Comparing multiple treatments 2: statistical methods for network meta-analysis

Dimitris Mavridis

University of Ioannina Medical School
Department of Hygiene and Epidemiology
Greece

Argie Veroniki

University of Ioannina Medical School
Department of Hygiene and Epidemiology
Greece

Acknowledgements: Georgia Salanti, Julian Higgins, Andrea Cipriani, Jo McKenzie & many clinicians
We have no actual or potential conflict of interest in relation to this presentation
12 new generation antidepressants

- paroxetine
- sertraline
- milnacipran
- reboxetine
- paroxetine
- mirtazapine
- duloxetine
- fluvoxamine
- escitalopram
- citalopram
- bupropion
- venlafaxine
- fluoxetine
12 new generation antidepressants
A plethora meta-analyses has been published in the last years

“Although Mirtazapine is likely to have a faster onset of action than Sertraline and Paroxetine no significant differences were observed...”

“...statistically significant differences in terms of efficacy .... between Fluoxetine and Venlafaxine, but the clinical meaning of these differences is uncertain...”

“Venlafaxine tends to have a favorable trend in response rates compared with duloxetine”

“...meta-analysis highlighted a trend in favour of Sertraline over other Fluoxetine”

Fluoxetine: 28€    Venlafaxine: 111€    Sertaline: 76 €
12 new generation antidepressants

- paroxetine
- sertraline
- milnacipran
- reboxetine
- paroxetine
- duloxetine
- escitalopram
- mirtazapine
- duloxetine
- bupropion
- venlafaxine
- fluoxetine
- citalopram
- fluvoxamine
Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis

Andrea Cipriani, Toshiaki A Furukawa, Georgios Salanti, John R Geddes, Julian P T Higgins, Rachel Churchill, Norio Watanabe, Atsuo Nakagawa, Ichiro M Omori, Hugh McGuire, Michele Tansella, Corrado Barbui

Summary

Background Conventional meta-analyses have shown inconsistent results for efficacy of second-generation antidepressants. We therefore did a multiple-treatments meta-analysis, which accounts for both direct and indirect comparisons, to assess the effects of 12 new-generation antidepressants on major depression.

Methods We systematically reviewed 117 randomised controlled trials (25,928 participants) from 1991 up to Nov 30, 2007, which compared any of the following antidepressants at therapeutic dose range for the acute treatment of unipolar major depression in adults: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine. The main outcomes were the proportion

Lancet 2009; 373: 746-58
Published Online January 29, 2009
DOI:10.1016/S0140-6736(09)60046-5
See Comment page 700
Department of Medicine and Public Health, Section of Psychiatry and Clinical Neurosciences, University of Oxford, Oxford, UK
Department of Psychiatry, Kyoto University, Kyoto, Japan
Department of Psychiatry, St George’s University of London, London, UK
Department of Cochrane, St George’s University of London, London, UK
Department of Psychiatry, Tokyo Metropolitan Research Institute of General Medical Services, Tokyo, Japan
Department of Psychiatry, St George’s University of London, London, UK
Department of Psychiatry, University of Oxford, Oxford, UK
National Institute of Mental Health, Tokyo, Japan
Department of Cochrane, St George’s University of London, London, UK
12 new generation antidepressants

paroxetine —— reboxetine

duloxetine —— mirtazapine
escitalopram —— fluvoxamine
milnacipran —— citalopram
sertraline —— venlafaxine
bupropion —— fluoxetine
milnacipran —— paroxetine
sertraline ? duloxetine
bupropion —— escitalopram
fluvoxamine —— fluoxetine
12 new generation antidepressants

paroxetine  reboxetine

 duloxetine  mirtazapine

 escitalopram  fluvoxamine

 milnacipran  citalopram

 sertraline  venlafaxine

 bupropion  fluoxetine

 milnacipran  paroxetine

 sertraline  duloxetine

 bupropion  escitalopram

 fluvoxamine  fluoxetine
12 new generation antidepressants

- Paroxetine
- Duloxetine
- Escitalopram
- Milnacipran
- Sertraline
- Bupropion
- Mirtazapine
- Venlafaxine
- Escitalopram
- Fluoxetine
- Fluvoxamine
- Duloxetine
- Milnacipran
- Bupropion

Probability of being the best:

- Paroxetine: 0%
- Sertraline: 7%
- Citalopram: 0%
- Escitalopram: 26%
- Fluoxetine: 0%
- Fluvoxamine: 0%
- Milnacipran: 1%
- Venlafaxine: 11%
- Reboxetine: 0%
- Bupropion: 0%
- Mirtazapine: 54%
- Duloxetine: 0%
4 Fluoride modalities for preventing dental carries: series of pairwise meta-analyses

