NMA Learning Live Webinar series

Question formulation and protocol development for systematic reviews with network meta-analysis

Anna Chaimani

Research Center of Epidemiology and Statistics Sorbonne Paris Cité (CRESS-UMR1153), Université de Paris, Inserm, France

18 November 2019

Acknowledgements: Georgia Salanti, Julian Higgins, Deborah Caldwell, Tianjing Li
Meta-analysis in the literature

- Systematic reviews and meta-analyses of randomized controlled trials have “transformed medicine”
  - Establish evidence-based practice
  - Resolve contradictory research outcomes
  - Support research planning and prioritization

- Massive production of meta-analyses assessing healthcare interventions
  - More than 10,000 meta-analyses of RCTs per year

Donnelly et al., Nature 2018
Sutherland et al. Nature 2018
Limitation of pairwise meta-analysis

Example: Antidepressants for major depression

Paroxetine versus other anti-depressive agents for depression

Marianna Purgato¹, Davide Papola¹, Chiara Gastaldon¹, Carlotta Trespidi¹, Laura R Magni², Carla Rizzo³, Toshi A Furukawa⁴, Norio Watanabe⁵, Andrea Cipriani⁶, Corrado Barbi¹

“Paroxetine was more effective than reboxetine…”

“…less effective than mirtazapine”

“…less effective than citalopram”

Purgato et al. Cochrane Database Syst Rev 2014
Limitation of pairwise meta-analysis

Example: Antidepressants for major depression

Paroxetine versus other anti-depressive agents for depression

Marianna Purgato¹, Davide Papola¹, Chiara Gastaldon¹, Carlotta Trespili¹, Laura R Magni², Carla Rizzo³, Toshi A Furukawa⁴, Norio Watanabe⁵, Andrea Cipriani⁶, Corrado Barbui¹

Duloxetine versus other anti-depressive agents for depression

Andrea Cipriani¹, Markus Koeser³, Toshi A Furukawa³, Michela Nosè⁴, Marianna Purgato¹, Ichiro M Omori⁵, Carlotta Trespili¹, Corrado Barbui¹

“…no statistically significant differences in efficacy when compared with other antidepressants…”

“…when compared with escitalopram or venlafaxine, there was a higher drop-out rate…”

“…more adverse events than paroxetine…”

Cipriani et al. Cochrane Database Syst Rev 2012
From pairwise to network meta-analysis

Example: Antidepressants for major depression

The most critical question raised by patients and clinicians at the point of care is

“what is the drug of choice for the given condition?”

Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis


Cipriani et al. Lancet 2018
Network meta-analysis in medical research

WHO (World Health Organization) guidelines now rely whenever possible on network meta-analysis

- All interventions
- Two interventions

Network meta-analysis of RCTs

- Meta-analysis of RCTs
- Randomized Controlled trials (RCTs)
- Cohort studies, Case-control studies

The highest possible level of clinical evidence

Faltinsen et al. BMJ Evid Based Med 2018
Indirect and mixed effects
Indirect and mixed effects

Indirect effect
Direct effect
Mixed effect
Indirect and mixed effects
Indirect and mixed effects
Indirect and mixed effects
Indirect and mixed effects

![Diagram with various medications and their connections]

- Paroxetine
- Fluoxetine
- Indirect and mixed effects
- Placebo
- Agomelatine
- Bupropion
- Citalopram
- Desvenlafaxine
- Duloxetine
- Escitalopram
- Fluoxetine
- Fluvoxamine
- Levomilnacipran
- Milnacipran
- Mirtazapine
- Nefazodone
- Paroxetine
- Sertraline
- Reboxetine
- Venlafaxine
- Vortioxetine
- Trazodone
- Vilazodone
- Venlafaxine
- Amitriptyline
- Clomipramine
- Bupropion
- Agomelatine
Indirect and mixed effects

