Synthesis without meta-analysis

Mhairi Campbell

MRC/CSO Social and Public Health Sciences Unit
University of Glasgow, UK

Session outline

1. Deciding whether or not to do a meta-analysis

2. Alternative methods to synthesis of quantitative effect data when meta-analysis is not conducted

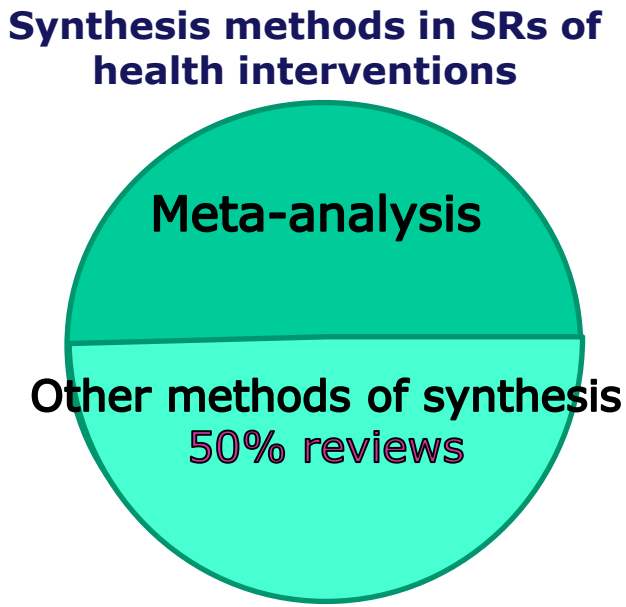
3. Transparency and rigour in synthesis which does not rely on meta-analysis

Session outline

1. Deciding whether or not to do a meta-analysis
2. Alternative methods to synthesis of quantitative effect data when meta-analysis is not conducted
3. Transparency and rigour in synthesis which does not rely on meta-analysis

To meta-analyse or not…

- Data synthesis (e.g. meta-analysis) is part of a systematic review
- Systematic review is **not** a method of synthesis
- 50% systematic reviews of RCT data included meta-analysis *(Ioannidis 2016, Millbank Quarterly)*



To meta-analyse or not…

- Very common
  ~~- lack of statistical expertise~~
  - lack of suitable data
- Different views about when not appropriate or useful to meta-analyse
  - heterogeneity
- Some say perform meta-analysis on appropriate groups and interpret cautiously

RESEARCH METHODOLOGY

Reasons or excuses for avoiding meta-analysis in forest plots

Heterogeneous data are a common problem in meta-analysis. **John Ioannidis**, **Nikolaos Patsopoulos**, and **Hannah Rothstein** show that final synthesis is possible and desirable in most cases

Some systematic reviews simply assemble the eligible studies without performing meta-analysis. This may be a legitimate choice. However, an interesting situation arises when reviews present forest plots (quantitative effects and uncertainty per study) but do not calculate a summary estimate (the diamond at the bottom). These reviews imply that it is important to visualise the quantitative data but final synthesis is inappropriate. For example, a review of sexual abstinence programmes for HIV prevention claimed that owing to "data unavailability, lack of intention-to-treat analyses, and heterogeneity in programme and trial designs… a statistical meta-analysis would be inappropriate."[1] As we discuss, options almost always exist for quantitative synthesis and sometimes they may offer useful insights. Reviewers and clinicians should be aware of these options, reflect carefully on their use, and understand their limitations.

Why meta-analysis is avoided

Of the 1739 systematic reviews that included at least one forest plot with at least two studies in issue 4 of the *Cochrane Database of Systematic Reviews* (2005), 135 reviews (8%) had 559 forest plots with no summary estimate.

The reasons provided for avoiding quantitative synthesis typically revolved around heterogeneity (table 1). The included studies were thought to be too different, either statistically or in clinical (including methodological) terms. Differences in interventions, metrics, outcomes, designs, participants, and settings were implied.

