Addressing multiple treatments II: introduction to network meta-analysis

Madrid, October 2011

Workshop outline

- The Basics: indirect comparisons
  - What are indirect comparisons & why are they necessary
  - Exercise: how to do an indirect comparison (calculator)
- Slightly more advanced:
  - Checking assumptions for IC (and NMA) with exercise
  - Checking consistency
  - What does an NMA look like?
- Advantages and examples of NMA
  - Meta-regression approach
  - Methodological challenges

Multiple treatment decision-making

- For many clinical indications there will often be several possible interventions.
- The Cochrane Database of Systematic Reviews
  - 22 interventions for adult smoking cessation
  - >12 interventions for chronic asthma in adults
- Health care decisions should be based on 'best available' evidence from systematic reviews & meta-analysis of RCTs
Problem...

- Systematic reviews focus on direct, head-to-head comparisons of interventions.
  - e.g. NRT vs placebo; Olanzapine vs placebo
  - A vs B; A vs C.
- The evidence base consists of a set of pairwise comparisons of interventions
  - Placebo comparisons of limited use to the practitioner or policy-maker who wants to know the 'best' treatment to recommend/prescribe.

Problem... (2)

- 'Best available' evidence is not always available or sufficient
  - Placebo controlled trials sufficient for regulatory approval of new drugs
  - Even when active comparisons have been made such direct evidence is often limited.
- Therefore, evidence base may not contain treatment comparisons of relevance for clinician or policy maker.

Example evidence structure #1

- Common situation is to have multiple competing treatments (often within class) each studied in placebo-controlled RCTs but none compared directly to each other.

<table>
<thead>
<tr>
<th>Placebo</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
</table>

- How do we know which treatment to use?
Evidence base: 3 treatment options; 2 comparisons

Summary of results from 2 separate enuresis meta-analyses

Case study: childhood nocturnal enuresis *

<table>
<thead>
<tr>
<th>Comparison</th>
<th>n/ N active</th>
<th>n/ N no treat</th>
<th>Relative Risk</th>
<th>CIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm vs no treatment</td>
<td>107/ 316</td>
<td>250/ 260</td>
<td>0.39</td>
<td>(0.33 to 0.46)</td>
</tr>
<tr>
<td>Imipramine vs no treatment</td>
<td>314/ 400</td>
<td>391/ 403</td>
<td>0.95</td>
<td>(0.87 to 0.99)</td>
</tr>
</tbody>
</table>

Outcome: failure to achieve 14 days consecutive dry nights

Indirect comparisons

• In absence of direct evidence for treatments A vs B, an indirect estimate of log risk ratio $lrr_{AB}$ can be obtained from RCTs comparing A vs C and B vs C:

$$LRR_{BC} - LRR_{AC} = LRR_{AB}$$

*Bucher HC, et al. (1997); Glenny et al (2005)*
Indirect comparisons

- In absence of direct evidence for treatments A vs B, an indirect estimate of log risk ratio $\ln r_{AB}$ can be obtained from RCTs comparing A vs C and B vs C.

\[
\begin{align*}
A & \quad B & \quad C \\
I & & \\
\end{align*}
\]

Consistency equation*


---

3 treatment network

\[
\begin{align*}
\mu_{AC} & = \mu_{AB} - \mu_{BC} \\
\mu_{BC} & = \mu_{AC} - \mu_{AB} \\
\mu_{AC} & = \mu_{AB} - \mu_{BC}
\end{align*}
\]

Three possible indirect comparisons, all equivalent:
- $\mu_{AC} = \mu_{AB} - \mu_{BC}$
- $\mu_{BC} = \mu_{AC} - \mu_{AB}$
- $\mu_{AC} = \mu_{AB} - \mu_{BC}$

---

Simple exercise

Comparison | RR  | CIs        |
------------|-----|-----------|
No treatment vs Imipramine | 0.95 | (0.87 to 0.99) |
No treatment vs Alarm | 0.59 | (0.33 to 0.46) |