Drugs:
- Paroxetine
- Fluoxetine
- Indirect and mixed effects
- Placebo
- Agomelatine
- Bupropion
- Citalopram
- Desvenlafaxine
- Milnacipran
- Nefazodone
- Sertraline
- Trazodone
- Vortioxetine
- Levomilnacipran
- Duloxetine
- Escitalopram
- Fluvoxamine
- Milnacipran
- Mirtazapine
- Venlafaxine
- Amitriptyline
- Venlafaxine
- vilazodone
- Reboxetine

13
NMA projects are long, cumbersome and challenging
Writing the protocol is a valuable opportunity to get things right from the start and get to know your collaborators

It involves
- long discussions (and disagreement!) between clinical experts
- even longer discussions between statisticians and clinicians

It ensures that
- all needed data will be extracted and formatted in a convenient way
- all team members learn to ‘speak the same language’

Updating the evidence
- much much easier and quicker
Systematic review protocols with multiple interventions

A RevMan template for NMA protocols is under preparation

Additional considerations are required when preparing a protocol for a systematic review with multiple interventions

Anna Chaimani\textsuperscript{a,d}, Deborah M. Caldwell\textsuperscript{b}, Tianjing Li\textsuperscript{c}, Julian P.T. Higgins\textsuperscript{b}, Georgia Salanti\textsuperscript{a,d,e}

\textsuperscript{a}Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, University Campus, Ioannina 45110, Greece
\textsuperscript{b}School of Social and Community Medicine, University of Bristol, Canynge Hall, 39 Whitley Road, Bristol BS8 2PS, UK
\textsuperscript{c}Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, E6011, Baltimore, MD 21205, USA
\textsuperscript{d}Institute of Social and Preventive Medicine (ISPM), University of Bern, Finkenhügelweg 11, Bern 3012, Switzerland
\textsuperscript{e}Institute of Primary Health Care (IPCAM), University of Bern, Gesellschaftsstrasse 49, Bern CH 3012, Switzerland

Accepted 30 November 2016; Published online 11 January 2017

Abstract

Objectives: The number of systematic reviews that aim to compare multiple interventions using network meta-analysis is increasing. In this study, we highlight aspects of a standard systematic review protocol that may need modification when multiple interventions are to be compared.

Study Design and Setting: We take the protocol format suggested by Cochrane for a standard systematic review as our reference and compare the considerations for a pairwise review with those required for a valid comparison of multiple interventions. We suggest new sections for protocols of systematic reviews including network meta-analyses with a focus on how to evaluate their assumptions. We provide example text from published protocols to exemplify the considerations.

Conclusion: Standard systematic review protocols for pairwise meta-analyses need extensions to accommodate the increased complexity of network meta-analysis. Our suggested modifications are widely applicable to both Cochrane and non-Cochrane systematic reviews involving network meta-analyses. © 2017 Elsevier Inc. All rights reserved.
Setting the rational for the review

• **Title:** Identify the review as one that compares multiple interventions

• **Clarify why a NMA is necessary**
  o lack of (many) direct comparisons between the treatments of interest
  o aim to comprehensively rank all treatments

• **Example:** Safety of antiepileptic drugs:

  “Some AEDs have been associated with increased risk of harm to the fetus and infants. [...] many studies have produced inconsistent findings regarding harm to the fetus and infant with use of other agents. As such, our objective is to evaluate the comparative safety of AEDs for infants and children who were exposed in utero or during breastfeeding through a systematic review and network meta-analysis”

  *Tricco et al. Syst Rev 2014*
Frame the research question

- Define the PICO
  - Keep in mind the transitivity assumption!

One major assumption
underlying network meta-analyses

- Conceptual definition
- Manifestation in the data
  - Transitivity
  - Coherence
Transitivity

The underlying assumption when B versus C is calculated *indirectly* is that we can learn about B versus C via A.

Validity depends on **transitivity** of treatment effects across trials making different treatment comparisons

\[
\text{advantage of B over C} = \text{advantage of B over A} + \text{advantage of A over C}
\]

Requires studies to be similar in ways other than the treatments being compared
Ways of thinking about transitivity...