<table>
<thead>
<tr>
<th>Treatment comparison</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placebo Toothpaste</td>
<td>69</td>
</tr>
<tr>
<td>Placebo Gel</td>
<td>13</td>
</tr>
<tr>
<td>Placebo Rinse</td>
<td>31</td>
</tr>
<tr>
<td>Placebo Varnish</td>
<td>3</td>
</tr>
<tr>
<td>Toothpaste Rinse</td>
<td>6</td>
</tr>
<tr>
<td>Toothpaste Varnish</td>
<td>1</td>
</tr>
<tr>
<td>Gel Rinse</td>
<td>1</td>
</tr>
<tr>
<td>Gel Varnish</td>
<td>?</td>
</tr>
</tbody>
</table>
Multiple treatments and series of meta-analyses

With pairwise meta-analyses we cannot answer the following questions:

– Which fluoride modality is the best?
– What is the ranking of fluoride treatments according to effectiveness?
– Which is better: Gel or Varnish (0 studies)
A new methodological framework

Other names:
Multiple treatments meta-analysis,
Mixed treatment comparisons

Network meta-analysis

Meta-analysis of RCTs

Randomized Controlled trials (RCTs)

Cohort studies, Case-control studies
Aims of the workshop

• To explain *indirect* and *mixed* comparison of interventions
  – Assumptions
  – Statistical methods

• To understand the statistical models for *network meta-analysis*

• To discuss presentation of results from network meta-analysis

• To understand inconsistency models
Indirect and mixed effects

- Direct effect
- Indirect effect
- Mixed effect
Indirect comparison

• We can obtain an *indirect* estimate for A vs B from RCTs comparing A vs C and B vs C:

\[
SMD_{AB} = SMD_{AC} + SMD_{CB}
\]

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

\[
\text{Var}(SMD_{AB}) = \text{Var}(SMD_{AC}) + \text{Var}(SMD_{BC})
\]
Example

How to compare Gel to Toothpaste?

- **Placebo vs Toothpaste**
  - SMD: -0.34
  - CIs: (-0.41, -0.28)

- **Placebo vs Gel**
  - SMD: -0.19
  - CIs: (-0.30, -0.10)

**Comparison**

Placebo Toothpaste Gel
Exercise

Indirect $SMD_{GvsT} = SMD_{PvsT} - SMD_{PvsG}$

Indirect $SMD_{GvsT} = -0.34 - (-0.19) = -0.15$

Variance Indirect $SMD_{GvsT} = Variance SMD_{PvsT} + Variance SMD_{PvsG}$

$Variance SMD_{PvsT} = ((\text{high CI} - \text{low CI})/3.92)^2$

$Variance SMD_{PvsT} = ((-0.28 - (-0.41))/3.92)^2 = 0.0011$

$Variance SMD_{GvsT} = ((-0.10 - (-0.30))/3.92)^2 = 0.0026$

Variance Indirect $SMD_{GvsT} = 0.0011 + 0.0026 = 0.0037$

SE Indirect $SMD_{GvsT} = \sqrt{0.0037} = 0.061$

95% CI for Indirect $SMD_{GvsT} = (-0.15 - 1.96 \times 0.061, -0.15 + 1.96 \times 0.061)$

95% CI for Indirect $SMD_{GvsT} = (-0.27, -0.03)$
Indirect SMD Gel vs Toothpaste: -0.15 (0.0037)
Indirect SMD Gel vs Toothpaste: -0.15 (0.0037)

Mixed effect!
Mixed comparison

- Summarize direct and indirect effect size into a single mixed effect

\[
\begin{align*}
\text{Mixed SMD} &= \frac{\text{SMD}_{\text{Direct}}}{\text{var(} \text{SMD}_{\text{Direct}} \text{)}} + \frac{\text{SMD}_{\text{Indirect}}}{\text{var(} \text{SMD}_{\text{Indirect}} \text{)}} \\
\text{var(} \text{Mixed SMD} \text{)} &= \frac{1}{\text{var(} \text{SMD}_{\text{Direct}} \text{)}} + \frac{1}{\text{var(} \text{SMD}_{\text{Indirect}} \text{)}}
\end{align*}
\]
Mixed comparison

**Indirect** \( SMD_{GvsT} = -0.15 \)
\[ \text{Var}(\text{Indirect } SMD_{GvsT}) = 0.004 \]

**Direct** \( SMD_{GvsT} = 0.04 \)
\[ \text{Var}(\text{Direct } SMD_{GvsT}) = 0.011 \]

**Mixed** \( SMD_{GvsT} = -0.10 \)
\[ \text{Var}(\text{Direct } SMD_{GvsT}) = 0.003 \]

We gain precision!

You can do this with any measure... InOR, lnRR, RD, mean difference, HR, Peto’s lnOR etc…
Extend the idea of mixed effect sizes in the entire network
Transitivity

Untestable assumption

….but you can evaluate clinically and epidemiologically its plausibility

The anchor treatment A is ‘transitive’
Transitivity requires... (1)

The ‘anchor’ treatment A to be similarly defined when it appears in AB and AC trials. e.g. a treatment given at different doses but no systematic difference in the average dose of A across AB and AC studies.

The ‘anchor’ treatment A may be different in AB and AC studies e.g. injection versus pill.
Transitivity requires... (1)

– However, placebo toothpaste and placebo rinse might not be comparable as the mechanical function of brushing might have a different effect on the prevention of caries.

