Answering complex hierarchy questions in network meta-analysis

Adriani Nikolakopoulou
Institute of Medical Biometry and Statistics,
Freiburg, Germany

15 November 2022
Network meta-analysis (NMA)

NMA synthesises both **direct and indirect evidence** in a network of trials that contain multiple interventions.

Can give valuable insight into the **comparative benefits and harms** of multiple alternative treatment options.
Network meta-analysis (NMA)

NMA synthesises both **direct and indirect evidence** in a network of trials that contain multiple interventions. It can give valuable insight into the **comparative benefits and harms** of multiple alternative treatment options.

**NMA output**
- All relative treatment effects
- A treatment hierarchy
Presentation of NMA treatment effects

| Treatment  | OR    | (95% CI) | Treatment  | OR    | (95% CI) |
|------------|-------|----------|------------|-------|----------|
| Treatment A1 | 1.66  | (0.96, 2.84) | Treatment B1 | 1.00  | (1.00, 1.00) |
| Treatment A2 | 1.37  | (0.96, 1.95) | Treatment B2 | 0.83  | (0.44, 1.55) |
| Treatment A3 | 1.33  | (1.10, 1.62) | Treatment B3 | 0.81  | (0.46, 1.41) |
| Treatment A4 | 1.27  | (1.05, 1.54) | Treatment B4 | 0.77  | (0.44, 1.34) |
| Treatment A5 | 1.24  | (1.05, 1.47) | Treatment B5 | 0.75  | (0.43, 1.30) |
| Treatment A6 | 1.21  | (0.98, 1.49) | Treatment B6 | 0.73  | (0.41, 1.28) |
| Treatment A7 | 1.19  | (1.04, 1.36) | Treatment B7 | 0.72  | (0.42, 1.24) |
| Treatment A8 | 1.19  | (1.02, 1.38) | Treatment B8 | 0.72  | (0.43, 1.21) |
| Treatment A9 | 1.14  | (0.90, 1.45) | Treatment B9 | 0.69  | (0.39, 1.22) |
| Treatment A10 | 1.12 | (0.88, 1.42) | Treatment B10 | 0.68 | (0.38, 1.21) |
| Treatment A11 | 1.11  | (0.94, 1.30) | Treatment B11 | 0.67  | (0.39, 1.16) |
| Treatment A12 | 1.06  | (0.87, 1.30) | Treatment B12 | 0.64  | (0.36, 1.13) |
| Treatment A13 | 1.04  | (0.71, 1.53) | Treatment B13 | 0.63  | (0.33, 1.21) |
| Treatment A14 | 1.00  |         | Treatment B14 | 0.60  | (0.35, 1.04) |
| Treatment A15 | 1.00  | (0.80, 1.24) | Treatment B15 | 0.60  | (0.34, 1.07) |
| Treatment A16 | 0.99  | (0.79, 1.24) | Treatment B16 | 0.60  | (0.34, 1.06) |
| Treatment A17 | 0.88  | (0.68, 1.15) | Treatment B17 | 0.53  | (0.30, 0.96) |
| Treatment A18 | 0.84  | (0.62, 1.15) | Treatment B18 | 0.51  | (0.28, 0.93) |
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

Andrea Cipriani, Toshi A Furukawa*, Georgia Solantti*, Anna Chaimani, Lauren Z Atkinson, Yusuke Ogawa, Stefan Leucht, Henricus G Ruhe, Erick H Turner, Julian P T Higgins, Matthias Egger, Nozomi Takeshima, Yu Hayasaka, Hissei Imai, Kiyomi Shinohara, Aran Tajika, John P A Ioannidis, John R Geddes