How large is too large heterogeneity?

This question of lumping versus splitting is difficult to answer objectively for clinical heterogeneity. Logic models based on the PICO (population-intervention–comparator-outcomes) framework may help to deal with the challenges of deciding what to include and what not. Still, different reviewers, readers, and clinicians may disagree on the (dis)similarity of interventions, outcomes, designs, participant characteristics, and settings.

No widely accepted quantitative measure exists to grade clinical heterogeneity. Nevertheless, it may be better to examine clinical differences in a meta-analysis rather than use them as a reason for not conducting one. For example, a review identified 40 trials of diverse interventions to prevent falls in elderly people.[2] Despite large diversity in the trials, the authors did a meta-analysis and also examined the effectiveness of different interventions. The analysis suggested that evidence was stronger for multifactorial risk assessment and management programmes and exercise and more inconclusive for environmental modifications and education.

Statistical heterogeneity can be measured—for example, by calculating I^2 and its uncertainty.[3-5] I^2, the proportion of variation between studies not due to chance, takes values from 0 to 100%. In the 22 forest plots including four or more studies that avoided synthesis because of heterogeneity, I^2 ranged between 35% and 98% with a median of 71% (figure). Yet, 86 of the 1011 forest plots where reviewers had no hesitation in performing meta-analysis had I^2 exceeding 71%.[5] The lower 95% confidence limit of I^2 was <25% in 11 of the 22 non-summarised forest plots—that is, for half of them we cannot exclude that statistical heterogeneity is limited. Therefore,

Table 1 | Reasons for not showing summary estimates in forest plots from systematic reviews in Cochrane database 2005 issue 4

Reason	No (%) of systematic reviews (n=135)*
Statistical heterogeneity too high	32 (24)
Different interventions compared	41 (30)
Different metrics or outcomes evaluated	26 (19)
Different metric of same outcome	7
Different outcome	20
Different study designs	21 (16)
Non-randomised studies	3

Alternative approaches to meta-analysis

Guidance: **conducting alternative synthesis**

- Cochrane Handbook Chapter 12: Synthesising and presenting findings using other methods (McKenzie & Brennan, 2021)

Guidance: **reporting synthesis without meta-analysis (SWiM)**

- Synthesis without meta-analysis (SWiM) in systematic reviews: a reporting guideline. (Campbell et al, 2020) BMJ

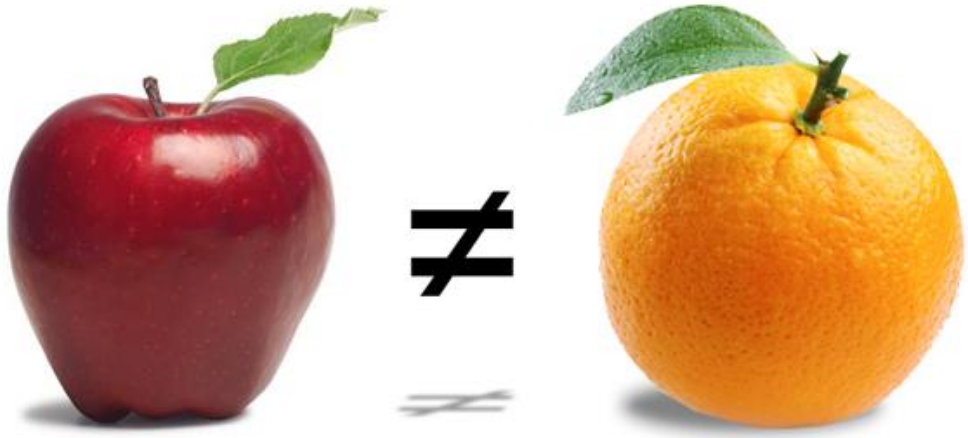
- **https://swim.sphsu.gla.ac.uk/**

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. Data presentation methods
8. Reporting results
9. Limitations of the synthesis methods