Outcome: failure to achieve 14 days consecutive dry nights
Simple exercise

<table>
<thead>
<tr>
<th>Comparison</th>
<th>RR</th>
<th>CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No treatment vs Imipramine</td>
<td>0.95</td>
<td>(0.87 to 0.99)</td>
</tr>
<tr>
<td>No treatment vs Alarm</td>
<td>0.39</td>
<td>(0.33 to 0.46)</td>
</tr>
</tbody>
</table>

A vs B is the effect of B relative to A: imipramine relative to placebo (or treated over control)

Pen and paper exercise.

\[
LRR_{BC} = LRR_{AC} - LRR_{AB}
\]

\[
lrr_{AB} = -0.06
\]

\[
lrr_{AC} = -0.93
\]

\[
lrr_{BC} = lrr_{AC} - lrr_{AB} = -0.93 - (-0.06) = -0.87
\]

Indirect RR_{BC} = \exp(lrr_{BC}) = 0.42
Confidence intervals and p-value

\[ \text{Var}(LRR_{AB}) = \text{Var}(LRR_{AC}) = 0.007 + 0.001 = 0.008 \]
\[ \text{SE}(LRR_{AB}) = \sqrt{0.008} = 0.09 \]

95% CI: LRR ± 1.96*SE = 0.35 to 0.50

\[ p = <0.0001 \quad (z = -9.66) \]

Note: Var(LRR_{AB}) = Var(LRR_{AC}) = Var(LRR_{BC})

Therefore, all things being equal (trials all of same size, equal variance and assuming a common treatment effect), a direct randomised trial is as precise as an indirect comparison based on 4 randomised trials (see Glenny, 2005 for more detail)

Online calculator:
http://www.cadth.ca/en/resources/ltc-user-guide

When is an indirect comparison sensible...

- Validity relies on the AB & AC RCTs being similar across factors which may affect the outcome (modify treatment effect).
- A clinical/epidemiological judgement:
  - No treatment by comparison interaction
  - Assuming inclusion/exclusion criteria same across comparisons
  - Patients, trial protocols, doses, administration etc. are similar in ways which might modify treatment effect.

"Between-trial comparisons [Indirect Comparisons] are unreliable. Patient populations may differ in their responsiveness to treatment. Therefore an apparently more effective treatment may have been tested in a more responsive population"

Cranney, Guyatt et al. End Rev 2002, 23; 570-8
Slide 19

“Placebo controlled trials lacking an active control give little useful information about comparative effectiveness. Such information cannot reliably be obtained from cross-study comparisons, as the conditions of the studies may have been quite different.”

International Council of Harmonisation E10 2.7.1.4

Slide 20

“Indirect comparisons are observational studies across trials, and may suffer the biases of observational studies, for example confounding.”

Cochrane Handbook for systematic reviews of interventions 4.2.5. Cochrane Library Issue 3

(Watch this space for CMIMG update...)

Slide 21

Checking assumptions

Exercise:
- Using the forest plots and study characteristics tables provided, work with a neighbour/ in small groups to discuss whether the AB and AC trials are similar enough across factors which may modify treatment effect.
- Suggested time: 10 minutes.
Handout: trial characteristics

Alarm vs placebo characteristics of studies

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Boys(%)</th>
<th>Exclusion</th>
<th>Previous treatment</th>
<th>Discrepancy</th>
<th>Baseline wetting (SD)</th>
<th>Recruitmentsetting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bollard (1981a)</td>
<td>81%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bollard (1981b)</td>
<td>63%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bollard (1981c)</td>
<td>7%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bennett (1985)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Houts (1986)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Jehu (1977)</td>
<td>7%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Lynch (1984)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Moffatt (1987)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Ronen (1992)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Nawaz (2002)</td>
<td>9%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Sacks (1974)</td>
<td>4%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Sloop (1973)</td>
<td>23%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Werry (1965)</td>
<td>43%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Wagner (1982)</td>
<td>82%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
</tbody>
</table>
| Test for overall effect: Z = 12.04 (P < 0.00001)
Heterogeneity: Chi² = 56.57, df = 13 (P < 0.00001); I² = 77%