| Treatment      | OR (95% CI) | Treatment      | OR (95% CI) |
|----------------|-------------|----------------|-------------|
| Vortioxetine   | 1.66 (0.96, 2.84) | Vortioxetine   | 1.00        |
| Bupropion      | 1.37 (0.96, 1.95) | Bupropion      | 0.83 (0.44, 1.55) |
| Escitalopram   | 1.33 (1.10, 1.62) | Escitalopram   | 0.81 (0.46, 1.41) |
| Mirtazapine    | 1.27 (1.05, 1.54) | Mirtazapine    | 0.77 (0.44, 1.34) |
| Amitriptyline  | 1.24 (1.05, 1.47) | Amitriptyline  | 0.75 (0.43, 1.30) |
| Agomelatine    | 1.21 (0.98, 1.49) | Agomelatine    | 0.73 (0.41, 1.28) |
| Paroxetine     | 1.19 (1.04, 1.36) | Paroxetine     | 0.72 (0.42, 1.24) |
| Venlafaxine    | 1.19 (1.02, 1.38) | Venlafaxine    | 0.72 (0.43, 1.21) |
| Duloxetine     | 1.14 (0.90, 1.45) | Duloxetine     | 0.69 (0.39, 1.22) |
| Milnacipran    | 1.12 (0.88, 1.42) | Milnacipran    | 0.68 (0.38, 1.21) |
| Sertraline     | 1.11 (0.94, 1.30) | Sertraline     | 0.67 (0.39, 1.16) |
| Citalopram     | 1.06 (0.87, 1.30) | Citalopram     | 0.64 (0.36, 1.13) |
| Nefazodone     | 1.04 (0.71, 1.53) | Nefazodone     | 0.63 (0.33, 1.21) |
| Fluoxetine     | 1.00         | Fluoxetine     | 0.60 (0.35, 1.04) |
| Clomipramine   | 1.00 (0.80, 1.24) | Clomipramine   | 0.60 (0.34, 1.07) |
| Fluvoxamine    | 0.99 (0.79, 1.24) | Fluvoxamine    | 0.60 (0.34, 1.06) |
| Trazodone      | 0.88 (0.68, 1.15) | Trazodone      | 0.53 (0.30, 0.96) |
| Reboxetine     | 0.84 (0.62, 1.15) | Reboxetine     | 0.51 (0.28, 0.93) |

Inferior to Fluoxetine Superior to Fluoxetine

Inferior to Vortioxetine Superior to Vortioxetine
Looking at all treatment effects is recommended.
Motivation - Outline

Producing a treatment hierarchy is very useful and at the same time debatable

43% of published NMAs present some form of treatment hierarchy

Petropoulou M, Nikolakopoulou A, Veroniki A-A, Rios P, Vafaei A, Zarin W, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. J Clin Epidemiol. 2016

Where does the usefulness of ranking comes from?

It is easier to highlight more clearly individual differences between treatments
Looking at all treatment effects is recommended

Ranking metrics summarise this table in different ways.
Ranking Metrics

Methodologists debate several issues underpinning the ranking metrics obtained from NMA

Main criticisms

- They are clinically not relevant
- They are difficult to interpret

Looking at all treatment effects is recommended

What question do ranking metrics answer?

- Is it clinically not relevant?
- Is it difficult to interpret?
Probability of being best
(or having the best mean outcome value)

| % probability | A   | B   | C   | D   |
|---------------|-----|-----|-----|-----|
| j=1           | 0.25| 0.50| 0.25| 0.00|
| j=2           | 0.25| 0.25| 0.50| 0.00|
| j=3           | 0.25| 0.25| 0.25| 0.25|
| j=4           | 0.25| 0.00| 0.00| 0.75|

i = A, B, C, D the treatment
j the rank

Compute for each treatment the probability of being at each possible position

Derived in a Bayesian or in a frequentist framework using a resampling method

What is the probability that A is first?

Treatment hierarchy question: Which treatment is most likely to have the best (most desirable) mean value on the studied outcome?

What is the probability that C is second?
What is the probability that A is first or second?

What is the probability that D is among the best three options?
Surface under the cumulative ranking curve

The areas under the cumulative curves (SUCRA) for the four treatments of the example above are
A=0.5
B=0.75
C=0.67
D=0.08

Treatment hierarchy question: Which treatment has the largest fraction of competitors that it beats?

Q5
Primary outcome: efficacy, defined as at least 50% reduction in the symptoms’ scales between baseline and 8 weeks of follow up.
What is the probability that Vortioxetine ranks first, Bupropion second and Escitalopram third?

What is the probability that Vortioxetine, Bupropion and Escitalopram are the best three treatments?

What is the probability that Vortioxetine has better outcome value than that of Bupropion and Bupropion has better outcome value than that of Escitalopram?

What is the probability that Vortioxetine, Bupropion and Escitalopram have an odds ratio of 1.25 or higher against Fluoxetine?

What is the probability that Vortioxetine, Bupropion and Escitalopram have an odds ratio of 1 or higher against Fluoxetine?
Approach

Perform NMA

Set a number of criteria

Which hierarchies satisfy these constraints?

How certain we are about each one of these hierarchies?

*Computing all $T!$ hierarchies is computationally intensive but is not needed. Only the most frequent ones are recorded.*

Answering complex hierarchy questions in network meta-analysis

Theodoros Papakonstantinou$^{1,2}$, Georgia Salanti$^2$, Dimitris Mavridis$^{3,4}$, Gerta Rücker$^1$, Guido Schwarzer$^1$ and Adriani Nikolakopoulou$^{1,2}$

Abstract

**Background:** Network meta-analysis estimates all relative effects between competing treatments and can produce a treatment hierarchy from the most to the least desirable option according to a health outcome. While about half of the published network meta-analyses present such a hierarchy, it is rarely the case that it is related to a clinically relevant decision question.