Conceptual/clinical heterogeneity

- Principles of synthesis: combining outcomes/ interventions etc that are conceptually similar



Synthesis heterogeneous data: level of similarity or commonality may vary

If you are synthesising it is implied that there is a level of commonality to justify the synthesis- this needs to be made clear

Grouping studies for synthesis

- Deciding how to group:
  - Populations, interventions, comparisons, outcomes (PICO)
  - Study designs
  - Risk of bias
- What will be useful to decision makers
- Important to clearly explain:
  - how studies are grouped
  - justify the grouping

Session outline

1. Deciding whether or not to do a meta-analysis
2. Alternative methods to synthesis of quantitative effect data when meta-analysis is not conducted
3. Transparency and rigour in synthesis which does not rely on meta-analysis

Standardised metric for synthesis

Synthesising:

- at some level something common to the studies/data (can't combine a mean with an odds ratio)
- in meta-analysis synthesising standardised effect sizes

Standardised metrics in quantitative studies

- effect sizes (unable to meta-analyse)
- direction of effect
- p values

Standardised metric

Effect sizes

- Examples: risk ratios, odds ratios, risk differences, mean differences, standardised mean differences, ratio of means

Standardised metric

Effect sizes

- Examples: risk ratios, odds ratios, risk differences, mean differences, standardised mean differences, ratio of means

Direction of effect

- Favour intervention / treatment
- Favour control
- No effect

Standardised metric

Effect sizes

- Examples: risk ratios, odds ratios, risk differences, mean differences, standardised mean differences, ratio of means

Direction of effect

- Favour intervention / treatment
- Favour control
- No effect

P values

- One-sided P values
- P values must all reflect same directional hypothesis

Standardised metric

Synthesis method

Standardised metric

- effect sizes
  (unable to meta-analyse)
- direction of effect
- p values

Synthesis method

- summarise effect sizes
- vote counting of studies
- combine p values

Alternative methods of synthesis

- **Summarise effect sizes**
  - Use when have estimates of intervention effect (but can't meta-analyse)
  - Descriptive statistics such as median, interquartile range, range

- **Vote counting based on direction of effect**
  - Use when have only direction of effect of studies, or no consistent effect measure or data reported across studies
  - Benefit or harm based on direction of effect (not statistical significance)

- **Combine p values**
  - Use when have p values and direction of effect of studies, outcomes and statistical tests differ across studies, or studies report non-parametric test results
  - Use (or convert to) 1-sided p values

Alternative methods of synthesis

McKenzie and Brennan (2021) Chapter 12: Synthesizing and presenting findings using other methods. Cochrane Handbook

Cochrane Methods Support Unit webinar recording

Development of materials to facilitate implementation of methods for presentation and statistical synthesis when data are not amenable to meta-analysis

https://www.youtube.com/watch?v=KPvAgxXZ1Qc

Questions answered by different synthesis methods

- **Meta-analysis of standardised effect sizes:** What is the average effect?

Other methods

- **Summarizing effect sizes:** What is the range and distribution of effects?

- **Vote counting based on direction of effect:** Is there any evidence of an effect?

- **Combining p values:** Is there evidence that there is an effect in at least one study?

Session outline

1. Deciding whether or not to do a meta-analysis
2. Alternative methods to synthesis of quantitative effect data when meta-analysis is not conducted
3. Transparency and rigour in synthesis which does not rely on meta-analysis

Transparency when reporting synthesis

Lack of reporting methods

Reduces ability to assess what was done to synthesise the data

Do not know whether can trust the review findings

- Even if the methods are robust – if not clearly reported, the review user is unaware robust methods have been used
- Results in lack of trust in otherwise high quality reviews

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. Data presentation methods
8. Reporting results
9. Limitations of the synthesis methods

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. **Grouping of studies for synthesis**
2. **Describe standardized metric and transformation methods used**
3. **Describe synthesis methods**
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. Data presentation methods
8. Reporting results
9. Limitations of the synthesis methods