Alarm versus no treatment

Handout: trial characteristics

Forest plot for AvB

Test for overall effect: Z = 12.04 (P < 0.00001)
Heterogeneity: Chi² = 56.57, df = 13 (P < 0.00001); I² = 77%

Alarm vs placebo characteristics of studies

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Boys(%)</th>
<th>Exclusion</th>
<th>Previous treatment</th>
<th>Discrepancy</th>
<th>Baseline wetting (SD)</th>
<th>Recruitmentsetting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bollard (1981a)</td>
<td>81%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bollard (1981b)</td>
<td>63%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bollard (1981c)</td>
<td>7%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Bennett (1985)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Houts (1986)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Jehu (1977)</td>
<td>7%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Lynch (1984)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Moffatt (1987)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Ronen (1992)</td>
<td>5%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Nawaz (2002)</td>
<td>9%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Sacks (1974)</td>
<td>4%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Sloop (1973)</td>
<td>23%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Werry (1965)</td>
<td>43%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
<tr>
<td>Wagner (1982)</td>
<td>82%</td>
<td>No details</td>
<td>No medication</td>
<td>No details</td>
<td>5.2 ± 1.8</td>
<td>Hospital clinic</td>
</tr>
</tbody>
</table>
| Test for overall effect: Z = 12.04 (P < 0.00001)
Heterogeneity: Chi² = 56.57, df = 13 (P < 0.00001); I² = 77%

Alarm versus no treatment

Handout: trial characteristics

Forest plot for AvB

Test for overall effect: Z = 12.04 (P < 0.00001)
Heterogeneity: Chi² = 56.57, df = 13 (P < 0.00001); I² = 77%
**Slide 24**

**Forest plot for AvC**

### Study or Subgroup
- Agarwala 1965
- Forsythe 1969
- Hodes 1973
- Khorana 1972
- Manhas 1967
- Poussaint 1965
- Schroder 1971
- Smellie 1976
- Tahmaz 2000
- Wagner 1982b

### Total (95% CI)

- Total events

### Heterogeneity
- Chi² = 269.99, df = 9 (P < 0.00001); I² = 97%

### Test for overall effect
- Z = 6.97 (P < 0.00001)

### Weight
- 10.1%
- 28.3%
- 10.6%
- 13.0%
- 9.2%
- 3.3%
- 10.2%
- 7.0%
- 4.7%
- 3.8%

### Risk Ratio
- M-H, Fixed, 95% CI
  - 0.93 [0.83, 1.05]
  - 0.99 [0.95, 1.02]
  - 0.96 [0.77, 1.18]
  - 0.55 [0.42, 0.73]
  - 0.36 [0.22, 0.59]
  - 0.44 [0.20, 0.96]
  - 1.04 [0.95, 1.15]
  - 0.21 [0.08, 0.53]
  - 0.64 [0.36, 1.13]
  - 0.73 [0.47, 1.12]
  - 0.77 [0.72, 0.83]

**Risk Ratio**
- 0.01
- 0.1
- 1
- 10
- 100

- Favours experimental
- Favours control

**Imipramine versus no treatment**

---

**Slide 25**

- Another common evidence structure is where we have some direct evidence on the relevant treatment comparisons (active vs active) but on its own its insufficient.

### Example evidence structure #2

- **Indirect evidence**
- **Direct evidence**

### No treatment
- Alarm
- Imipramine

### Evidence base: 3 treatment options; 3 comparisons

### Summary of results from 3 enuresis meta-analyses

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Relative Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alarm vs no treatment</td>
<td>0.39 (0.33 to 0.46)</td>
</tr>
<tr>
<td>Imipramine vs no treatment</td>
<td>0.95 (0.87 to 0.99)</td>
</tr>
<tr>
<td>Alarm vs imipramine</td>
<td>0.77 (0.64 to 0.93)</td>
</tr>
</tbody>
</table>
Slide 27

Network meta-analysis

Combines direct and indirect evidence. Also known as:
1) Mixed treatment comparison
2) Multiple treatment meta-analysis

ALL 3 mean the same thing – simultaneous comparison of multiple competing treatments using direct & indirect evidence (usually from RCTs) in a single analysis.