**Methods:** We first define treatment hierarchy and treatment ranking in a network meta-analysis and suggest a simulation method to estimate the probability of each possible hierarchy to occur. We then propose a stepwise approach to express clinically relevant decision questions as hierarchy questions and quantify the uncertainty of the criteria that constitute them. The steps of the approach are summarized as follows: a) a question of clinical relevance is defined, b) the hierarchies that satisfy the defined question are collected and c) the frequencies of the respective hierarchies are added; the resulting sum expresses the certainty of the defined set of criteria to hold. We then show how the frequencies of all possible hierarchies relate to common ranking metrics.

**Results:** We exemplify the method and its implementation using two networks. The first is a network of four treatments for chronic obstructive pulmonary disease where the most probable hierarchy has a frequency of 28%. The second is a network of 18 antidepressants, among which Vortioxetine, Bupropion and Escitalopram occupy the first three ranks with frequency 19%.

**Conclusions:** The developed method offers a generalised approach of producing treatment hierarchies in network meta-analysis, which moves towards attaching treatment ranking to a clear decision question, relevant to all or a subset of competing treatments.

**Keywords:** Clinically relevant question, Indirect evidence, Probabilistic ranking, Evidence synthesis
Set a number of criteria
Which hierarchies satisfy these constraints?
How certain we are about each one of these hierarchies?

Perform NMA

What is the probability that Vortioxetine ranks first, Bupropion second and Escitalopram third?

Derive all possible hierarchies & filter those that satisfy the desired criterion

Add the frequencies of the hierarchies that satisfy the set criterion
The nmarank package

nmarank: Complex Hierarchy Questions in Network Meta-Analysis

Derives the most frequent hierarchies along with their probability of occurrence. One can also define complex hierarchy criteria and calculate their probability. Methodology based on Papakonstantinou et al. (2021) [doi:10.21203/rs.3.rs-858140/v1].

Version: 0.2-3
Depends: R (≥ 3.3.1), meta (≥ 4.19-1), netmeta (≥ 1.5-0), data.tree, mvtnorm, tidyverse
Imports: dplyr, tibble, rlang
Suggests: testthat
Published: 2021-09-19
Author: Adriani Nikolakopoulou [aut], Guido Schwarzer [aut], Theodoros Papakonstantinou [aut, cre]
Maintainer: Theodoros Papakonstantinou <dev at tpapk.com>
License: GPL-3
URL: https://github.com/ESM-ispM-unibe-ch/nmarank
NeedsCompilation: no
Materials: NEWS
In views: MetaAnalysis
CRAN checks: nmarank results

Documentation:

Reference manual: nmarank.pdf

Downloads:

Package source: nmarank_0.2-3.tar.gz
Windows binaries: r-devel: nmarank_0.2-3.zip, r-release: nmarank_0.2-3.zip, r-oldrel: nmarank_0.2-3.zip
macOS binaries: r-release (arm64): nmarank_0.2-3.tar.gz, r-oldrel (arm64): nmarank_0.2-3.tar.gz, r-release (x86_64): nmarank_0.2-3.tar.gz, r-oldrel (x86_64): nmarank_0.2-3.tar.gz

Linking:

Please use the canonical form https://CRAN.R-project.org/package=nmarank to link to this page.
The nmarank package

nmarank: This function specifies the frequencies of hierarchies along with their estimated probabilities and the probability that a specified criterion holds.

Arguments in nmarank
- TE.nma: An object of class netmeta or a matrix with network effects
- condition: Condition that should be satisfied (see later)
- VCOV.nma: Variance-covariance matrix for network estimates
- pooled: A character indicating whether the hierarchy is calculated for the fixed effects (“fixed”) or random effects (“random”) model.
- nsim: Number of simulations
- small.values: A character string specifying whether small treatment effects indicate a “good” or “bad” effect

Output of nmarank
- An object of class nmarank: A list containing:
  - hierarchies: A list of the most frequent hierarchies along with their estimated probability of occurrence
  - probabilityOfSelection: Combined probability of all hierarchies that satisfy the defined condition
**The nmarank package**

**nmarank**: This function specifies the frequencies of hierarchies along with their estimated probabilities and the probability that a specified criterion holds.

**condition**: This function defines a condition that is of interest to be satisfied involving a set of treatments in the network.

