Introduction to meta-analysis

Presented by Catrin Tudur Smith and Sarah Nevitt

Cochrane Colloquium 2018
Steps of a Cochrane Review

1. define the question
2. plan eligibility criteria
3. plan methods
4. search for studies
5. apply eligibility criteria
6. collect data
7. assess studies for risk of bias
8. analyse and present results
9. interpret results and draw conclusions
10. improve and update review
Session outline

• principles of meta-analysis
• steps in a meta-analysis
• presenting your results

See Chapter 9 of the Handbook
Source: Jo McKenzie & Miranda Cumpston
What is a meta-analysis?

- combines the results from two or more studies
- estimates an ‘average’ or ‘common’ effect
- optional part of a systematic review

Source: Julian Higgins
Why perform a meta-analysis?

- quantify treatment effects and their uncertainty
- increase power
- increase precision
- explore differences between studies
- settle controversies from conflicting studies
- generate new hypotheses

Source: Julian Higgins
When not to do a meta-analysis

- mixing apples with oranges
  - each included study must address same question
    - consider comparison and outcomes
    - requires your subjective judgement
    - combining a broad mix of studies answers broad questions
  - answer may be meaningless and genuine effects may be obscured if studies are too diverse

Source: Julian Higgins
When not to do a meta-analysis

- **garbage in – garbage out**
  - a meta-analysis is only as good as the studies in it
  - if included studies are biased:
    - meta-analysis result will also be incorrect
    - will give more credibility and narrower confidence interval
  - if serious reporting biases present:
    - unrepresentative set of studies may give misleading result

Source: Julian Higgins
When can you do a meta-analysis?

- more than one study has measured an effect
- the studies are sufficiently similar to produce a meaningful and useful result
- the outcome has been measured in similar ways
- data are available in a format we can use
Session outline

- principles of meta-analysis
- steps in a meta-analysis
- presenting your results
Steps in a meta-analysis

- identify comparisons to be made
- identify outcomes to be reported and statistics to be used
- collect data from each relevant study
- combine the results to obtain the summary of effect
- explore differences between the studies
- interpret the results
Selecting comparisons

Hypothetical review: Caffeine for daytime drowsiness

- caffeinated coffee vs decaffeinated coffee

- break your topic down into pair-wise comparisons
- each review may have one or many
- use your judgement to decide what to group together, and what should be a separate comparison
Selecting outcomes & effect measures

Hypothetical review: Caffeine for daytime drowsiness

- caffeinated coffee vs decaffeinated coffee

- asleep at end of trial (RR)
- irritability (MD/SMD)
- headaches (RR)

- for each comparison, select outcomes
- for each outcome, select an effect measure
  - may depend on the available data from included studies
Common types of outcome data

(1) Binary (or dichotomous) e.g. Survival status (Alive, Dead)
   - For revman: Enter number of participants with events and total number of participants in experimental and control groups.

(2) Continuous e.g. blood pressure measurement
   - For revman: Enter mean, standard deviation and number of participants in experimental and control groups
Types of effect measure

(1) Binary (or dichotomous) data
   Risk Ratio (or Relative Risk)
   Odds Ratio
   Risk Difference

RR and OR are ratio measures - the ‘null’ value is 1

(2) Continuous data:
   Mean Difference
   Standardised Mean Difference

RD, MD and SMD are difference measures – the ‘null’ value is 0
Calculating the summary result

- collect a summary statistic from each contributing study
- how do we bring them together?
  - treat as one big study – add intervention & control data?
    - breaks randomisation, will give the wrong answer
  - simple average?
    - weights all studies equally – some studies closer to the truth
  - weighted average
Weighting studies

- more weight to the studies which give more information
  - more participants, more events, narrower confidence interval
  - calculated using the effect estimate and its variance

- inverse-variance method:

  \[
  \text{weight} = \frac{1}{\text{variance of estimate}} = \frac{1}{SE^2}
  \]

  \[
  \text{pooled estimate} = \frac{\text{sum of (estimate} \times \text{weight})}{\text{sum of weights}}
  \]
For example

<table>
<thead>
<tr>
<th>Headache</th>
<th>Caffeine</th>
<th>Decaf</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amore-Coffea 2000</td>
<td>2/31</td>
<td>10/34</td>
<td></td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10/40</td>
<td>9/40</td>
<td></td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12/53</td>
<td>9/61</td>
<td></td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3/15</td>
<td>1/17</td>
<td></td>
</tr>
<tr>
<td>Norscafe 1998</td>
<td>19/68</td>
<td>9/64</td>
<td></td>
</tr>
<tr>
<td>Oohlahlazza 1998</td>
<td>4/35</td>
<td>2/37</td>
<td></td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8/35</td>
<td>6/37</td>
<td></td>
</tr>
</tbody>
</table>
For example

