equipment requirements)
- Integrity of interventions (i.e. the degree to which specified procedures or components of the intervention were implemented as planned)
- Description of co-interventions
- Definition of 'control' groups (e.g. no intervention, placebo, minimally active comparator, or components of usual care)
- Components, dose, timing, frequency
- For observational studies: description of how intervention status was assessed; length of exposure, cumulative exposure

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

4. Screen and select: data to collect

Outcomes

For each pre-specified outcome domain (e.g. anxiety) in the systematic review:

- Whether there is evidence that the outcome domain was assessed (especially important if the outcome was assessed but the results not presented; see Chapter 13)
- Measurement tool or instrument (including definition of clinical outcomes or endpoints); for a scale, name of the scale (e.g. the Hamilton Anxiety Rating Scale), upper and lower limits, and whether a high or low score is favourable, definitions of any thresholds if appropriate
- Specific metric (e.g. post-intervention anxiety, or change in anxiety from baseline to a post-intervention time point, or post-intervention presence of anxiety (yes/no))
- Method of aggregation (e.g. mean and standard deviation of anxiety scores in each group, or proportion of people with anxiety)
- Timing of outcome measurements (e.g. assessments at end of eight-week intervention period, events occurring during the eight-week intervention period)
- Adverse outcomes need special attention depending on whether they are collected systematically or non-systematically (e.g. by voluntary report)

Results

For each group, and for each outcome at each time point: number of participants randomly assigned and included in the analysis; and number of participants who withdrew, were lost to follow-up or were excluded (with reasons for each)

Summary data for each group (e.g. 2×2 table for dichotomous data; means and standard deviations for continuous data)

Between-group estimates that quantify the effect of the intervention on the outcome, and their precision (e.g. risk ratio, odds ratio, mean difference)

If subgroup analysis is planned, the same information would need to be extracted for each participant subgroup

Miscellaneous

Key conclusions of the study authors

Reference to other relevant studies

Correspondence required

Miscellaneous comments from the study authors or by the review authors

(Cochrane Handbook for Systematic Reviews of Interventions, 2019)

4. Screen and select the studies: define measures

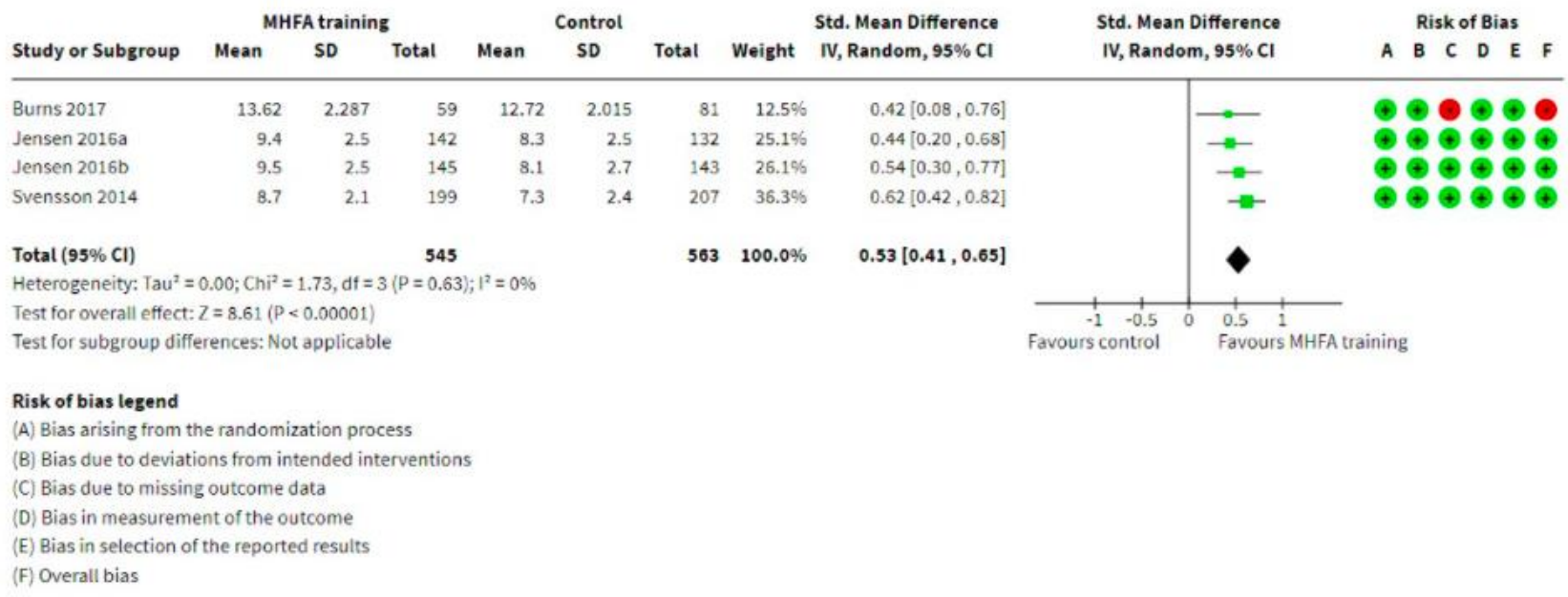
- At the end of data collection, it should be possible to:
  - Choose effect measures, for quantitative studies:
    - «Effect measures are either ratio measures (e.g. risk ratio, odds ratio) or difference measures (e.g. mean difference, risk difference)» (Cochrane Handbook, 2019)
    - Computing estimates of effect, after extracting data from studies
  - Define the meta-synthesis method (for qualitative studies)