**Arguments in condition**
- **fn**: Character string specifying type of condition
- **...**: Function arguments

**Output of condition**
A list with the defined function and its arguments
Details

The following types of conditions are available.

The condition `fn = "sameHierarchy"` checks whether a specific hierarchy occurs. One additional unnamed argument has to be provided in `\ldots`: a vector with a permutation of all treatment names in the network.

The condition `fn = "specificPosition"` checks whether a treatment ranks in a specific position. Two additional unnamed arguments have to be provided in `\ldots`: (1) name of the treatment of interest and (2) a single numeric specifying the rank position.

The condition `fn = "betterEqual"` checks whether a treatment has a position better or equal to a specific rank. Two additional unnamed arguments have to be provided in `\ldots`: (1) name of the treatment of interest and (2) a single numeric specifying the rank position.

The condition `fn = "retainOrder"` checks whether a specific order of two or more treatments is retained anywhere in the hierarchy. One additional unnamed argument has to be provided in `\ldots`: a vector with two or more treatment names providing the order of treatments.

The condition `fn = "biggerCIV"` checks whether the effect of a treatment is bigger than that of a second treatment by more than a given clinically important value (CIV) on an additive scale (e.g. log odds ratio, log risk ratio, mean difference). Three additional unnamed arguments have to be provided in `\ldots`: (1) name of the first treatment, (2) name of the second treatment and (3) a numerical value for the CIV. Note that the actual value of the relative effect is considered independently of whether small values is "good" or "bad".

**Composition of conditions for more complex queries:**

Conditions can be combined to express more complex decision trees. This can be done by using the special operators `%AND%`, `%OR%`, `%XOR%` and the opposite function. The combination should be defined as a binary tree with the use of parentheses. If A, B, C and D are conditions, we can for example combine them into a complex condition E:

\[
E = A \%AND\% (B \%OR\% (\text{opposite(C)} \%XOR\% D))
\]
Example 1: network of 21 antidepressants

```r
#rm(list=ls())
library(devtools)
load_all()

# example 2: depression
# Load data
data("depression")

# Prepare data
p1 <- pairwise(treat = drug_name, event = Responders,
               n = Ntotal, data = depression, studlab = studyID, sm = "OR")

# Conduct network meta-analysis
netp1 <- netmeta(p1)

# Calculate probabilities of hierarchies that satisfy a set of criteria

# criterion A and its probability
sel1 <- condition("specificPosition", "Vortioxetine", 1)
sel2 <- condition("specificPosition", "Bupropion", 2)
sel3 <- condition("specificPosition", "Escitalopram", 3)
criterionA <- (sel1 %AND% (sel2 %AND% sel3))

ranksA <- mmrank(netp1, nsim = 10000, small.values = "bad", condition = criterionA)
print(c("probability of Vor in 1st, Bup in 2nd and Esci in 3"
        + "probabilityOfSelection"))
```

Perform NMA

Set a number of criteria

Which hierarchies satisfy these constraints?

How certain we are about each one of these hierarchies?

"probability of Vor in 1st, Bup in 2nd and Esci in 3"

"0.0939"
What is the probability that Vortioxetine ranks first, Bupropion second and Escitalopram third? 9%

What is the probability that Vortioxetine, Bupropion and Escitalopram are the best three treatments? 19%

What is the probability that Vortioxetine has better outcome value than that of Bupropion and Bupropion has better outcome value than that of Escitalopram? 33%

What is the probability that Vortioxetine, Bupropion and Escitalopram have an odds ratio of 1.25 or higher against Fluoxetine? 45%

What is the probability that Vortioxetine, Bupropion and Escitalopram have an odds ratio of 1 or higher against Fluoxetine? 92%
Welcome to nmarank

Complex Hierarchy Questions in Network Meta-Analysis

This is a demonstration of the nmarank CRAN package

You can proceed to the main page

Upload NMA effects matrix txt file

Upload Variance-Covariance matrix txt file

For netmeta users for the hypothetical net1 netmeta object you can use for example write.table(net1$TE.random) and write(net1$S Cov.random)

You can download the example tables taken from:

Woods BS, Hawkins N, Scott DA (2010): Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: A tutorial. BMC Medical Research Methodology, 10, 54

Network effects matrix

|        | Fluticasone | Placebo  | Salmeterol | SFC  |
|--------|-------------|----------|------------|------|
| Fluticasone | 0.00       | -0.60    | 0.27       | 0.65 |
| Placebo   | 0.60       | 0.00     | 0.86       | 1.25 |
| Salmeterol| -0.27      | -0.86    | 0.00       | 0.39 |
| SFC       | -0.65      | -1.25    | -0.39      | 0.00 |