Treatment A must be similar when it appears in AB and AC trials

For example, is it plausible
- when A is placebo given in different forms (e.g. injection versus pill)?
- when A is a drug given in different doses?
Ways of thinking about transitivity...

- **Example:** When comparing different fluoride treatments, comparison between fluoride toothpaste and fluoride rinse can be made via placebo
  - 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

The definition of the nodes in the treatment network is a **challenging issue**

*Salanti et al., JCE 2009*
Transitivity at the protocol stage

• Consider whether ‘missing’ arms are likely missing at random
  o AC trials do not have B arms and AB trials do not have treatment C
  o Is this reasonable? In some clinical areas patients would never receive alternative treatments
    o e.g. Sequencing of drugs

• Consider if all treatments are “jointly randomizable”
  o The treatments need to be genuinely competing alternatives
  o It should possible to imagine a randomized trial comparing all treatments in the network
  o Could patients have been randomly allocated to any of the treatments?
    o e.g. first- and second-line chemotherapy regimens
Transitivity at the protocol stage

- Consider the distribution of possible effect modifiers of the relative treatment effects in AC and AB trials
  - identify a priori potential effect modifiers and compare how they are distributed across comparisons (see data extraction)
  - e.g. patients, trial protocols, doses, administration, etc. should be similar in ways which might modify the treatment effect

![Graphs showing distribution of age in Placebo vs B and Placebo vs C comparisons.]
What to keep in mind for the eligible interventions

• Restricting your review to compare few interventions
  o limits its usefulness and applicability
  o you must justify your choice
  o risk to have unconnected networks
  o few data, low power (depends on the setting)

• Expanding the database too much to include many treatments
  o jeopardizes the transitivity assumptions (or at least makes its defense challenging)
  o renders review process long and data management difficult

• **Watch out for:** old and new treatments, ad-on treatments, intransitive legacy treatments

• **What will you do if you identify new interventions while scanning the literature?**

• **How to deal with different doses or drug class and co-interventions?**
  o Merging versus splitting
The treatments we compare are in principle jointly randomizable. The groups of studies that compare them do not differ with respect to the distribution of effect modifiers. Direct and indirect treatment effects are in statistical agreement.

At the outset

The treatments we compare are in principle jointly randomizable.

They have the same indication, we can imagine a mega-trial with all treatments being compared etc.

Looking at the studies

The groups of studies that compare them do not differ with respect to the distribution of effect modifiers.

Can be tested with enough studies per comparison.

Analysing the data

Direct and indirect treatment effects are in statistical agreement.

Various statistical tests if there is both direct and indirect evidence.

Thinking about transitivity

Searching, selecting studies and extracting data

• Any study comparing at least two of the eligible interventions should be considered
  o i.e. all available direct comparisons between the eligible interventions should be included

• Describe you will extract data on
  o **Outcomes:** study-level or arm-level preferable?
  o **Potential effect modifiers:**
    o population and study characteristics that may act as effect modifiers selected based on bibliography and clinical understanding
    o required to evaluate statistically the transitivity assumption and clinical/methodological heterogeneity
    o used also in additional analyses to explain statistical heterogeneity/incoherence
  o **Risk of bias data, etc.**
Selecting effect measures

• Usual considerations between OR/RR/RD or MD/SMD

• State which **measure will be used to rank the treatments** (if an objective of the review)
  - Avoid probability of being the best
  - Use SUCRAs/P-scores/mean ranks instead


<table>
<thead>
<tr>
<th>Treatment</th>
<th>Prob of being best</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40%</td>
</tr>
<tr>
<td>B</td>
<td>33%</td>
</tr>
<tr>
<td>C</td>
<td>27%</td>
</tr>
</tbody>
</table>
Selecting effect measures

• Usual considerations between OR/RR/RD or MD/SMD

• State which **measure will be used to rank the treatments** (if an objective of the review)
  o Avoid probability of being the best
  o Use SUCRAs/P-scores/mean ranks instead