– If this is the case, the transitivity assumption is doubtful (Salanti 2009).
Transitivity means that...

AC and BC trials do not differ with respect to the distribution of effect modifiers.

Difficult to defend when you have older and newer treatments, and variables are often unobserved.
Distribution of mean dose of the active intervention in ten studies

Placebo vs A

Placebo vs B
Compare the distribution of important characteristics across treatments (Salanti et al 2009)
If all three A, B and C are transitive then the loop is consistent.
Consistency Equation

Placebo

Toothpaste

Gel

\[
\text{Placebo} \rightarrow \text{Toothpaste} \rightarrow \text{Gel}
\]

69, 3, 13
Inconsistency Factor

Indirect \( SMD_{GvST}^{ind} = -0.15 \)
\( \text{var}(SMD_{GvST}^{ind}) = 0.004 \)

Direct \( SMD_{GvST}^{dir} = 0.04 \)
\( \text{var}(SMD_{GvST}^{dir}) = 0.011 \)

\[
IF = \left| SMD_{GvST}^{dir} - SMD_{GvST}^{ind} \right| = \left| 0.04 - (-0.15) \right| = 0.19
\]

You can do this with any measure... lnOR, lnRR, RD, mean difference, HR etc.
Consistency and heterogeneity

a) Fixed effects analysis

b) Random effects analysis

Consistency and heterogeneity
Fit a network meta-analysis model

- Meta-analysis is a weighted regression with no covariates
- Network meta-analysis is a weighted regression with dummy variables for the treatments
- You should take into account correlations in multi-arm trials
Network and meta-regression

• Meta-regression using the treatments as ‘covariates’ and without intercept
• With 3 treatments and AC, AB, BC studies, chose C as reference, so AC and BC are *basic parameters*

\[
y_i = \mu_{AC} I_{iAC} + \mu_{BC} I_{iBC} + \delta_i + \varepsilon_i
\]

• The AC studies have \((I_{iAC}, I_{iBC}) = (1,0)\), the BC studies \((I_{iAC}, I_{iBC}) = (0,1)\) [*basic*]
• AB studies have \((I_{iAC}, I_{iBC}) = (1,-1)\) [*functional*] because AB=AC-BC
Toothpaste
Placebo
Varnish
Rinse
No treat
Gel

Basic Parameter
Functional Parameter

SMD_{TG} = SMD_{PG} - SMD_{PT}

Consistency equation
Choose a space of basic parameters

How many basic parameters?

Consistency equations

\[ SMD_{TG} = SMD_{PG} - SMD_{PT} \]
\[ SMD_{TR} = SMD_{PR} - SMD_{PT} \]
\[ SMD_{TV} = SMD_{PV} - SMD_{PT} \]
\[ SMD_{RG} = SMD_{PG} - SMD_{PR} \]

\( T - 1 \)
\[ y_i = \mu^{PT}T_i + \mu^{PG}G_i + \mu^{PR}R_i + \mu^{PV}V_i + \mu^{PN}N_i \]

<table>
<thead>
<tr>
<th>No. studies</th>
<th>Placebo</th>
<th>Toothpaste</th>
<th>Gel</th>
<th>Rinse</th>
<th>Varnish</th>
<th>NoTreatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>31</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
\[
y = X (\mu^{PT}, \mu^{PG}, \mu^{PR}, \mu^{PV}, \mu^{PN})' + \delta + \varepsilon
\]

- **\( \varepsilon \sim N(0, \text{diag}(v_i)) \)**
  - Variances matrix (for the observed SMD)

- **\( \delta \sim N(0, \text{diag}(\tau^2)) \)**
  - We assume equal heterogeneities for all comparisons

**Matrix of all observations**
- **Design matrix**
- **Vector of summary effects**
- **Random effects**
- **Random errors**
What’s the problem with multi-arm trials?

• We need to take into account the correlations between the estimates that come from the same study

• The random effects that refer to the same trial are correlated as well

• You have to build in the correlation matrix for the observed effects, and the correlation matrix for the random effects

\[ \varepsilon \sim N(0, S) \quad \delta \sim N(0, \Delta) \]
<table>
<thead>
<tr>
<th>Study</th>
<th>No. arms</th>
<th>#</th>
<th>Data</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>i=1</td>
<td>$T_1=2$</td>
<td>1</td>
<td>$y_{1,1}, v_{1,1}$</td>
<td>AB</td>
</tr>
<tr>
<td>i=2</td>
<td>$T_2=2$</td>
<td>1</td>
<td>$y_{2,1}, v_{2,1}$</td>
<td>AC</td>
</tr>
<tr>
<td>i=3</td>
<td>$T_3=2$</td>
<td>1</td>
<td>$y_{3,1}, v_{3,1}$</td>
<td>BC</td>
</tr>
<tr>
<td>i=4</td>
<td>$T_4=3$</td>
<td>2</td>
<td>$y_{4,1}, v_{4,1}$, $y_{4,2}, v_{4,2}$, $\text{cov}(y_{4,1}, y_{4,2})$</td>
<td>AB, AC</td>
</tr>
</tbody>
</table>

