Using network meta-analysis (NMA) for decision making

Sofia Dias

with thanks to: Nicky Welton, Tony Ades, Debbi Caldwell, Alex Sutton

ISCB, August 2013, Munich
Outline

• What is NMA
  • Aka Mixed Treatment Comparisons (MTC) or Multiple Treatment Meta-analysis (MTM)
  • Brief review of methods and assumptions

• The National Institute for Health and Care Excellence (NICE) framework (UK)
  • Technology Appraisals and Clinical Guidelines
  • Principles of evidence-based decision making

• Use of NMA in NICE appraisals and guidelines
  • Technical Support Documents in Evidence Synthesis
What is an Indirect Comparison?

Pair-wise MA

A ─── B

Indirect Comparisons

Letters represent treatments; Lines represent comparisons made in RCTs
NMA: can combine direct and indirect evidence

Mixed/Multiple Treatment Comparisons (MTC)
Multiple Treatment Meta-analysis (MTM)
Network MA (NMA, NWMA)
NMA: Direct vs indirect

• When SEVERAL treatments A, B, C, D, … are to be compared, evidence that is “direct” for some comparisons is “indirect” for others, and the distinction becomes meaningless.

• Health care decisions should be based on ‘best available’ evidence from a systematic review & meta-analysis of ALL relevant RCTs.

• NMA provides a coherent summary of ALL the evidence
Consistency

- Suppose a decision maker is considering which of 3 treatments A, B, C is best for a specific (perhaps non-homogeneous) group of patients.
- The following statement about the true treatment effects must be correct:
  \[ d_{BC} = d_{AC} - d_{AB} \]

\[
\begin{array}{c}
\text{A} \quad \underline{d_{AB}} \quad \text{B} \quad \underline{d_{BC}} \quad \text{C}
\end{array}
\]
Coherent data

• In a decision problem, consistency is a fact!
• For a given patient population, inconsistency in the TRUE effects is IMPOSSIBLE.
• But inconsistent DATA is not
• The consistency assumption follows from exchangeability/homogeneity assumption made in pairwise MA (Lu & Ades 2006, 2009, NICE TSD2).
• As any modelling assumption, it should be checked whenever possible
  • check whether $\hat{d}_{AC} = \hat{d}_{AB} + \hat{d}_{BC}$ is close to being true, given the statistical error and heterogeneity (NICE TSD4)
Direct and Indirect evidence

• The existence of “evidence loops” means that there is both DIRECT evidence and INDIRECT evidence on the same contrast

• More data
  • estimates more precise
  • more robust (less sensitive to any one source of data)

• Possible, and indeed necessary, to check “consistency” of the direct and indirect.
Example: Early thrombolysis for AMI*

<table>
<thead>
<tr>
<th></th>
<th>RCTs</th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>Sk+tPA</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

**6 treatments:** Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated t-PA (Acc t-PA), Tenecteplase (TNK), Reteplase (r-PA)

**14 trials; 7 comparisons** made; **15 possible** pairwise comparisons

* Boland et al, *Health Technology Assessment*, 2003
**Results: Thrombolysis Pairwise MA**

Odds ratios (95% CrI), Fixed effects analysis

<table>
<thead>
<tr>
<th></th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>t-PA+SK</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK</td>
<td>X</td>
<td>1.00 (0.94-1.06)</td>
<td>0.86 (0.78-0.94)</td>
<td>0.96 (0.87-1.05)</td>
<td>0.95 (0.79-1.12)</td>
<td></td>
</tr>
<tr>
<td>t-PA</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc t-PA</td>
<td>X</td>
<td>1.12 (1.00-1.25)</td>
<td>1.02 (0.90-1.16)</td>
<td>1.01 (0.88-1.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-PA+SK</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r-PA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TNK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
Upper right: pair-wise OR (95% CrI)
Lower left: NMA OR (95% CrI)

