Quantitative Synthesis

Chapter 1. Decision to combine trials

Prepared for:
The Agency for Healthcare Research and Quality (AHRQ)
Training Modules for Systematic Reviews Methods Guide
www.ahrq.gov
Purpose:
- To consolidate and update guidance from prior methods guides
- Focuses on *Comparative Effectiveness Reviews (CERs)*:
  - Systematic reviews comparing effectiveness and harms of alternative clinical options
  - Help clinicians, policy-makers, patients make informed treatment choices
- Focuses on interventional studies, not diagnostic, individual-level, or observational studies
• Quantitative synthesis (meta-analysis) is a critical component of comparative effectiveness reviews (CERs).
• Quantitative synthesis should be conducted transparently, consistently, with methodology explicitly reported.
• This guide supports this process, but is not a comprehensive review or a text.
• Addresses the issues commonly encountered conducting CERs, in the order they occur.
Systematic review process overview

Prepare Topic
- Refine Topic
- Develop analytic framework

Search for and Select Studies for Inclusion
- Identify study eligibility criteria
- Search for relevant studies
- Select evidence for inclusion

Extract Data from Studies

Analyze and Synthesize Studies
- Assess the quality of individual studies
- Assess applicability
- Present findings
  - Synthesize quantitative data
- Grade strength of evidence

Report Systematic Review
Learning objective: Describe the basic principles of combining data and when this is appropriate.
The first fundamental question in quantitative synthesis is, “Is it appropriate to pool the results of the identified studies?”

Many factors must be considered when deciding whether to combine studies in meta-analysis:

- Clinical factors
- Methodological factors
- Statistical factors

Studies must be reasonably similar to be pooled in meta-analysis.

Decision to combine trials

• Assessing whether studies are reasonably similar requires considering heterogeneity.

• **Clinical Heterogeneity** and **Methodological Heterogeneity** describe the variations in the population, intervention and study factors.²,³

  ► **Statistical heterogeneity** is discussed later

• A **pooling decision tree** guides consideration of these factors in the decision to combine trials.

---


Pooling decision tree: Step A

A. Identify groups of included studies that are clinically and methodological similar enough to pool.

Are there trials that are similar enough for an average effect to be meaningful?

Yes → No → Do not pool
Step A: Explore clinical and methodological heterogeneity

- Step A: Exploring clinical and methodological heterogeneity of identified studies.
  - The goal is to identify groups of trials that are similar enough that an average effect is a sensible summary
  - There is no objective measure or consensus standard
- It is recommended to explore variability:
  - Clinical intervention
  - Control condition
  - Participants
  - Study design
  - Outcome
  - Follow-up time
Heterogeneity

- **Clinical heterogeneity** relates to participants, interventions, outcomes, study setting.

- **Methodological heterogeneity** relates to study methods (e.g., study design, measures, conduct).
  - E.g., can individually-randomized studies be combined with cluster-randomized studies?
  - Often yes, if adjusted for clustering\(^4\)
  - If results systematically differ between the two, then subgroups are more appropriate

---

Pooling decision tree: Steps B and C

B. Is evidence best characterized by a small subset (e.g., 1-4) of large or best-quality trials?

Yes: Consider not pooling, or pooling only “best evidence” trials

No:

C. Explore the risk of misleading results if the trials are pooled through tables, forest plots, and other preliminary analyses.

- Are there wide-ranging effect sizes suggesting both possible benefit and possible harm?
- Is there suspicion of reporting bias?
- Is there evidence of small studies effect?

Yes: Consider not pooling
• Step B: Best evidence or all evidence?
  ► Cherry-picking studies to include in meta-analysis can introduce bias. Decisions to exclude trials because of subjective quality assessments must be justified carefully.

• Step C: Will combining evidence produce misleading results?
  ► An unmeasured source of heterogeneity may also explain widely varying results.
  ► If an intervention truly benefits some patients and harms others, a summary estimate of effect may be meaningless.
Publication bias, reporting bias

• If trials with null results are not published, this results in *publication bias*.

• If trial results are published only in part, this results in *reporting bias*.

• These missing results introduce bias into meta-analysis and reduce precision.\(^5\)
  - Use standard tests (e.g., Egger test) to examine small study effects if number of trials permits (i.e., >10)
  - Do not report meta-analysis results that result from small study effects due to bias. (Instead, describe the evidence you are relying on to determine that the results would be biased.)