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. **Data presentation methods**
8. Reporting results
9. Limitations of the synthesis methods

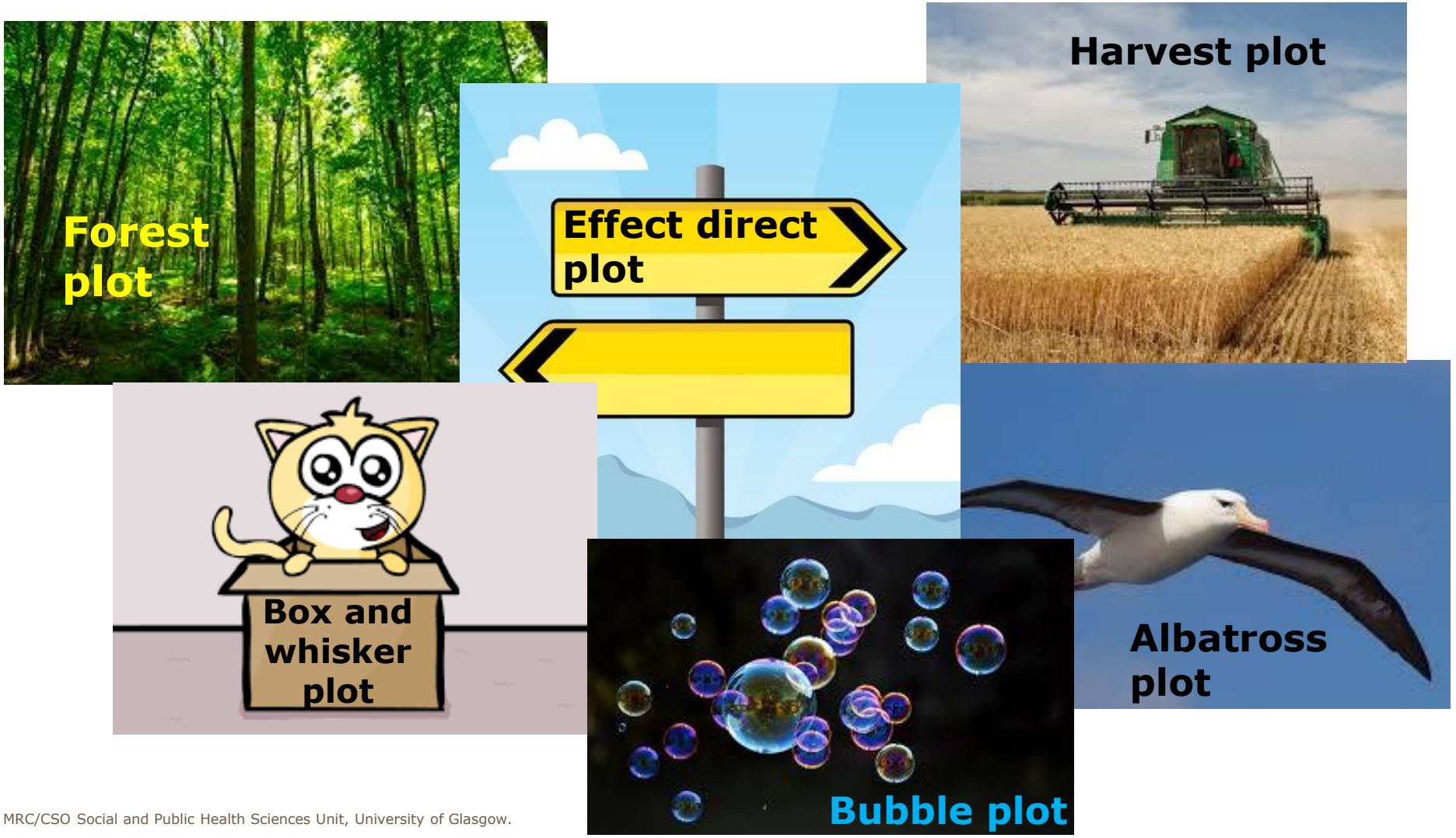
Transparent links between data & synthesis findings: tables