<table>
<thead>
<tr>
<th></th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>t-PA+SK</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK</td>
<td>X</td>
<td>1.00 &lt;br&gt;(0.94-1.06)</td>
<td>0.86 &lt;br&gt;(0.78-0.94)</td>
<td>0.96 &lt;br&gt;(0.87-1.05)</td>
<td>0.95 &lt;br&gt;(0.79-1.12)</td>
<td></td>
</tr>
<tr>
<td>t-PA</td>
<td>1.00 &lt;br&gt;(0.94-1.06)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc t-PA</td>
<td>0.87 &lt;br&gt;(0.78-0.94)</td>
<td>0.87 &lt;br&gt;(0.78-0.97)</td>
<td>X</td>
<td>1.12 &lt;br&gt;(1.00-1.25)</td>
<td>1.02 &lt;br&gt;(0.90-1.16)</td>
<td>1.01 &lt;br&gt;(0.88-1.14)</td>
</tr>
<tr>
<td>t-PA+SK</td>
<td>0.96 &lt;br&gt;(0.88-1.05)</td>
<td>0.97 &lt;br&gt;(0.86-1.08)</td>
<td>1.11 &lt;br&gt;(1.00-1.24)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r-PA</td>
<td>0.90 &lt;br&gt;(0.80-1.01)</td>
<td>0.91 &lt;br&gt;(0.79-1.03)</td>
<td>1.04 &lt;br&gt;(0.94-1.16)</td>
<td>0.94 &lt;br&gt;(0.82-1.14)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TNK</td>
<td>0.87 &lt;br&gt;(0.75-1.01)</td>
<td>0.88 &lt;br&gt;(0.74-1.03)</td>
<td>1.01 &lt;br&gt;(0.89-1.14)</td>
<td>0.91 &lt;br&gt;(0.77-1.07)</td>
<td>0.97 &lt;br&gt;(0.82-1.14)</td>
<td>X</td>
</tr>
</tbody>
</table>
Upper right: pair-wise OR (95% CrI)  
Lower left: NMA OR (95% CrI)

<table>
<thead>
<tr>
<th></th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>t-PA+SK</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-PA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc t-PA</td>
<td>0.87 (0.78-0.94)</td>
<td>0.87 (0.78-0.97)</td>
<td>X</td>
<td>1.12 (1.00-1.25)</td>
<td>1.02 (0.90-1.16)</td>
<td>1.01 (0.88-1.14)</td>
</tr>
<tr>
<td>t-PA+SK</td>
<td>0.96 (0.88-1.05)</td>
<td>0.97 (0.86-1.08)</td>
<td>1.11 (1.00-1.24)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r-PA</td>
<td>0.90 (0.80-1.01)</td>
<td>0.91 (0.79-1.03)</td>
<td>1.04 (0.94-1.16)</td>
<td>0.94 (0.82-1.14)</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TNK</td>
<td>0.87 (0.75-1.01)</td>
<td>0.88 (0.74-1.03)</td>
<td>1.01 (0.89-1.14)</td>
<td>0.91 (0.77-1.07)</td>
<td>0.97 (0.82-1.14)</td>
<td>X</td>
</tr>
</tbody>
</table>
### Example: Early thrombolysis for AMI*

#### Table:

<table>
<thead>
<tr>
<th>RCTs</th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>Sk+tPA</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

#### Diagram:

- **t-PA**
- **r-PA**
- **SK**
- **Acc t-PA**
- **SK+t-PA**
- **TNK**

#### Text:

**6 treatments:** Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated t-PA (Acc t-PA), Tenecteplase (TNK), Retepase (r-PA)

**14 trials; 7 comparisons** made; **15 possible** pairwise comparisons

* Boland et al, *Health Technology Assessment*, 2003
<table>
<thead>
<tr>
<th></th>
<th>SK</th>
<th>t-PA</th>
<th>Acc t-PA</th>
<th>t-PA+SK</th>
<th>r-PA</th>
<th>TNK</th>
</tr>
</thead>
<tbody>
<tr>
<td>SK</td>
<td>×</td>
<td>1.00 (0.94-1.06)</td>
<td>0.86 (0.78-0.94)</td>
<td>0.96 (0.87-1.05)</td>
<td>0.95 (0.79-1.12)</td>
<td></td>
</tr>
<tr>
<td>t-PA</td>
<td>1.00 (0.94-1.06)</td>
<td>×</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acc t-PA</td>
<td>0.87 (0.78-0.94)</td>
<td>0.87 (0.78-0.97)</td>
<td>×</td>
<td>1.12 (1.00-1.25)</td>
<td>1.02 (0.90-1.16)</td>
<td>1.01 (0.88-1.14)</td>
</tr>
<tr>
<td>t-PA+SK</td>
<td>0.96 (0.88-1.05)</td>
<td>0.97 (0.86-1.08)</td>
<td>1.11 (1.00-1.24)</td>
<td>×</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r-PA</td>
<td>0.90 (0.80-1.01)</td>
<td>0.91 (0.79-1.03)</td>
<td>1.04 (0.94-1.16)</td>
<td>0.94 (0.82-1.14)</td>
<td>×</td>
<td></td>
</tr>
<tr>
<td>TNK</td>
<td>0.87 (0.75-1.01)</td>
<td>0.88 (0.74-1.03)</td>
<td>1.01 (0.89-1.14)</td>
<td>0.91 (0.77-1.07)</td>
<td>0.97 (0.82-1.14)</td>
<td>×</td>
</tr>
</tbody>
</table>
**Example: Early thrombolysis for AMI**