SAME assumption as made for indirect comparison alone: the consistency assumption.

Slide 28

Combining direct and indirect evidence

Simple approach to pooling direct and indirect evidence on \( \text{BC} \)

1. \( IRR_{\text{direct}}^0 \)
2. \( IRR_{\text{indirect}}^0 \)
3. \( IRR_{\text{NMA}}^0 = \left( \frac{w_1 IRR_{\text{direct}}^0 + w_2 IRR_{\text{indirect}}^0}{w_1 + w_2} \right) \)

\( w = \frac{1}{(w_1 w_2)^{\frac{1}{2}}} \)

Indirect evidence given less weight than direct evidence

Slide 29

Using GIV to combine in RevMan

<table>
<thead>
<tr>
<th>Study on Reference</th>
<th>Log(RR) Value</th>
<th>95% CI</th>
<th>SE</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct B vs C</td>
<td>-0.257 ± 0.095</td>
<td>0.64, 0.93</td>
<td>0.091</td>
<td>47.9%</td>
</tr>
<tr>
<td>Indirect B vs C</td>
<td>-0.87 ± 0.091</td>
<td>0.42, 0.93</td>
<td>0.091</td>
<td>52.1%</td>
</tr>
</tbody>
</table>

Heterogeneity: Not applicable

Test for overall effect: \( Z = 2.71 \) (P = 0.007)

Total (95% CI) \( 0.77 [0.64, 0.93] \)

Test for subgroup differences: \( \chi^2 = 21.71 \), df = 1 (P < 0.00001), I² = 95.4%

Test for overall effect: \( Z = 8.78 \) (P < 0.00001)
IC and NMA assume that the "Direct" and "Indirect" evidence estimate the same parameter, i.e. are CONSISTENT.

That the Treatment effect $\mu_{BC}$ estimated by the BC trials, would be the same as the treatment effect estimated by the AC and AB trials (if they had included B and C arms).

Nearly all the doubts about IC and NMA are doubts about this assumption.

**Slide 31**

**Discussion of indirect and direct estimates**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Risk Ratio</th>
<th>Risk Ratio</th>
<th>SE</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1 Direct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct B vs C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity:</td>
<td>Not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect:</td>
<td>Z = 2.71 (P = 0.007)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2.2 Indirect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indirect B vs C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity:</td>
<td>Not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect:</td>
<td>Z = 9.56 (P &lt; 0.00001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (95% CI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneity:</td>
<td>Not applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for overall effect:</td>
<td>Z = 8.78 (P &lt; 0.00001)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test for subgroup differences:</td>
<td>Chi² = 21.71, df = 1 (P &lt; 0.00001), I² = 95.4%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Slide 32**

**Bucher approach to checking consistency**

The difference $\omega$ between direct $\text{LRR}_{BC}$ and indirect $\text{LRR}_{BC}$

$\omega = -0.257 - (-0.87) = 0.61$

To calculate the standard error of the difference we sum the SE from the direct and indirect log risk ratios

$SE(\omega) = \sqrt{SE(\text{LRR}_{BC})^2 + SE(\text{LRR}_{BC})^2}$

$= \sqrt{0.091^2 + 0.095^2} = 0.13$
**Slide 33**

Bucher approach to checking consistency

Calculate confidence intervals & p-values for:

\[ 95\% \text{ CI} = \hat{\omega} \pm (1.96 \times \text{SE}) = \exp[0.36] \text{ to } \exp[0.86] \]

\[ z\text{-score} = \frac{\hat{\omega}}{\text{SE}} = 4.64 \quad p\text{-value} = <0.000002 \]