<table>
<thead>
<tr>
<th>Headache</th>
<th>Caffeine</th>
<th>Decaf</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amore-Coffea 2000</td>
<td>2/31</td>
<td>10/34</td>
<td>6.6%</td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10/40</td>
<td>9/40</td>
<td>21.9%</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12/53</td>
<td>9/61</td>
<td>22.2%</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3/15</td>
<td>1/17</td>
<td>2.9%</td>
</tr>
<tr>
<td>Norscafe 1998</td>
<td>19/68</td>
<td>9/64</td>
<td>26.4%</td>
</tr>
<tr>
<td>Oohlahlazza 1998</td>
<td>4/35</td>
<td>2/37</td>
<td>5.1%</td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8/35</td>
<td>6/37</td>
<td>14.9%</td>
</tr>
</tbody>
</table>
Meta-analysis options

• for dichotomous or continuous data
  • inverse-variance
    • straightforward, general method
• for dichotomous data only
  • Mantel-Haenszel (default)
    • good with few events – common in Cochrane reviews
    • weighting system depends on effect measure
  • Peto
    • for odds ratios only
    • good with few events and small effect sizes (OR close to 1)
Meta-analysis options
Session outline

- principles of meta-analysis
- steps in a meta-analysis
- presenting your results
A forest of lines
Forest plots

Headache at 24 hours

- headings explain the comparison
Forest plots

Headache at 24 hours

- list of included studies
### Forest plots

**Headache at 24 hours**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Caffeinated coffee</th>
<th>Decaffeinated coffee</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Amore-Coffea 2000</td>
<td>2</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10</td>
<td>40</td>
<td>9</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12</td>
<td>53</td>
<td>9</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Norscave 1998</td>
<td>19</td>
<td>68</td>
<td>9</td>
</tr>
<tr>
<td>Oohlahlazzza 1998</td>
<td>4</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8</td>
<td>35</td>
<td>6</td>
</tr>
</tbody>
</table>

**Total (95% CI)**
- Caffeinated: 277 events, 58 total events
- Decaffeinated: 290 events, 46 total events

**Heterogeneity:**
- Chi² = 8.58, df = 6 (P = 0.26), I² = 30%

**Test for overall effect:** Z = 1.73 (P = 0.08)

- raw data for each study
Forest plots

Headache at 24 hours

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Caffeinated coffee</th>
<th>Decaffeinated coffee</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amore-Coffea 2000</td>
<td>2</td>
<td>10</td>
<td>0.22 [0.05, 0.92]</td>
</tr>
<tr>
<td>Delicioza 2004</td>
<td>10</td>
<td>9</td>
<td>1.11 [0.51, 2.44]</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12</td>
<td>9</td>
<td>1.53 [0.70, 3.35]</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3</td>
<td>1</td>
<td>3.40 [0.39, 29.31]</td>
</tr>
<tr>
<td>Norscape 1998</td>
<td>19</td>
<td>9</td>
<td>1.99 [0.97, 4.07]</td>
</tr>
<tr>
<td>Oohlalhaza 1998</td>
<td>4</td>
<td>2</td>
<td>2.11 [0.41, 10.83]</td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8</td>
<td>6</td>
<td>1.41 [0.54, 3.65]</td>
</tr>
</tbody>
</table>

Total (95% CI)          | 277                | 290                  | 1.38 [0.96, 2.00]           |

Total events: 58 vs 46

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 29%
Test for overall effect: Z = 1.73 (P = 0.08)

- total data for all studies
Forest plots

Headache at 24 hours

- weight given to each study
## Forest plots

### Headache at 24 hours

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Caffeinated coffee</th>
<th>Decaffeinated coffee</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
<td>Total</td>
</tr>
<tr>
<td>Amore-Cofea 2000</td>
<td>2</td>
<td>31</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10</td>
<td>40</td>
<td>9</td>
<td>40</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12</td>
<td>53</td>
<td>9</td>
<td>61</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3</td>
<td>15</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Norscafe 1998</td>
<td>19</td>
<td>68</td>
<td>9</td>
<td>64</td>
</tr>
<tr>
<td>Oohlahlaza 1998</td>
<td>4</td>
<td>35</td>
<td>2</td>
<td>37</td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8</td>
<td>35</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td>277</td>
<td></td>
<td>290</td>
<td></td>
</tr>
<tr>
<td><strong>Total events</strong></td>
<td>58</td>
<td></td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>