<table>
<thead>
<tr>
<th>Treatment</th>
<th>Prob of being best</th>
<th>SUCRA/ P-score</th>
<th>Mean Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>40%</td>
<td>67%</td>
<td>2</td>
</tr>
<tr>
<td>B</td>
<td>33%</td>
<td>67%</td>
<td>2</td>
</tr>
<tr>
<td>C</td>
<td>27%</td>
<td>67%</td>
<td>2</td>
</tr>
</tbody>
</table>

Treatments with large uncertainty can be favoured by P(best)!
Cautious note about ranking

- Ranking measures are not substitutes for relative effect estimates

- Ranking based on SUCRAs or mean ranks accounts better for the uncertainty in relative ranking
  - Using $P(\text{best})$ to rank treatments can be misleading

- Ranking measures are conditional on the set of treatments being compared
  - SUCRAs and mean ranks will change when only a subset of interventions are compared

- Avoid ranking when there is a lot of uncertainty in the effect estimates or when there are important differences in the uncertainty across comparisons

- Methods that allow more information in ranking are available
Evaluating transitivity

• Describe how you will evaluate the clinical and methodological comparability of studies (**heterogeneity**)
  o as in standard meta-analysis

• Describe how you will evaluate the plausibility of the **transitivity assumption**
  o the comparability/similarity of studies evaluating *different* comparisons
  o we can compare the distribution of effect modifiers across sets of studies grouped by comparison
  o in practice this is often difficult – be prepared and remember lack of evidence is not evidence of lack
Evaluating transitivity

• Example: psychological interventions for bipolar disorder

“To infer about the assumption of transitivity:
1. We will assess whether the included interventions are similar when they are evaluated in RCTs with different designs; for example, whether interventions are administered the same way in studies comparing active treatments to usual care (or no treatment) and in those comparing active treatments to other active treatments.
2. We will compare the distribution of the potential effect modifiers across the different pairwise comparisons (see ‘Data extraction and management’ for the list of potential effect modifiers). If the distributions are balanced across comparisons we will conclude against evidence of intransitivity.”

(not against intransitivity!)

Cipriani et al. PROSPERO 2015:CRD42015016085
Describing the statistical analysis

• Two possible types of analyses:
  o A series of independent pairwise meta-analyses (usually as the first step of NMA)
  o Network meta-analysis
  o State whether both types of analyses will be performed
    o if the required assumptions are plausible

• Describe the statistical model
  o Bayesian or frequentist setting
  o fixed or random effects
  o common or different heterogeneity across the comparisons

• Report the modelling details (e.g. method for heterogeneity, prior distributions)

• Explain how you will handle variability in treatment definition (e.g. different doses or modalities)
  o Analyze as separate treatment modes nodes
  o Model explicitly their variability
  o Additive/multiplicative models for complex interventions?

• Report the software of the analysis
  o e.g. STATA, R, BUGS
  o give the codes
Evaluating incoherence

Report on methods for:

- **Assessment of incoherence locally**
  - identify pairwise comparisons or loops of evidence that might be important sources of incoherence
  - e.g. node-splitting approach

\[
\mu_{dir} - \mu_{ind} = IF
\]

...using information from the entire network
Evaluating incoherence

Report on methods for:

- **Assessment of incoherence locally**
  - identify pairwise comparisons or loops of evidence that might be important sources of incoherence
  - e.g. node-splitting approach

- **Assessment of incoherence globally**
  - evaluate the presence of incoherence in the entire network
  - e.g. design-by-treatment interaction model
  - Compare coherence vs incoherence models

The coherence model

\[
\mu_{BC} = \mu_{AC} - \mu_{AB}
\]
Evaluating incoherence

Report on methods for:

- **Assessment of incoherence locally**
  - identify pairwise comparisons or loops of evidence that might be important sources of incoherence
  - e.g. node-splitting approach

- **Assessment of incoherence globally**
  - evaluate the presence of incoherence in the entire network
  - e.g. design-by-treatment interaction model
  - Compare coherence vs incoherence models