**Slide 34**

Limitations of simple approach

Straightforward & conceptually intuitive

- Extension of pairwise meta-analysis
- Checking consistency of evidence

BUT it is very LIMITED:

- Pool separately for each treatment comparison (separate meta-analyses)

What happens when

<table>
<thead>
<tr>
<th>Treatments</th>
<th>4 5 6 7 8 9 10 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pairwise</td>
<td>6 10 15 21 28 36 45 55</td>
</tr>
<tr>
<td>Indirect</td>
<td>12 30 60 105 168 252 360 495</td>
</tr>
</tbody>
</table>

**Slide 35**

Using Network Meta-analysis Methods to Compare Multiple Interventions

Part II

Tianjing Li, MD, MHS, PhD
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
19th Cochrane Colloquium
Madrid, Spain
October, 2011
Key Messages

- Network meta-analysis is an extension of standard, pair-wise meta-analysis; meta-regression, generalized linear model, and Bayesian approaches could be used.
- To ensure validity of findings from meta-analyses, the systematic review, whether it involves a standard, pair-wise meta-analysis or a network meta-analysis, must be designed rigorously and conducted carefully.

An Overview of Meta-regression

- In primary studies we use regression to examine the relationship between one or more covariates and a dependent variable.
- The same approach can be used with meta-analysis, except that:
  - Unit of analysis: each observation in the regression model is usually a study;
  - Dependent variable is the summary estimate in each primary study rather than outcomes measured in individual participants;
  - Covariates are at level of the study rather than the level of the participant.
**Why do a Meta-regression?**

- Examine the relationship between study-level characteristics and intervention effect.
- Study potential effect modification: Does the intervention effect (association) vary with different population or study characteristics?
- Explore and explain between study variation.

### Bacillus Calmette-Guérin (BCG) Vaccine to Prevent Tuberculosis Dataset

<table>
<thead>
<tr>
<th>ID</th>
<th>Study</th>
<th>Country</th>
<th>Year</th>
<th>Latitude</th>
<th>TB Vaccinated</th>
<th>TB Unvaccinated</th>
<th>RR</th>
<th>SE(ln RR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ferguson_1949</td>
<td>UK</td>
<td>1949</td>
<td>55</td>
<td>0.205</td>
<td>0.441</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Hart_1977</td>
<td>UK</td>
<td>1977</td>
<td>52</td>
<td>0.237</td>
<td>0.141</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Aronson_1948</td>
<td>UK</td>
<td>1948</td>
<td>44</td>
<td>0.411</td>
<td>0.571</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Stein_1953</td>
<td>USA</td>
<td>1953</td>
<td>44</td>
<td>0.456</td>
<td>0.083</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Rosenthal_1960</td>
<td>USA</td>
<td>1960</td>
<td>42</td>
<td>0.254</td>
<td>0.270</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Rosenthal_1961</td>
<td>USA</td>
<td>1961</td>
<td>42</td>
<td>0.260</td>
<td>0.644</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Comstock_1976</td>
<td>USA</td>
<td>1976</td>
<td>33</td>
<td>0.983</td>
<td>0.267</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Comstock_1969</td>
<td>USA</td>
<td>1969</td>
<td>33</td>
<td>1.562</td>
<td>0.730</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Coetzee_1968</td>
<td>South Africa</td>
<td>1968</td>
<td>27</td>
<td>0.625</td>
<td>0.238</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Vandiviere_1973</td>
<td>USA</td>
<td>1973</td>
<td>19</td>
<td>0.198</td>
<td>0.472</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Frimodt_1973</td>
<td>Denmark</td>
<td>1973</td>
<td>13</td>
<td>0.804</td>
<td>0.226</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>TB Prevention_1980</td>
<td>USA</td>
<td>1980</td>
<td>13</td>
<td>1.012</td>
<td>0.063</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

1. RR < 1.0 indicates the vaccine decreased the risk of TB.
2. The higher the latitude the farther away the study location was from the equator (used as surrogate for climates).