---

Pooling decision tree: Steps D and E

D. Pooling a small number of studies (e.g., <10)?
   Yes

E. Extra caution is required. High risk situations include:
   - If an outcome is a rare event that the included trials are underpowered to detect in sufficient numbers for a valid result
   - If the body of evidence is limited to small studies
   - If there are wide-ranging effect sizes with an insufficient number of trials for an accurate determination of the standard error of the pooled estimate, so coverage of the confidence intervals of the pooled estimate may substantially differ from 95% coverage

Low risk
   Yes, pool

High risk
   Consider not pooling
Steps D and E: Small number of studies

- **Steps D & E: Consider these high risk of bias situations when pooling a small number of studies:**
  - **Small sample sizes:** Results from small trials are less reliable even when risk of bias is low.\(^6\)
    - If sample sizes are small, pooled effects are less likely to reflect true effects
  - **Rare outcomes:** If included trials are underpowered for a valid result, do not pool.\(^7\)
    - Consider calculating the Optimal Information Size to see if combined studies have sufficient power
  - **Wide-ranging effect sizes:** Statistical heterogeneity is underestimated with few studies (especially <7).\(^8\)

---


Pooling decision tree: Step F

F(1). Non-negligible statistical heterogeneity present?
  - No
    - No further exploration of heterogeneity
  - Yes
    - Explore heterogeneity

F(2). Result shows potentially important statistical heterogeneity (e.g., I2 >50%)?
  - No
    - Report pooled results
  - Yes
    - Consider not reporting unless special methods were used to ensure accuracy of the CI in the presence of high statistical heterogeneity.
Step F: Statistical heterogeneity

• Step F: Consider *statistical heterogeneity* next if clinical and methodological heterogeneity are not concerns.
  - Chapter 4 covers statistical heterogeneity in detail

• Statistical heterogeneity is assessed by conducting a preliminary meta-analysis. ⁹
  - Then decide whether meta-analysis results should be presented, or just forest plots without pooled results
  - Methods for assessing: $I^2$, Cochrane’s Q, and $\tau^2$

---

Assessing statistical heterogeneity

- $I^2$: proportion of total variance in pooled trials due to inter-study variance (i.e., not random).
  - 0-40%: might not be important; 30-60%: may represent moderate heterogeneity; 50-90%: may represent substantial heterogeneity; 75-100%: considerable heterogeneity
  - Critiques of $I^2$: Underestimates heterogeneity in random effects models; increases as sample size increases; unreliable if too few studies

- Cochrane’s $Q$ and $\tau^2$: no standardized scale.

- No method is perfect and no method can make the determination of whether to pool alone.

Pooling decision tree (1)

A. Identify groups of included studies that are clinically and methodological similar enough to pool. Are there trials that are similar enough for an average effect to be meaningful?

- Yes
- No

- No: Do not pool
- Yes

B. Is evidence best characterized by a small subset (e.g., 1-4) of large or best-quality trials?

- Yes
- No

- Yes: Consider not pooling, or pooling only “best evidence” trials
- No

C. Explore the risk of misleading results if the trials are pooled through tables, forest plots, and other preliminary analyses.

- Are there wide-ranging effect sizes suggesting both possible benefit and possible harm?
- Is there suspicion of reporting bias?
- Is there evidence of small studies effect?

- Yes: Consider not pooling
- No
Pooling decision tree (2)

E. Extra caution is required. High risk situations include:
- If an outcome is a rare event that the included trials are underpowered to detect in sufficient numbers for a valid result
- If the body of evidence is limited to small studies
- If there are wide-ranging effect sizes with an insufficient number of trials for an accurate determination of the standard error of the pooled estimate, so coverage of the confidence intervals of the pooled estimate may substantially differ from 95% coverage

D. Pooling a small number of studies (e.g., <10)?

No

F(1). Non-negligible statistical heterogeneity present?

Yes

F(2). Result shows potentially important statistical heterogeneity (e.g., I² >50%)?

No

Low risk

Yes, pool

Consider not pooling

Report pooled results

Yes

Explore heterogeneity

Consider not reporting unless special methods were used to ensure accuracy of the CI in the presence of high statistical heterogeneity

No

No further exploration of heterogeneity

High risk

Yes

High risk

Consider not pooling

Report pooled results

No
Use the pooling decision tree (shown in slides 19 and 20) when deciding whether to combine data.
• This presentation was prepared by Jonathan Snowden, Ph.D.
• The presentation is based on the chapter entitled “Decision to Combine Trials” in the Methods Guide for Comparative Effectiveness Reviews (available at: https://doi.org/10.23970/AHRQEPCMETHGUIDE3