**Meta-regression**

\[
\begin{bmatrix}
y_{1,1} \\
y_{2,1} \\
y_{3,1} \\
y_{4,1} \\
y_{4,2}
\end{bmatrix} =
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
-1 & 1 \\
1 & 0 \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
\mu_{AB} \\
\mu_{AC}
\end{bmatrix} +
\begin{bmatrix}
\delta_{1,1} \\
\delta_{2,1} \\
\delta_{3,1} \\
\delta_{4,1} \\
\delta_{4,2}
\end{bmatrix} +
\begin{bmatrix}
\epsilon_{1,1} \\
\epsilon_{2,1} \\
\epsilon_{3,1} \\
\epsilon_{4,1} \\
\epsilon_{4,2}
\end{bmatrix}
\]
<table>
<thead>
<tr>
<th>Study</th>
<th>No. arms</th>
<th>#</th>
<th>Data</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>i=1</td>
<td>T₁=2</td>
<td>1</td>
<td>y₁,₁, v₁,₁</td>
<td>AB</td>
</tr>
<tr>
<td>i=2</td>
<td>T₂=2</td>
<td>1</td>
<td>y₂,₁, v₂,₁</td>
<td>AC</td>
</tr>
<tr>
<td>i=3</td>
<td>T₃=2</td>
<td>1</td>
<td>y₃,₁, v₃,₁</td>
<td>BC</td>
</tr>
<tr>
<td>i=4</td>
<td>T₄=3</td>
<td>2</td>
<td>y₄,₁, v₄,₁ \ y₄,₂, v₄,₂</td>
<td>AB AC</td>
</tr>
</tbody>
</table>

\[
\begin{pmatrix}
  y_{1,1} \\
y_{2,1} \\
y_{3,1} \\
y_{4,1} \\
y_{4,2}
\end{pmatrix} =
\begin{pmatrix}
  1 & 0 \\
  0 & 1 \\
-1 & 1 \\
 1 & 0 \\
 0 & 1
\end{pmatrix}
\begin{pmatrix}
  \mu_{AB} \\
  \mu_{AC}
\end{pmatrix} +
\begin{pmatrix}
  \varepsilon_{1,1} \\
  \varepsilon_{2,1} \\
  \varepsilon_{3,1} \\
  \varepsilon_{4,1} \\
  \varepsilon_{4,2}
\end{pmatrix} +
\begin{pmatrix}
  \delta_{1,1} \\
  \delta_{2,1} \\
  \delta_{3,1} \\
  \delta_{4,1} \\
  \delta_{4,2}
\end{pmatrix}
\begin{pmatrix}
  \varepsilon_{1,1} \\
  \varepsilon_{2,1} \\
  \varepsilon_{3,1} \\
  \varepsilon_{4,1} \\
  \varepsilon_{4,2}
\end{pmatrix}
\]

\[
\begin{pmatrix}
  \varepsilon_{1,1} \\
  \varepsilon_{2,1} \\
  \varepsilon_{3,1} \\
  \varepsilon_{4,1} \\
  \varepsilon_{4,2}
\end{pmatrix} \sim N
\begin{pmatrix}
  0 \\
  0 \\
  0 \\
  0 \\
  0
\end{pmatrix},
\begin{pmatrix}
  0 & 0 & 0 & 0 & 0 \\
  0 & 0 & v_{2,1} & 0 & 0 \\
  0 & 0 & 0 & v_{3,1} & 0 \\
  0 & 0 & 0 & 0 & v_{4,1} \\
  0 & 0 & 0 & \text{cov}(y_{4,1}, y_{4,2}) & v_{4,2}
\end{pmatrix}
\]

Take into account correlation in observations.
<table>
<thead>
<tr>
<th>Study</th>
<th>No. arms</th>
<th>#</th>
<th>Data</th>
<th>Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>i=1</td>
<td>$T_1=2$</td>
<td>1</td>
<td>$y_{1,1}, v_{1,1}$</td>
<td>AB</td>
</tr>
<tr>
<td>i=2</td>
<td>$T_2=2$</td>
<td>1</td>
<td>$y_{2,1}, v_{2,1}$</td>
<td>AC</td>
</tr>
<tr>
<td>i=3</td>
<td>$T_3=2$</td>
<td>1</td>
<td>$y_{3,1}, v_{3,1}$</td>
<td>BC</td>
</tr>
<tr>
<td>i=4</td>
<td>$T_4=3$</td>
<td>2</td>
<td>$y_{4,1}, v_{4,1}$ $y_{4,2}, v_{4,2}$</td>
<td>AB AC</td>
</tr>
</tbody>
</table>

$$
\begin{pmatrix}
y_{1,1} \\
y_{2,1} \\
y_{3,1} \\
y_{4,1} \\
y_{4,2}
\end{pmatrix} =
\begin{pmatrix}
1 & 0 \\
0 & 1 \\
-1 & 1 \\
1 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
\delta_{1,1} \\
\delta_{2,1} \\
\delta_{3,1} \\
\delta_{4,1} \\
\delta_{4,2}
\end{pmatrix} +
\begin{pmatrix}
\varepsilon_{1,1} \\
\varepsilon_{2,1} \\
\varepsilon_{3,1} \\
\varepsilon_{4,1} \\
\varepsilon_{4,2}
\end{pmatrix}
$$