---

**Meta-regression Model Specification**

\[
\ln(RR) = a + b \times \text{latitude} + \epsilon_i
\]

- \(a\) – intercept, \(\ln(RR)\) at latitude=0 (equator)
- \(b\) – slope, the average change in \(\ln(RR)\) for every unit change in latitude
- \(\tau^2\) – between study variance

Parameters to estimate:
- \(a\) – intercept, \(\text{ln}(RR)\) at latitude=0 (equator)
- \(b\) – slope, the average change in \(\ln(RR)\) for every unit change in latitude
- \(\tau^2\) – between study variance
Variance (Heterogeneity) Explained by a Covariate

The spread of this distribution reflects the amount of between-study variance (τ²) without any covariate.

The spread of this distribution reflects the amount of between-study variance with a covariate assumed to be the same at each level of covariate.

The decrease in spread from the top to the bottom pane illustrates how a covariate explains some of the between-studies variance.


Network Meta-analysis using Meta-regression and Other Approaches

What is a Network Meta-analysis?

Network (multiple treatments comparison) meta-analysis:
Meta-analysis, in the context of a systematic review, in which three or more treatments have been compared using both direct and indirect evidence from several studies.

We observe $y_i$ in each study (e.g. the log(OR)).

Network meta-analysis and indirect comparison could be conducted under the meta-regression framework where treatments are treated as “covariates” in the model.

Slides 11-16 were adapted from workshop given previously by Georgia Salanati.

**Meta-regression Parameterization**

AC, AB, BC studies, chose C as reference, then:

$$y_i = [\beta_{AC} (\text{treat}=A) + \beta_{BC} (\text{treat}=B) + \beta_i]$$

Coding for indicator variables (treat=A, treat=B):

- AC studies: (1, 0)
- BC studies: (0, 1)
- AB studies: (-1, 1)

**Parameterization of the Network**

Choose basic parameters. Write all other contrasts as linear functions of the basic parameters to build the design matrix.
**Slide 48**

**Log(OR) for Death in Treatments for MI**

\[ \ln(OR) \text{ for Death in Treatments for MI} \]

<table>
<thead>
<tr>
<th>Treatment</th>
<th>No. studies</th>
<th>Streptokinase t-PA</th>
<th>Anistreplase</th>
<th>Accelerated t-PA</th>
<th>Angioplasty</th>
<th>Reteplase</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Lumley 2002, Stat Med

**Log(OR) compared to Streptokinase (RE Model)**

\[ Y = \left( \mu^A, \mu^B, \mu^C, \mu^D, \mu^E \right) \times X + \Delta \]

Y ~ N(μX, V)

\[ \Delta \sim N(0, \text{diag}(t^2)) \]

Variance-covariance matrix (for the observed LOR)

**Slide 49**

**Log(OR) for Death in Treatments for MI**

\[ \ln(OR) \text{ compared to Streptokinase (RE Model)} \]

<table>
<thead>
<tr>
<th>Treatment</th>
<th>LOR(SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PA</td>
<td>-0.02 (0.03)</td>
</tr>
<tr>
<td>Anistreplase</td>
<td>-0.00 (0.03)</td>
</tr>
<tr>
<td>Accelerated t-PA</td>
<td>-0.15 (0.05)</td>
</tr>
<tr>
<td>Angioplasty</td>
<td>-0.43 (0.20)</td>
</tr>
<tr>
<td>Reteplase</td>
<td>-0.11 (0.06)</td>
</tr>
</tbody>
</table>

**Slide 50**

**Log(OR) compared to Streptokinase (RE Model)**

\[ Y = \left( \mu^A, \mu^B, \mu^C, \mu^D, \mu^E \right) \times X + \Delta \]