5. Assess included studies: risk of bias

- Before proceeding to the synthesis of results, it is necessary to assess the quality of the studies, particularly for the risk of bias (ROB)
- At least two reviewers should conduct ROB assessment independently
- Several methods may be used. For example:
  - Rob 2 (Cochrane)
  - GRADE

5. Assess included studies: Risk of Bias 2

Figure 7.4.a Forest plot displaying RoB 2 risk-of-bias judgements for each randomized trial included in a meta-analysis of mental health first aid (MHFA) knowledge scores. Adapted from Morgan et al (2018).



Risk of bias legend
(A) Bias arising from the randomization process
(B) Bias due to deviations from intended interventions
(C) Bias due to missing outcome data
(D) Bias in measurement of the outcome
(E) Bias in selection of the reported results
(F) Overall bias

Rob 2 (Cochrane)

5. Assess included studies: GRADE

- «For authors of systematic reviews:
- The quality of evidence reflects the extent to which we are confident that an estimate of the effect is correct.
- Although the quality of evidence represents a continuum, the GRADE approach results in an assessment of the quality of a body of evidence in one of four grades» (GRADE Handbook)

5. Assess included studies: GRADE

Table 5.1: Quality of Evidence Grades	
Grade	Definition
High	We are very confident that the true effect lies close to that of the estimate of the effect.
Moderate	We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low	Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.
Very Low	We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

5. Assess included studies: GRADE

The GRADE approach to rating the quality of evidence begins with the study design (trials or observational studies) and then addresses five reasons to possibly rate down the quality of evidence and three to possibly rate up the quality. The subsequent sections of the handbook will address each of the factors in detail.

Table 5.2: Factors that can reduce the quality of the evidence

Factor	Consequence
Limitations in study design or execution (risk of bias)	↓ 1 or 2 levels
Inconsistency of results	↓ 1 or 2 levels
Indirectness of evidence	↓ 1 or 2 levels
Imprecision	↓ 1 or 2 levels
Publication bias	↓ 1 or 2 levels

5. Assess included studies: GRADE

Table 5.3: Factors that can increase the quality of the evidence	
Factor	Consequence
Large magnitude of effect	↑ 1 or 2 levels
All plausible confounding would reduce the demonstrated effect or increase the effect if no effect was observed	↑ 1 level
Dose-response gradient	↑ 1 level

Tabella 4 – Esempio di tabella delle evidenze, tratto da GRADEpro (www.gradepro.org)

Quality assessment							Summary of findings				Quality	Importance
							No of patients		Effect			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Receiving drug X	Control	Relative (95% CI)	Absolute		
Outcome A: Death												
5	randomized trial	serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	45/1715	76/1253	OR 0.42 (0.29 to 0.61)	29 fewer per 1,000	moderate	

6. Summarize the results

- The final synthesis may be:
  - Quantiative, using statistical methods (low level of etherogenity of results)
  - Quantitative, descriptive (high level of etherogeneity of results)
  - Qualitative (different techniques for meta-synthesis)

BFM
BIBLIOTECA FEDERATA DI MEDICINA
FERDINANDO ROSSI

UNIVERSITÀ
DEGLI STUDI
DI TORINO

7. References

- Cochrane Handbook for Systematic Reviews of Interventions. Wiley-Blackwell; 2019
- Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should i conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol. 2018;18(1):1–9
- Schünemann H, Brożek J, Guyatt G, Oxman AD. GRADE Handbook [Internet]. 2013. Available from: http://gdt.guidelinedevelopment.org/app/handbook/handbook.html