\[
\begin{pmatrix}
\delta_{1,1} \\
\delta_{2,1} \\
\delta_{3,1} \\
\delta_{4,1} \\
\delta_{4,2}
\end{pmatrix} \sim N
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix},
\begin{pmatrix}
\tau_{AB}^2 & 0 & 0 & 0 \\
0 & \tau_{AC}^2 & 0 & 0 \\
0 & 0 & \tau_{BC}^2 & 0 \\
0 & 0 & 0 & 0
\end{pmatrix}
\]

Take into account correlation in random effects

Multivariate meta-analysis

• Studies typically report many outcomes
  – Efficacy and acceptability in antidepressants

• **Multivariate meta-analysis** allows a joint synthesis of the multiple end points

• Different between-treatment contrasts are viewed as different outcomes

• White et al estimate NMA models by expressing them as multivariate random-effects meta-regressions (mvmeta in STATA)
Data: \( n \) studies with 2 outcomes

\[
\begin{align*}
\text{Efficacy/AB} & \quad \text{Acceptability/AC} \\
\text{Study 1:} & \quad y_{11}, y_{12}, \begin{pmatrix} s_{11}^2 & S_{112} \\ S_{112} & s_{12}^2 \end{pmatrix} \\
\text{Study i:} & \quad y_{i1}, y_{i2}, \begin{pmatrix} s_{i1}^2 & S_{i12} \\ S_{i12} & s_{i2}^2 \end{pmatrix} \quad S_{i12} = \rho_i s_{i1} s_{i2} \\
\text{Study n:} & \quad y_{n1}, y_{n2}, \begin{pmatrix} s_{n1}^2 & S_{n12} \\ S_{n12} & s_{n2}^2 \end{pmatrix}
\end{align*}
\]
Network meta-analysis and multivariate approaches

- We can look at network meta-analysis as either a multivariate meta-regression or a multivariate meta-analysis

- Multivariate meta-regression:
  - extends the meta-regression approach to allow for multi-arm trials
  - dummy 1, -1 and 0 codes for treatments (with a reference in mind)
  - assumes a common heterogeneity variance

- Multivariate meta-analysis:
  - no covariates required
  - Flexible modelling of the between-study variance matrix
  - requires a common reference arm for every study
    - a problem that is surmountable using data augmentation
How to fit network meta-analysis?

- R mvmeta, metasem, netmeta
- STATA using metareg (no multi-arm studies)
- STATA mvmeta
- To my knowledge only netmeta in R and mvmeta in STATA model properly the matrix $\Delta$
- Using MCMC (WinBUGS)
Presenting results from network meta-analysis

• With many treatments judgments based on pairwise effect sizes are difficult to make
Antidepressants

<table>
<thead>
<tr>
<th>Efficacy (response rate) (95% CI)</th>
<th>Comparison</th>
<th>Acceptability (dropout rate) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BUP</strong></td>
<td><strong>CIT</strong></td>
<td><strong>DUL</strong></td>
</tr>
<tr>
<td>0.98 (0.78-1.28)</td>
<td>0.75 (0.55-1.02)</td>
<td>1.07 (0.86-1.31)</td>
</tr>
<tr>
<td>1.09 (0.83-1.43)</td>
<td>0.84 (0.70-1.01)</td>
<td>1.43 (1.09-1.85)</td>
</tr>
<tr>
<td>0.82 (0.67-1.01)</td>
<td>0.75 (0.60-0.93)</td>
<td>0.75 (0.70-1.01)</td>
</tr>
<tr>
<td>1.08 (0.90-1.29)</td>
<td>1.10 (0.93-1.31)</td>
<td>1.32 (1.12-1.55)</td>
</tr>
<tr>
<td>1.10 (0.83-1.47)</td>
<td>1.13 (0.86-1.47)</td>
<td>1.35 (1.02-1.76)</td>
</tr>
<tr>
<td>1.07 (0.77-1.48)</td>
<td>1.09 (0.78-1.50)</td>
<td>0.97 (0.69-1.38)</td>
</tr>
<tr>
<td>0.79 (0.72-1.00)</td>
<td>0.80 (0.63-1.01)</td>
<td>0.96 (0.76-1.19)</td>
</tr>
<tr>
<td>1.06 (0.87-1.30)</td>
<td>1.08 (0.90-1.30)</td>
<td>1.30 (1.10-1.53)</td>
</tr>
<tr>
<td>1.60 (1.22-2.16)</td>
<td>1.63 (1.25-2.14)</td>
<td>1.46 (1.05-2.02)</td>
</tr>
<tr>
<td>0.87 (0.72-1.05)</td>
<td>0.88 (0.62-1.01)</td>
<td>1.06 (0.88-1.27)</td>
</tr>
<tr>
<td>0.85 (0.70-1.01)</td>
<td>0.86 (0.71-1.05)</td>
<td>0.77 (0.60-0.99)</td>
</tr>
</tbody>
</table>