6 treatments: Streptokinase (SK), Tissue-plasminogen activator (t-PA), Accelerated t-PA (Acc t-PA), Tenecteplase (TNK), Reteplase (r-PA)

14 trials; 7 comparisons made; 15 possible pairwise comparisons

* Boland et al, *Health Technology Assessment*, 2003
NICE principles of evidence-based decision making

• NICE must consider the “broad balance of clinical benefits and costs” and the need to allocate resources fairly.

• This must be done in a consistent way both across conditions and across appraisals.
  • regardless of number of treatments being assessed or type of outcome

• Must adhere to principles of accountability, fairness and transparency.
NICE framework for technology appraisals (TA)

- In the UK new technologies go through a two-stage system of approval
  1. Licensing, usually based on efficacy against Placebo for a given indication
  2. NICE TA, to determine if technology is cost-effective for the National Health Service (NHS)
     - Single TA or Multiple TA
- NICE guidance is mandatory
  - If technology approved, it must be adopted within 3 months (available to all patients with condition)
NICE framework for clinical guidelines (CG)

• NICE also produces CG which recommend best practice for a range of clinical conditions
  • Recommend from the range of available treatments
  • Unless already deemed not cost-effective in a TA
  • Usually only recommends treatments licensed for the indication, but can consider off-label use in some cases

• CG recommendations are not mandatory
  • But evidence that uptake is generally good
Cost-effectiveness

- NICE guidance is based on cost-effectiveness to ensure more health-related quality of life per £ spent on the NHS
  - tax funded, free at the point of use
- NICE recommends an incremental cost-utility approach
- Technologies should be compared to “the most appropriate comparator(s)"
  NICE, Guide to the methods of TA 2013
“Appropriate Comparator(s)"

- Usually an established treatment or treatments currently being used or recommended for use on the NHS
  - NOT usually a Placebo
- Although direct evidence from head-to-head RCTs is preferred (base case), this is often not available
  - Even when it is available, there may not be much of it.
- Indirect comparisons and NMA are then required to inform decisions
Treatments for rheumatoid arthritis

- Placebo + MTX (1)
- Tocilizumab + MTX (7)
- Infliximab + MTX (5)
- Etanercept + MTX (4)
- Adalimumab + MTX (3)
- Rituximab + MTX (6)
- CZP + MTX (2)

NICE TA186 Certolizumab Pegol for the treatment of Rheumatoid Arthritis, 2010
Discussion of NMA/MTC results

• “The Committee noted that there was a high degree of uncertainty around the effectiveness point estimates, demonstrated by the wide overlapping confidence intervals. (...)

• “The Committee recognised the heterogeneity of the studies, highlighted by the mixed-treatment comparison, as well as the potential methodological limitations and concluded that there was no convincing evidence that certolizumab pegol was more or less effective than other TNF inhibitors.”
Discussion of Cost-effectiveness

Results

• “(...) certolizumab pegol had the highest probability of being cost effective at £20,000 per QALY gained.”
NMA in Guidelines

- NMA used in NICE CG on a variety of topics
  - Colorectal cancer, Advanced Breast Cancer, Schizophrenia, Hyperphosphataemia, Social Anxiety, Urinary Incontinence, Neuropathic Pain etc
- NICE CG Technical Support Unit to advise and support use of complex evidence synthesis methods
  - Based at the University of Bristol
  - Advises on models for NMA (amongst other things...)

University of Bristol
Treatments for Schizophrenia

- Risperidone
  - Olanzapine (2)
  - Ziprasidone (6)
  - Placebo (1)
  - Paliperidone (7)
- Haloperidol (8)
  - Amisulpride (3)
  - Zotepine (4)
  - Aripiprazole (5)

Value in Health 2010 Ades et al
NMA for Schizophrenia

• Data available from 17 RCTs on
  • Relapse
  • Drug discontinuation due to intolerable side effects
  • Drug discontinuation due to other reasons

• Outcomes are not independent

• Analysed in a single NMA using ‘competing risks’ logistic regression model assuming a multinomial distribution of data
Brief History of NMA

- The original idea was published in 1996
  - Higgins & Whitehead, Stat Med
- It was later picked up by the Health Technology Assessment community due to need to compare multiple treatments
  - Lu & Ades, Stat Med. 2004
  - Initially used in the UK (NICE)
  - Recently also used in other countries (Canada, Germany, USA ...)
- Regular feature of NICE technology appraisals
  - Since Jan 2009 59/129 (46%) included indirect or mixed treatment comparisons (NMA)
Recent Developments

• **NICE principles of accountability, fairness, transparency** require a **uniform** approach to evidence synthesis
  • regardless of medical condition, data format or number of treatments being assessed.

