Heterogeneity

- What do we mean when we talk about heterogeneity?
- Invariably, studies brought together in a systematic review will differ, in some way, to each other.
- Within a systematic review, this is termed *heterogeneity*. 
Does it make conceptual sense to combine studies?

Shouldn’t combine apples and oranges (usually)—we should only combine studies that are similar enough to be comparable.

Such decisions require qualitative judgements.

3 hypothetical systematic reviews

1. A systematic review of all methods of assessment of fetal wellbeing for pregnant women, intended to generate a single estimate of the impact of these methods on perinatal mortality.

2. A systematic review that examines the effect of different antibiotics, such as tetracyclines, penicillins and chloramphenicol, on improvement in days of mechanical ventilation in infants admitted to neonatal intensive care.

3. A systematic review that examines the effects of epidural analgesia on the mother and the baby, when compared with non-epidural or no pain relief during labour.
2. A systematic review that examines the effect of different antibiotics, such as tetracyclines, penicillins and chloramphenicol, on improvement in days of mechanical ventilation in infants admitted to neonatal intensive care.

- What aspects of the primary studies must be similar if their results are to be combined in the above systematic review?

My thoughts...

<table>
<thead>
<tr>
<th>Population</th>
<th>Interventions</th>
<th>Outcomes</th>
<th>Study Methodologies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestation at birth;</td>
<td>Same antibiotic in all studies;</td>
<td>Type of ventilation;</td>
<td>All randomized trials</td>
</tr>
<tr>
<td>Reason for admission to intensive care;</td>
<td>Same class of antibiotic in all studies;</td>
<td>Oxygen support;</td>
<td>Only blinded randomized trials</td>
</tr>
<tr>
<td>Mode of birth;</td>
<td>Comparison of antibiotic with placebo;</td>
<td>Number of days to wean;</td>
<td>Cohort studies</td>
</tr>
<tr>
<td></td>
<td>Comparison of one antibiotic with another;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Keypoint

- Judgements as to whether the primary studies are similar enough to be combined in a systematic review is based on whether there is a plausible rational to expect a similar treatment effect (or result) across the range of infants, interventions, outcomes and methodologies of the primary studies.

Different types of heterogeneity

Clinical heterogeneity
- Variation in participants, interventions and outcomes across studies.

Methodological heterogeneity
- Variation in study design and risk of bias across studies.

Statistical heterogeneity
- Consequence of clinical and/or methodological heterogeneity;
- Variation in the treatment effects of the intervention being evaluated across the studies;
- Observed intervention effects being more different from each other than one would expect due to random error (chance) alone.
Assessing statistical heterogeneity

- If there is substantial heterogeneity among studies in a systematic review, it might be inappropriate to do a meta-analysis.

- How do we know if there is ‘substantial’ heterogeneity?
  1. Visual inspection of a forest plot of studies included in the review;
  2. Assessment of results of tests for statistical heterogeneity.

Results of the studies in 2 hypothetical systematic reviews.

For which systematic review does it make sense to combine results?

- Hatala R et al. CMAJ 2005; 172: 661-665
Visual criteria AGAINST pooling

- Reject combining studies into a single summary estimate if:
  1. highly disparate point estimates and
  2. confidence intervals with little overlap
If most studies are similar to each other and show a similar result (low heterogeneity), this increases confidence that the effect being measured is real.

If results from different studies are vastly different from each other, this suggests that each study is measuring something slightly different from the other studies.
Startegies for addressing heterogeneity

- There are a number of options if (statistical) heterogeneity is identified among a group of studies that would otherwise be considered suitable for a meta-analysis;

1. Do review without statistical meta-analysis
2. Use subgroup analyses
3. Use sensitivity analyses
4. Random effects meta-analysis can incorporate heterogeneity into meta-analysis (but doesn’t fix it)

Subgroup analyses

- Splitting all the participant data into subgroups, often so as to make comparisons between them.
- Subgroup analyses may be done for subsets of participants (such as primiparous and multiparous women), or for subsets of studies (such as different models of midwife-led care).
  - Known as effect modifiers, explanatory variables or covariates
- In addition to investigating heterogeneity, subgroup analyses can be used to answer specific questions about particular groups of participants, types of intervention or types of study.
Selection of study characteristics for subgroup analyses

- Specifying characteristics *apriori*;
- Keep number of characteristics as small as possible to minimise likelihood of a false positive result;
- Selection of characteristics should be justified and ideally supported by evidence from sources other than the included studies;

Subgroup analyses

- Be wary:
  - Not based on randomized comparisons
  - Therefore, observational by nature
  - Interpret cautiously
Subgroup Analyses

Table 1: Subgroup Analyses

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Restrict. square</th>
<th>Restrict. square</th>
<th>Risk Ratio</th>
<th>Weight</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Mutation</td>
<td>100.0%</td>
<td>10.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2nd Mutation</td>
<td>10.0%</td>
<td>100.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Subgroup Analyses

Table 2: Subgroup Analyses

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Restrict. square</th>
<th>Restrict. square</th>
<th>Risk Ratio</th>
<th>Weight</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Mutation</td>
<td>100.0%</td>
<td>10.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2nd Mutation</td>
<td>10.0%</td>
<td>100.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Subgroup Analyses

Table 3: Subgroup Analyses

<table>
<thead>
<tr>
<th>Study or subgroup</th>
<th>Restrict. square</th>
<th>Restrict. square</th>
<th>Risk Ratio</th>
<th>Weight</th>
<th>Risk Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Mutation</td>
<td>100.0%</td>
<td>10.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>2nd Mutation</td>
<td>10.0%</td>
<td>100.0%</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Subgroup Analyses

Sensitivity analyses

- A sensitivity analysis asks the question, “Are the findings robust to the decisions made in the process of obtaining them?”

- Examples of need for a sensitivity analysis:
  - Studies at high versus low risk of bias judgements
  - Different methods of imputation of data
  - Fixed or random effects models of meta analyses

- Should be determined a priori where possible but often it is not possible to prespecify;

- Reporting of effects of sensitivity analyses usually done narratively without forest plots
References