OR > 1 means the treatment in top-left is better.
Probabilities

- Estimate for each treatment *the probability of being the best*
- Rankings are constructed by drawing the coefficients a large number of times from their approximate posterior density
- For each draw, the effect sizes are estimated and the largest effect size is noted
# 12 new generation antidepressants

<table>
<thead>
<tr>
<th>Drug 1</th>
<th>Drug 2</th>
<th>Probability of being the best</th>
</tr>
</thead>
<tbody>
<tr>
<td>paroxetine</td>
<td>reboxetine</td>
<td>0%</td>
</tr>
<tr>
<td>duloxetine</td>
<td>mirtazapine</td>
<td>7%</td>
</tr>
<tr>
<td>escitalopram</td>
<td>fluvoxamine</td>
<td>0%</td>
</tr>
<tr>
<td>milnacipran</td>
<td>citalopram</td>
<td>26%</td>
</tr>
<tr>
<td>sertraline</td>
<td>venlafaxine</td>
<td>0%</td>
</tr>
<tr>
<td>bupropion</td>
<td>fluoxetine</td>
<td>0%</td>
</tr>
<tr>
<td>milnacipran</td>
<td>paroxetine</td>
<td>1%</td>
</tr>
<tr>
<td>sertraline</td>
<td>duloxetine</td>
<td>11%</td>
</tr>
<tr>
<td>bupropion</td>
<td>escitalopram</td>
<td>0%</td>
</tr>
<tr>
<td>fluvoxamine</td>
<td>milnacipran</td>
<td>54%</td>
</tr>
</tbody>
</table>

*The probability of being the best does not convey the spread of the rank probabilities....*
<table>
<thead>
<tr>
<th>% probability</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j=1$</td>
<td>0.25</td>
<td>0.50</td>
<td>0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>% probability</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>---------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>( j=1 )</td>
<td>0.25</td>
<td>0.50</td>
<td>0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>( j=2 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>( j=3 )</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>( j=4 )</td>
<td>0.25</td>
<td>0.0</td>
<td>0.0</td>
<td>0.75</td>
</tr>
</tbody>
</table>

\( i \) the treatment
\( j \) the rank
Ranking for efficacy (solid line) and acceptability (dotted line). Ranking: probability of being the best treatment, of being the second best, the third best and so on, among the 12 comparisons.
Surface under the cumulative ranking curve

Use posterior probabilities for each treatment to be among the $n$-best options

$$\text{Cumulative ranking curve}_i = \frac{\sum_{j=1}^{T-1} p_{ji}}{T-1}$$

$T$ Total number of treatments

[J Clin Epidemiol. 2010 Salanti et al]
<table>
<thead>
<tr>
<th>% probability</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>j=1</td>
<td>0.25</td>
<td>0.50</td>
<td>0.25</td>
<td>0.00</td>
</tr>
<tr>
<td>j=2</td>
<td>0.50</td>
<td>0.75</td>
<td>0.75</td>
<td>0.00</td>
</tr>
<tr>
<td>j=3</td>
<td>0.75</td>
<td>1.00</td>
<td>1.00</td>
<td>0.25</td>
</tr>
<tr>
<td>j=4</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

$i$ the treatment

$j$ the rank

The areas under the cumulative curves for the four treatments of the example above are:

- A = 0.5
- B = 0.75
- C = 0.67
- D = 0.08
Warning: measures based on probabilities are attractive, but can be unstable and should be presented along with the effect sizes!
Validity of network meta-analysis

- The validity of a network meta-analysis depends on *transitivity* of effect size parameters:
  - For any pair A and B, typical (or mean) advantage of A over B = advantage of A over C − advantage of B over C

- In a simple *indirect* comparison, we **cannot** test this assumption empirically.
- In a *network meta-analysis*, we sometimes **can**.
- We call this looking at *inconsistency*. 

Evaluate the assumption of consistency
What is inconsistency?

- **Consistency** = The data fit together according to the laws of transitivity
  
i.e.
  - for each pair of interventions A and B, all sources of evidence about A vs B agree with each other
    - (this means direct evidence (if available) and different routes to indirect evidence)

- **Inconsistency** = Lack of consistency

- Only *closed loops* can tell us about (in)consistency
Example: a simple loop of treatments

<table>
<thead>
<tr>
<th>Direct Evidence</th>
<th>Summary Treatment effect</th>
<th>Variance</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SMD_{GvsT}$</td>
<td>0.04</td>
<td>0.011</td>
<td>(–0.43, –0.25)</td>
</tr>
</tbody>
</table>

![Diagram showing the loop of treatments with Placebo, Toothpaste, and Gel]

<table>
<thead>
<tr>
<th>Indirect Evidence</th>
<th>Summary Treatment effect</th>
<th>Variance</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SMD_{PvsT}$</td>
<td>-0.34</td>
<td>0.002</td>
<td>(–0.43, –0.25)</td>
</tr>
<tr>
<td>$SMD_{PvsG}$</td>
<td>-0.19</td>
<td>0.002</td>
<td>(–0.28, –0.10)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indirect comparison</th>
<th>Summary Treatment effect</th>
<th>Variance</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SMD_{GvsT_ind}$</td>
<td>-0.15</td>
<td>0.004</td>
<td>(–0.27, –0.03)</td>
</tr>
</tbody>
</table>
How much inconsistency?