**Heterogeneity:** Chi² = 8.58, df = 6 (P = 0.20); I² = 30%

**Test for overall effect:** Z = 1.73 (P = 0.08)

- effect estimate for each study, with CI
### Forest plots

**Headache at 24 hours**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Caffeinated coffee</th>
<th>Decaffeinated coffee</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Events</td>
<td>Total</td>
<td>Events</td>
</tr>
<tr>
<td>Amore-Coffea 2000</td>
<td>2</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10</td>
<td>40</td>
<td>9</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12</td>
<td>53</td>
<td>9</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3</td>
<td>15</td>
<td>1</td>
</tr>
<tr>
<td>Norscafe 1998</td>
<td>19</td>
<td>68</td>
<td>9</td>
</tr>
<tr>
<td>Oohlahlzaza 1998</td>
<td>4</td>
<td>35</td>
<td>2</td>
</tr>
<tr>
<td>Piazza-Alerta 2003</td>
<td>8</td>
<td>35</td>
<td>6</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>277</strong></td>
<td></td>
<td><strong>290</strong></td>
</tr>
</tbody>
</table>

Total events: 58
Total events: 46

Heterogeneity: Chi² = 8.58, df = 6 (P = 0.20); I² = 30%
Test for overall effect: Z = 1.73 (P = 0.08)

- effect estimate for each study, with CI
Forest plots

Headache at 24 hours

- scale and direction of benefit
Forest plots

Headache at 24 hours

- pooled effect estimate for all studies, with CI
### Forest plots

#### Headache at 24 hours

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Caffeinated coffee</th>
<th>Decaffeinated coffee</th>
<th>Risk Ratio IV, Fixed, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amore-Coffea 2000</td>
<td>2</td>
<td>10</td>
<td>0.22 [0.05, 0.92]</td>
</tr>
<tr>
<td>Deliciozza 2004</td>
<td>10</td>
<td>9</td>
<td>1.11 [0.51, 2.44]</td>
</tr>
<tr>
<td>Mama-Kaffa 1999</td>
<td>12</td>
<td>9</td>
<td>1.53 [0.70, 3.35]</td>
</tr>
<tr>
<td>Morrocona 1998</td>
<td>3</td>
<td>1</td>
<td>3.40 [0.39, 29.31]</td>
</tr>
<tr>
<td>Norscafe 1998</td>
<td>19</td>
<td>9</td>
<td>1.99 [0.97, 4.07]</td>
</tr>
<tr>
<td>Oollahlazza 1998</td>
<td>4</td>
<td>2</td>
<td>2.11 [0.41, 10.83]</td>
</tr>
<tr>
<td>Piazza-Allerta 2003</td>
<td>8</td>
<td>6</td>
<td>1.41 [0.54, 3.65]</td>
</tr>
</tbody>
</table>

**Total (95% CI)**

- **Events**: 277
- **Total**: 290
- **Weight**: 100.0%
- **Risk Ratio**: 1.38 [0.96, 2.00]

**Heterogeneity**

- Chi² = 8.58, df = 6 (P = 0.20); I² = 30%
- Test for overall effect: Z = 1.73 (P = 0.08)

---

- **Heterogeneity**
Interpreting confidence intervals

- always present estimate with a confidence interval

- precision
  - point estimate is the best guess of the effect
  - CI expresses uncertainty – range of values we can be reasonably sure includes the true effect

- significance
  - if the CI includes the null value
    - rarely means evidence of no effect
    - effect cannot be confirmed or refuted by the available evidence
  - consider what level of change is clinically important
Statistical and clinical significance
(from Berry G. (1986), Med. J. Aust, 144: 618-619)

Null hypothesis
0

Difference

Clinically important

(a) STATISTICALLY SIGNIFICANT
(b) Important

(c) STATISTICALLY NOT SIGNIFICANT
(d) Not important

Inconclusive True negative result
Presenting data in your review

• present outcomes in consistent order throughout
  • Abstract, Methods, Results, data

• forest plots
  • key forest plots linked as figures
    • usually primary outcomes
  • all forest plots will be published as supplementary data
  • avoid forest plots with only one study

• may also add other data tables
  • results of single studies
    • summary data for each group, effect estimates, confidence intervals
  • non-standard data
What to include in the protocol

- how will you decide whether a meta-analysis is appropriate?
- meta-analysis model to be used
Take home message

- there are several advantages to performing a meta-analysis but it is not always possible (or appropriate)
- plan your analysis carefully, including comparisons, outcomes and meta-analysis methods
- forest plots display the results of meta-analyses graphically
- interpret your results with caution
References


Acknowledgements

• Compiled by Miranda Cumpston
• Based on materials by Sally Hopewell, Julian Higgins, the Cochrane Statistical Methods Group and Cochrane Netherlands
• Approved by the Cochrane Methods Board