Key considerations when designing data tables

- Data tabulated along with key **relevant** characteristics
  - population
  - outcome
  - study design and/or quality/risk of bias
  - study size, location etc.as relevant and as space allows

  - Reflect the order/grouping of the synthesis to promote transparency (more helpful than alphabetical lists of studies)
  - Allow comparison across studies in relevant groupings

Data presentation options



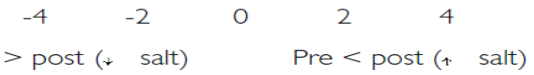
Analysis 3.1. Comparison 3 Mean change in salt intake (grams/d) from pre-intervention to post-intervention - WOMEN ONLY, Outcome 1 Salt intake in grams per day.

Review: Population-level interventions in government jurisdictions for dietary sodium reduction

Comparison: 3

Outcome: 1 Sa

Study or subgroup	Post-intervention N	Mean(SD)	Pre-intervention N	Mean(SD)	Mean Difference IV,Random,95% CI	Mean Difference IV,Random,95% CI
Austria	232	7.6 (3.11)	1345	7.6 (2.81)		0.0 [-0.43, 0.43]
Canada	5612	7.41 (1.44)	2566	5 (2.71)		2.41 [2.30, 2.52]
China	3605	11.43 (6.35)	3584	12.19 (6.86)		-0.76 [-1.07, -0.45]
Finland	220	9.5 (3.78)	327	10.4 (4.15)		-0.90 [-1.57, -0.23]
France	1082	6.65 (1.88)	732	6.93 (2.16)		-0.28 [-0.47, -0.09]
Netherlands	188	7.42 (2.49)	180	7.89 (2.96)		-0.47 [-1.03, 0.09]
Switzerland	742	7.8 (3.3)	95	7.3 (2.9)		0.50 [-0.13, 1.13]
United Kingdom - England	2343	4.12 (2.47)	933	4.73 (3.12)		-0.61 [-0.83, -0.39]
United Kingdom - Great Britain	398	7.66 (4.77)	891	8.1 (3.88)		-0.44 [-0.97, 0.09]
United Kingdom - Scotland	598	5.7 (4.37)	640	6.1 (3.87)		-0.40 [-0.86, 0.06]
United States	267	7.55 (3.84)	604	6.79 (6.18)		0.76 [0.09, 1.43]



Source: McLaren et al 2016 https://doi.org/10.1002/14651858.CD010166.pub2

Effect direction plot

Author Year	Study design	Study quality	Final Sample Int/Cont	Time since interv'n	Housing condition	General health	Respiratory
Intervention: Warmth & Energy Efficiency improvements (post 1980)							
Osman et al 2010	RCT	A	45/51	5 months	▲	▼ a	▼4 a
Howden-Chapman et al 2008 **	RCT	A	175/174	4-5 months	▲	▲	▲11
Braubach et al 2008	CBA	A	~210/165	5-8 months	Λ	Λ	Λ
Barton et al 2007 *, ****	RCT	A	193/254	3-10 months	◄►	<>	◄► 3
Howden-Chapman et al 2007 *	RCT	A	1689/1623	<1 year	▲	▲3	▲5
Platt et al 2007	CBA	A	1281/1084	1-2 years	▲	▲3	◄►4
Lloyd et al 2008	CBA	B	9/27	1-2.5 years			
Shortt et al 2007	CBA	B	46/54	1-3.5 years	▲ b		◄►3
Somerville et al 2000 **	UBA	B	72	3 months	▲ b		▲ b 7
Hopton et al 1996 **	CBA	B	55/77	5-11 months	▲ b		◄►2
Allen 2005	UBA	C	16	<1 year	Λ b		
Allen 2005 a	UBA	C	24	<3 years	Λ b	▼	
Health Action Kirklees 2005	R	C	102	2-8 months	Λ b		
Iversen et al 1986	CBA	C	106/535	3-6 months	Λ b		▲