• Taking into account the previous evidence,
• the difference between direct and indirect estimates is

\[ IF = \left| SMD_{GvST}^{dir} - SMD_{GvST}^{ind} \right| = |0.04 - (-0.15)| = 0.19 \]

• and we add the variances (since the sources of evidence are independent):

\[ \text{Var}(\text{difference between direct and indirect}) = \]
\[ var(IF) = var(SMD_{GvST}^{dir}) + var(SMD_{GvST}^{ind}) = 0.004 + 0.011 = 0.015 \]
How much inconsistency?

\[ Z = \frac{IF}{\sqrt{\text{var}(IF)}} \sim N(0,1) \]

95\% Confidence Interval for Inconsistency

\[ IF \pm 1.96\sqrt{\text{var}(IF)} \]
\[ 0.19 \pm 1.96\sqrt{0.015} \]
\[ 0.19 \pm 0.24 \]
\[ (-0.05, 0.43) \]
Example: fluoride treatments

- Placebo
- Varnish
- Toothpaste
- Rinse
- Gel

Numbers represent the counts or frequencies of treatments: 69, 31, 13, 3, 1, 6, 3, 4, 4, 9, 1
Evaluation of consistency within closed loops

Estimates with 95% confidence intervals

Closed loops

NGV  NGR  NRV  PTG  PTV  PTR  TGV  TGR  TRV  PGV  PGR  PRV  GRV  AGRV  PTGV  PTGR  PTRV  TGRV  PGRV  PTGRV

Drawback: dependence between loops

R routine in www.mtm.uoi.gr/howotodoanmtm.html

[Clin Epidemiol 2009, Salanti et al]
Are networks typically inconsistent?

**Triangular networks**
- Song et al. BMJ 2011 found 16/112 (14%) inconsistent triangles
  - The same authors evaluated the assumption of consistency in Cochrane Reviews separately (Xiong et al. JCE 2013) and found 16/94 (17%) triangles inconsistent

**Complex networks**
- Veroniki et al. (IJE 2013) published network meta-analyses with binary data that involve at least 4 treatments and at least one closed loop
  - so far 40 networks, 303 loops
  - Inconsistency was detected in between 2% and 10% of the tested loops, depending on the effect measure and heterogeneity estimation method
  - About one eighth of the networks was found to be inconsistent.
Approaches for exploring inconsistency

Evaluation of local inconsistency

- Loop-Specific: examine each closed loop separately

Evaluation of global inconsistency

- Use a network meta-analysis model that allows for inconsistency (Lu & Ades JASA 2005)
- Compare model fit between consistency and inconsistency models
- Apply a ‘design by treatment’ interaction model (White et al RSM 2012, Higgins et al RSM 2012)
Inconsistency models: introduction

\[ SMD_{B vs C} = SMD_{A vs C} - SMD_{A vs B} + W \]
Model for consistency

Modelled log odds ratios
(basic parameters $\mu_{AB}$ and $\mu_{AC}$);
$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB}+ \delta_i$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC}+ \delta_i$</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>$\mu_{AB}+ \delta_i$</td>
<td>$\mu_{AC}+ \delta_i$</td>
</tr>
</tbody>
</table>
Model for inconsistency
Lu and Ades model

Modelled log odds ratios
(basic parameters $\mu_{AB}$ and $\mu_{AC}$);
$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>BC</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i + w$</td>
<td></td>
</tr>
</tbody>
</table>
Model for consistency with a three-arm trial

Modelled log odds ratios
(basic parameters $\mu_{AB}$ and $\mu_{AC}$);
$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
</tbody>
</table>
Issues with the Lu and Ades model

- In the presence of multi-arm trials, the Lu and Ades inconsistency model is not uniquely defined

- Problems arise because multi-arm trials *must* be consistent, so a network with multi-arm trials will have a mixture of *consistent* and *inconsistent* loops
Lu and Ades model

Modelled log odds ratios
(basic parameters $\mu_{AB}$ and $\mu_{AC}$);
$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i + w$</td>
</tr>
</tbody>
</table>
A model that is completely general is one that allows for all types of inconsistency
– inconsistency within loops made up of different trials
– inconsistency between two-arm and three-arm trials
– and beyond...