• **NICE Decision Support Unit Technical Support Documents in Evidence Synthesis** propose
  • generalised linear modelling framework
  • uniform set of criteria for model fit, selection and diagnostics, applicable to any type of outcome (or outcomes)
  • Bayesian approach

• **Pairwise meta-analysis and indirect comparisons** are special cases of NMA (same model and code are used)
NICE Technical Support Documents in Evidence Synthesis

• Details of implementation of pairwise and NMA under the framework in which NICE operates in the UK
  • 7 TSDs produced, including a checklist for reviewers
  • Also a tutorial series in Medical Decision Making (Vol. 33, July 2013)
    • Model details, examples, discussions and WinBUGS code
• Applicable to most decision problems based on effectiveness or cost-effectiveness.
• Although not compulsory, provide a guide to synthesis approach
NICE DSU TECHNICAL SUPPORT DOCUMENT 2: A GENERALISED LINEAR MODELLING FRAMEWORK FOR PAIRWISE AND NETWORK META-ANALYSIS OF RANDOMISED CONTROLLED TRIALS

REPORT BY THE DECISION SUPPORT UNIT

Last updated August 2011

Sofia Dias¹, Nicky J Welton¹, Alex J Sutton², AE Ades¹
Advantages of Bayesian approach

- Code available for WinBUGS for a variety of different data formats (binary, continuous, ordered)
  - Correctly accounts for correlation in trials with >2 arms (NICE TSD2)
- Easily adaptable to more complex models and different data formats
  - some studies report Hazard ratios, others report counts at different time points etc
  - covariates\(^1\), bias adjustment\(^2\)
  - Individual Participant Data for some or all trials

\(^1\)NICE TSD3
\(^2\)Dias et al JRSSA 2010
Advantages of Bayesian approach (cont)

• Decision making and cost-effectiveness or risk/benefit analyses are inherently Bayesian
  • Output from NMA easy to use in decision models
  • Correctly propagating uncertainty and correlations in all parameters

• Can make probability statements, such as probability treatment A is best, or in top 5 etc
• Can produce treatment rankings directly
Symptoms of SA
Social Anxiety: Data

• 41 different interventions in 17 classes
  • Psychological interventions, drugs, combinations
  • 3 different “controls”: Wait list, Pill Placebo, Attention-matched control
  • 820 possible pairwise comparisons

• 100 RCTs

• Sparse network
  • Most comparisons only made in one or two trials

• Data reported on various continuous and dichotomous scales
  • Clinicians want results on SMD
  • But Probability of Recovery required for economic model (obtained from OR)
Social Anxiety: Analysis

- Single model incorporating all data, on all treatments and outcomes – easy to do using BUGS language
- Class effects model
  - treatment effects around a class mean
  - random effects within class, in addition to between study variance
- Produced coherent outputs on both SMD and OR scales
  - Treatment ranks, and hence decisions, were compatible
- Results more precise and more robust
- Allowed a better estimate of between-trial variability (heterogeneity) using all evidence
  - To better account for uncertainty
Class effects (SMD)

-2 -1.5 -1 -.5 0 .5
SMD

REFERENCE
Practical Advantages of NMA

• Allows Technology Appraisal Committees or Guideline Development Groups to assess direction of effects and uncertainty on **ALL** comparisons

• Naturally incorporated into Decision Models

• If separate pairwise MA conducted, can be faced with several tables and forest plots providing relative estimates which are not consistent/coherent
  • And may not even include the comparisons of interest

• Cannot make coherent decisions based on inconsistent evidence!!
“From the NMA results we can now see that the drug we intended to recommend is not the best one after all, although it looked pretty good against [standard comparator].

“It’s like in football, if Liverpool beat Sunderland, they still might not beat Manchester United [current champions]...”
THANK YOU!
References

Our website:  http://www.bristol.ac.uk/social-community-medicine/projects/mpes/
NICE DSU website: http://www.nicedsu.org.uk/

Evidence Synthesis for Decision Making 1-7, Medical Decision Making  Vol 33, 2013. http://mdm.sagepub.com/content/33/5.toc