Source: Thomson et al 2013 https://doi.org/10.1002/14651858.CD008657.pub2

Effect direction plot

Author Year	Study design	Study quality	Final Sample Int/Cont	Time since interv'n	Housing condition	General health	Respiratory
Intervention: Warmth & Energy Efficiency improvements (post 1980)							
Osman et al 2010	RCT	A	45/51	5 months	▲	▼ a	▼4 a
Howden-Chapman et al 2008 **	RCT	A	175/174	4-5 months	▲	▲	▲11
Braubach et al 2008	CBA	A	~210/165	5-8 months	Λ	Λ	Λ
Barton et al 2007 *, ****	RCT	A	193/254	3-10 months	◄►	<>	◄► 3
Howden-Chapman et al 2007 *	RCT	A	1689/1623	<1 year	▲	▲3	▲5
Platt et al 2007	CBA	A	1281/1084	1-2 years	▲	▲3	◄►4
Lloyd et al 2008	CBA	B	9/27	1-2.5 years			
Shortt et al 2007	CBA	B	46/54	1-3.5 years	▲ b		◄►3
Somerville et al 2000 **	UBA	B	72	3 months	▲ b		▲ b 7
Hopton et al 1996 **	CBA	B	55/77	5-11 months	▲ b		◄►2
Allen 2005	UBA	C	16	<1 year	Λ b		
Allen 2005 a	UBA	C	24	<3 years	Λ b	▼	
Health Action Kirklees 2005	R	C	102	2-8 months	Λ b		
Iversen et al 1986	CBA	C	106/535	3-6 months	Λ b		▲

Source: Thomson et al 2013 https://doi.org/10.1002/14651858.CD008657.pub2

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. Data presentation methods
8. **Reporting results**
9. Limitations of the synthesis methods

Item 8: Reporting results

Information that should be included in the text summary:

- **Clear reporting of the nature of the question(s) addressed**
  For example, based on effect direction rather than effect size
- **Standardised metric(s) and synthesis method(s) used**
- **Reference the studies used in the synthesis**
  For each outcome/comparison

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. Criteria used to prioritize results for summary and synthesis
5. Investigation of heterogeneity in reported effects
6. Certainty of evidence
7. Data presentation methods
8. Reporting results
9. **Limitations of the synthesis methods**

Item 9: Limitations of the synthesis

- Standardised metric used
- Synthesis method used
- Changes to groups used in synthesis

Implications for what questions can be answered and how synthesis can be interpreted

For example

- if the standardised metric used is **direction of effect**:

  Review question is about ‘is there any evidence of an effect?’

  rather than ‘what is the average intervention effect size?’

- lack of studies or reported outcomes in studies may change how the synthesis is structured - how the studies are grouped

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. **Criteria used to prioritize results for summary and synthesis**
5. **Investigation of heterogeneity in reported effects**
6. **Certainty of evidence**
7. Data presentation methods
8. Reporting results
9. Limitations of the synthesis methods

Reading list

Ioannidis et al. (2008). Reasons or excuses for avoiding meta-analysis in forest plots. BMJ, 336(7658), 1413-1415.

McKenzie & Brennan (2021). Synthesising and presenting findings using other methods (Chapter 12). Cochrane Cochrane handbook for systematic reviews of interventions (eds: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V), 321-347.

Campbell et al. (2020) Synthesis without meta-analysis (SWiM) in systematic reviews: a reporting guideline. BMJ, 368

SWiM webpage: **https://swim.sphsu.gla.ac.uk/**

Cochrane Methods Support Unit webinar recording

Development of materials to facilitate implementation of methods for presentation and statistical synthesis when data are not amenable to meta-analysis **https://www.youtube.com/watch?v=KPvAgxXZ1Qc**

Boon, M. H., & Thomson, H. (2021). The effect direction plot revisited: Application of the 2019 Cochrane Handbook guidance on alternative synthesis methods. Research synthesis methods, 12(1), 29-33.