The incoherence model

\[
\mu_{BC} = \mu_{AC} - \mu_{AB} + w_{ABC}
\]
Investigating heterogeneity and incoherence

• Heterogeneity & incoherence → caused by differences in populations and study characteristics within and across comparisons

• Specify the additional analyses that will be performed to explain heterogeneity and inconsistency
  o e.g. subgroup analyses, network meta-regression (if sufficient data are available)

• Pre-specify the variables that will be considered as possible sources of heterogeneity and incoherence
  o choose a subset of the potential effect modifiers listed earlier (see also Data Extraction)
Reporting bias and small-study effects

• It is as much of a threat as in pairwise meta-analysis

• Use **contour-enhanced funnel plots** (per comparison)
  *Peters et al. JCE 2008;61(10):991-6*

• Use **comparison-adjusted funnel plots** (for the entire network)
  *Chaimani et al. PlosOne 2013*

• Use **network meta-regression models**
  *Chaimani and Salanti ResSynthMeth 2012*

Require assumptions about the direction of potential small study effects!

• Judge how comprehensive was the literature search and whether unpublished studies have been identified

• Use **selection models** in the case of serious reporting bias
Evaluating confidence in the evidence

RESEARCH METHOD

A GRADE Working Group approach for assessing confidence in the evidence of treatment effect estimates from network meta-analysis

Network meta-analysis (NMA), combining direct and indirect evidence, is now widely used to examine the comparative effectiveness of medical interventions. However, there is little guidance on how to rate the quality of evidence supporting treatment comparisons. We present a four-step approach to rate the quality of evidence of treatment effects from NMA estimates, based on methods developed by the GRADE Working Group. We illustrate this approach with an example in which the quality of evidence for a published NMA, we show that the quality of evidence supporting the treatment effects was very low across comparisons, and that quality ratings given to the NMA were not very high and likely to mislead.

New Results

Assessing Confidence in the Results of Network Meta-Analysis (Cinema)

Adriani Nikolakopoulou, Julian PT Higgins, Theodore Papakonstantinou, Anna Chaimani, Cnzia Del Giovane, Matthias Egger, Georgia Salanti

doi: https://doi.org/10.1101/597047

This article is a preprint and has not been peer-reviewed [what does this mean?].
Network meta-analysis in Cochrane

• 104 Cochrane reviews and 50 protocols have NMA in the title, abstract, or keyword as of October 16, 2019

• Examples of recently published NMAs
  o Antibiotic treatment for spontaneous bacterial peritonitis in people with decompensated liver cirrhosis: a network meta-analysis
  o Interventions for maintenance of surgically induced remission in Crohn’s disease: a network meta-analysis
  o Treatment for hepatorenal syndrome in people with decompensated liver cirrhosis: a network meta-analysis
  o Interventions for unexplained infertility: a systematic review and network meta-analysis
  o Pharmacological interventions for the treatment of delirium in critically ill adults
Comparing Multiple Interventions Methods Group

• Registered in October 2010; with a little over 100 members in Archie
• Covers methodology of (1) NMA and (2) overviews of reviews
• Current co-covenors:

Deborah Caldwell
University of Bristol, UK

Anna Chaimani
Paris Descartes University, France

Tianjing Li
University of Colorado Anschutz Medical Campus, USA

Lisa Hartling
University of Alberta, Canada
Training events

• We strongly recommend that all reviews include NMAs have a statistician in the review team

• Numerous workshops at Cochrane Colloquia

• Upcoming short courses
  • 3-day NMA course in Paris: From planning to publication (4-6 December 2019, registration deadline 20 November 2019) (http://livenetworkmetaanalysis.com/nma-training/)
  • 2-day course at University of Bristol (11-12 December) (https://www.epi-winterschool.org)
  • 3-day course at Swiss Epidemiology Winterschool (January 2020) (https://www.epi-winterschool.org)
  • 2-day course at EPIsummer of Columbia University (June 2020) (https://www.mailman.columbia.edu/research/episummercolumbia)