Such a model has been termed a design-by-treatment interaction model
Forms of Inconsistency

Loop Inconsistency

If they statistically differ: \textit{Inconsistency!}
Forms of Inconsistency

**Design Inconsistency**

- **Design BC**
- **Design ABC**

If they statistically differ: **Inconsistency!**
## Design-by-treatment interaction model

Modelled log odds ratios

(basic parameters $\mu_{AB}$ and $\mu_{AC}$);

$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i + w_{AB}$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC} + \delta_i + w_{AC}$</td>
</tr>
<tr>
<td>BC</td>
<td></td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i + w_{BC}$</td>
</tr>
</tbody>
</table>

[Higgins et al/ RSM 2012], [White et al/ RSM 2012]
Lu and Ades model for inconsistency with a three-arm trial

Modelled log odds ratios
(basic parameters $\mu_{AB}$ and $\mu_{AC}$);
$\delta_i$ is the heterogeneity random effect

<table>
<thead>
<tr>
<th>Trial</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>AB</td>
<td>ref</td>
<td>$\mu_{AB} + \delta_i$</td>
<td></td>
</tr>
<tr>
<td>AC</td>
<td>ref</td>
<td></td>
<td>$\mu_{AC} + \delta_i$</td>
</tr>
<tr>
<td>BC</td>
<td>$\mu_{AB} + \delta_i$</td>
<td>$\mu_{AC} + \delta_i + w$</td>
<td></td>
</tr>
</tbody>
</table>
Modelling the $w$ parameters

- When we have several inconsistency ($w$) parameters, we could let them have a random-effects distribution across comparisons
  \[ w_j \sim N(0, \sigma^2) \]
- Comparing $\sigma^2$ with $\tau^2$ (heterogeneity) allows us to assess the magnitude of the inconsistency
- I prefer to use fixed effects for the $w$ parameters
  - can interpret them individually
  - and it’s easier to fit the model using Stata
Example: Survival with 11 chemotherapy regimens in colorectal cancer
\[ \mu_{IO} = \mu_{IF} - \mu_{OF} + w_1 \]
Results: colorectal cancer network

- $w_1 = -0.08$, $w_2 = -0.07$, $w_3 = -0.06$, $w_4 = -0.03$
  - No loop is remarkably inconsistent

- $\sigma^2 = 0.11$ (SD 0.04), $\tau^2 = 0.19$ (SD 0.18)

- $P(\sigma^2 > \tau^2) = 0.41$
  - No important changes in posterior HRs or fit
What if we find inconsistency?

• Try to explain inconsistency!

• Use network meta-regression

• Might consider
  – presenting results from the inconsistency model
  – presenting a variety of separate direct, indirect and mixed comparisons

• Be careful! Selective inclusion of evidence pieces might lead to bias
Comparison of assumptions (random effects models)

<table>
<thead>
<tr>
<th><strong>Meta-analysis</strong></th>
<th><strong>Network meta-analysis</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Similarity of participants, interventions and outcomes</td>
<td>Similarity of participants, outcomes; ‘random selection’ of interventions</td>
</tr>
<tr>
<td>Appropriate modelling of study data (within-study variances often assumed known, uncorrelated with effects)</td>
<td>Appropriate modelling of study data (within-study variances often assumed known, uncorrelated with effects)</td>
</tr>
<tr>
<td>Normal distribution for random effects</td>
<td>Normal distribution for random effects</td>
</tr>
<tr>
<td>Possibly covariates to explain heterogeneity</td>
<td>Possibly covariates to explain heterogeneity and/or inconsistency</td>
</tr>
<tr>
<td></td>
<td>Possible assumptions about different τ values for different comparisons</td>
</tr>
<tr>
<td></td>
<td>Possible extra parameters to allow for inconsistency across comparisons</td>
</tr>
</tbody>
</table>
Multiple-Treatments Meta-analysis (MTM)

Meta-analysis is the statistical technique used to synthesize evidence from experiments addressing the same research question. It is often used to combine data from clinical trials regarding the relative effectiveness of two interventions in order, for example, to infer about whether antihypertensives A and B are equally effective in lowering blood pressure.

The main drawback of the current state of the art is that meta-analysis focuses on comparing only two alternatives. However, clinicians and patients need to know the relative ranking of a set of alternative options and not only whether option A is better than B.

The statistical methodology applied to synthesize information over a network of comparisons involving all alternative treatment options for the same condition is called Multiple-Treatments Meta-Analysis.

This site provides

- an introduction to statistical and methodological issues related to MTM
- links to training material
- support to statisticians with the analysis of networks of interventions
Hands on

- [www.mtm.uoi.gr](http://www.mtm.uoi.gr)
- Go to ‘how to do an MTM’ tab
- Use R routine `mtmnetwork.plot` to plot a network
- Use the R routine `ifplot.fun` to plot inconsistency in all closed loops
- In WinBUGS: read the description of models (e.g. [www.mtm.uoi.gr/3.binarymodeldescription.pdf](http://www.mtm.uoi.gr/3.binarymodeldescription.pdf)) download the data and the WinBUGS code
- Use the R routine `sucraplot.fun` to get rankograms and SUCRA
- Go to ‘STATA routines for Network Meta-Analysis’ tab for an implementation of network meta-analysis
Research Synthesis Methods
The official journal of the Society for Research Synthesis Methodology
http://www.srsm.org/

A special issue for Network Meta-analysis published in 2012
A new methods group has been recently established to support reviews that aim to compare multiple interventions.
Thank you!
Questions?