SWiM: guidance for reporting Synthesis Without Meta-analysis

1. Grouping of studies for synthesis
2. Describe standardized metric and transformation methods used
3. Describe synthesis methods
4. **Criteria used to prioritize results for summary and synthesis**
5. **Investigation of heterogeneity in reported effects**
6. **Certainty of evidence**
7. Data presentation methods
8. Reporting results
9. Limitations of the synthesis methods

Item 4: Criteria used to prioritize results for summary and synthesis

Some study findings may be prioritized over others.

If studies are given priority in the synthesis or results, the criteria for this should be reported.

- study design (e.g. only randomised trials)
- risk of bias assessment (e.g. only studies at a low risk of bias)
- sample size
- relevance of the evidence addressing the review question (e.g. outcome, population/context or intervention)

Item 5: Investigation of heterogeneity in reported effects

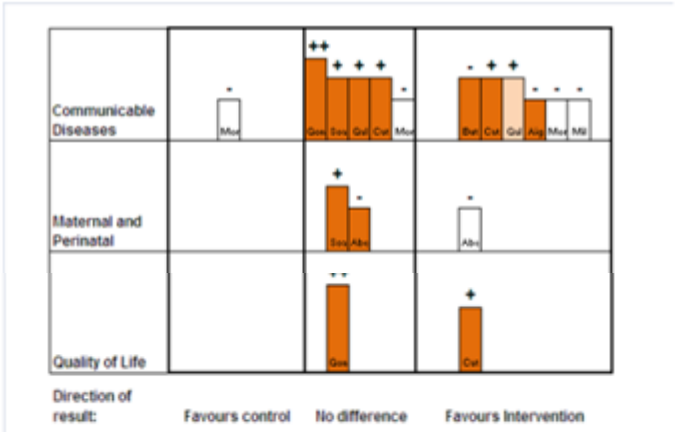
Methods to examine differences in results – when statistical methods such as meta-regression are not possible

Visual examination of tables ordered by modifiers, e.g.:

- study design
- subpopulations (e.g. sex, age)
- intervention components
- context/setting

Graphs such as harvest plots

Figure 5. Harvest plot to show health and QoL findings across main and supporting studies.



Item 6: Certainty of evidence

Assess certainty of the evidence, considering:

- risk of bias
- precision (confidence intervals, or number of studies and participants)
- consistency of effects across studies
- how directly studies address review question

GRADE (Grading of Recommendations, Assessment, Development and Evaluations)

Guyatt et al. GRADE Working Group. (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.

Guyatt et al. (2011) GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables.

Reading list

- Ioannidis et al. (2008). Reasons or excuses for avoiding meta-analysis in forest plots. BMJ, 336(7658), 1413-1415.
- McKenzie & Brennan (2021). Synthesising and presenting findings using other methods (Chapter 12). Cochrane Cochrane handbook for systematic reviews of interventions (eds: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, Welch V), 321-347.
- Campbell et al. (2020) Synthesis without meta-analysis (SWiM) in systematic reviews: a reporting guideline. BMJ, 368
- SWiM webpage: **https://swim.sphsu.gla.ac.uk/**
- Cochrane Methods Support Unit webinar recording
- Development of materials to facilitate implementation of methods for presentation and statistical synthesis when data are not amenable to meta-analysis **https://www.youtube.com/watch?v=KPvAgxXZ1Qc**
- Guyatt et al. (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.
- Guyatt et al. (2011) GRADE guidelines: 1. Introduction - GRADE evidence profiles and summary of findings tables.
- Boon, M. H., & Thomson, H. (2021). The effect direction plot revisited: Application of the 2019 Cochrane Handbook guidance on alternative synthesis methods. Research synthesis methods, 12(1), 29-33.