<table>
<thead>
<tr>
<th>Treatment</th>
<th>LOR(SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-PA</td>
<td>-0.02 (0.03)</td>
</tr>
<tr>
<td>Anistreplase</td>
<td>-0.00 (0.03)</td>
</tr>
<tr>
<td>Accelerated t-PA</td>
<td>-0.15 (0.05)</td>
</tr>
<tr>
<td>Angioplasty</td>
<td>-0.43 (0.20)</td>
</tr>
<tr>
<td>Reteplase</td>
<td>-0.11 (0.06)</td>
</tr>
</tbody>
</table>
Example: Inhaled Drugs to Reduce Exacerbations in Patients with COPD


“We performed a logistic regression arm-level analysis with the presence of exacerbation as dependent and the different treatment options as independent variables. To preserve randomization within each trial, we included a dummy variable for each of the studies.”

Generalized Linear Model for Network Meta-analysis

Methodologic Challenges and Research Opportunities for Network Meta-analysis

Challenge of Considering Risk of Bias and Quality of Evidence

With particular thanks to Dr. Milo Puhan for the next 3 slides – drawing on his ideas
Overall (I-squared = 0.0%, p = 0.833)

Conventional meta-analysis: Entire evidence for 1 estimate

Network meta-analysis: Trials contribute to different estimates

Quality of evidence likely to be heterogeneous across network

Within and across comparisons
Challenge of Reporting Bias

Evidence Network of Comparative Efficacy and Acceptability of 12 New Generation Antidepressants

Cipriani et al. Lancet 2009; 373:746-58

117 RCTs
25,928 participants

Efficacy and Acceptability of 12 New-generation Antidepressants
Cipriani et al. Lancet 2009; 373:746-58

Best?
Worst?
ORs < 1 favor the row-defining treatment
Slide 60

Ranking of Efficacy and Acceptability of 12 New-generation Antidepressants

Cipriani et al. Lancet 2009; 373:746-58

Pr(mirtazapine) is the best treatment is high

Best
Worst

mirtazapine being ranked at each of 12 possible positions

Slide 61

Ranking of Efficacy and Acceptability of 12 New-generation Antidepressants

Cipriani et al. Lancet 2009; 373:746-58

Slide 62

Potential Bias in Study and Data Selection - Publication Bias

“Among placebo-controlled antidepressant trials registered with the FDA, most negative results are unpublished or published as positive.”

- 5 sertraline trials registered with FDA
  - 1 positive trial was published
  - 1 negative trial was published as positive
  - 3 were never published

Correspondence: Ioannidis JP. Lancet 2009; 373:1759-1760
Slide 63

Discrepant Rankings of Effect Sizes for Effectiveness of Antidepressants

Correspondence: Ioannidis JP. Lancet 2009; 373:1759-1760

Potential Bias in Study and Data Selection
- Publication Bias (cont'd)

Slide 64

Define the review question and eligibility criteria
Search for and select studies
Assess risk of bias, collect data
Synthesize evidence qualitatively
Synthesize evidence quantitatively
Interpret results and draw conclusions
Report findings

Methodologic considerations in doing a conventional systematic review
Challenges and areas of research for indirect comparison and network meta-analysis

- Define "network"
- Inclusion of observational studies for harms?
- Rely on studies included in published systematic reviews vs. a new comprehensive literature search?
- Different sources of data?
- Quality of indirect and combined evidence?
- Efficiency
- Workforce
- Extremely important but often overlooked
- Heterogeneity, inconsistency
- Subgroup analysis, meta-regression, sensitivity analysis
- Rare events, missing data
- More/less bias? Adjustment of bias
- Implementation and user-friendly software
- Interpretability and recommendations
- Reporting standards, peer review


Slide 65

Key Messages
- Network meta-analysis is an extension of standard, pair-wise meta-analysis; meta-regression, generalized linear model, and Bayesian approaches could be used.
- To ensure validity of findings from meta-analyses, the systematic review, whether it involves a standard, pair-wise meta-analysis or a network meta-analysis, must be designed rigorously and conducted carefully.