Variance-Covariance matrix

|                        | Fluticasone:Placebo | Fluticasone:Salmeterol | Fluticasone:SFC | Placebo:Salmeterol | Placebo:SFC | Salmeterol:SFC |
|------------------------|---------------------|------------------------|----------------|--------------------|-------------|----------------|
| Fluticasone:Placebo    | 0.38                | 0.29                   | 0.25           | -0.09              | -0.13       | -0.04          |
| Fluticasone:Salmeterol | 0.29                | 0.50                   | 0.25           | 0.22               | -0.04       | -0.25          |
| Fluticasone:SFC        | 0.25                | 0.25                   | 0.76           | 0.00               | 0.50        | 0.50           |
| Placebo:Salmeterol     | -0.09               | 0.22                   | 0.00           | 0.31               | 0.09        | -0.22          |
| Placebo:SFC            | -0.13               | -0.04                  | 0.50           | 0.09               | 0.63        | 0.54           |
| Salmeterol:SFC         | -0.04               | -0.25                  | 0.50           | -0.22              | 0.54        | 0.75           |
Example 2: Treatments for chronic obstructive pulmonary disease (COPD)

**Primary Outcome:** mortality

- The hierarchy is exactly “SFC, Salmeterol, Placebo, Fluticasone”
- SFC is better than Fluticasone and Fluticasone is better than Placebo. The order “SFC, Fluticasone, Placebo” is retained anywhere in the hierarchy
- Salmeterol is 2\textsuperscript{nd}
- Fluticasone is among the two best options
Methodologists debate several issues underpinning the ranking metrics obtained from NMA.

Main criticisms

- They are clinically not relevant
- They are difficult to interpret
- They are not accompanied by a measure of uncertainty

Uncertainty within each ranking metric

Uncertainty of the entire treatment hierarchy

Research and Reporting Methods

Annals of Internal Medicine

Uncertainty in Treatment Rankings: Reanalysis of Network Meta-analyses of Randomized Trials

Ludovic Trinquart, PhD; Nassima Attliche, MSc; Aïda Bafeta, PhD; Raphaël Porcher, PhD; and Philippe Ravaud, MD, PhD

Background: Ranking of interventions is one of the most appealing elements of network meta-analysis. There is, however, little evidence about the reliability of these rankings.

Purpose: To empirically evaluate the extent of uncertainty in intervention rankings from network meta-analysis.

Data Sources: Two previous systematic reviews that involved searches of the Cochrane Library, MEDLINE, and Embase up to July 2012 for articles that included networks of at least 3 interventions.

Study Selection: 58 network meta-analyses involving 1308 randomized trials and 404 interventions with available aggregated outcome data.

Data Analysis: Each network was analyzed with a Bayesian approach. For each intervention, the surface under the cumulative ranking curve (SUCRA) and its 95% credible interval (95% CI) were estimated. Through use of the SUCRA values, the interven-
Uncertainty of the entire treatment hierarchy

Preliminary suggestion:
look at the shape of rankograms

Idea:
formalize this using our approach

B, C, D, A: higher probability

B, C, D, A: smaller probability

Alternatives:
- Magnitude of most frequent hierarchy
- Summary of the “hierarchy matrix” (e.g. their variance)
- Ratio of most frequent hierarchy to the rest
Uncertainty of the entire treatment hierarchy

Drawback:
All these depend on the number of treatments

Alternative way of judging precision:
Looking at the certainty of the specified criteria of interest
(could be the derived hierarchy by SUCRAs)

Scenario:
Examples with imprecise results but associated with certainty around specific criteria relevant for decision making
Ranking Metrics

Methodologists debate several issues underpinning the ranking metrics obtained from NMA

Main criticisms

- They are clinically not relevant
- They are difficult to interpret
- They are not accompanied by a measure of uncertainty
- They do not account for multiple outcomes
Future directions: multiple outcomes & benefit-harm considerations

a) For the selected hierarchies examine their precision for other outcomes

| Hierarchy | Outcome 1 | Outcome 2 | Outcome 3 |
|-----------|-----------|-----------|-----------|
| B, C, D, A | 28% | 10% | 35% |

b) Sample separately or simultaneously from two or more outcomes and measure the frequency for each one of the possible hierarchies for all outcomes

\[ P(A = 4 \cap B = 1 \cap C = 2 \cap D = 3)_{01} \cap P(A = 4 \cap B = 1 \cap C = 2 \cap D = 3)_{02} \]

but only if the treatments are exactly the same which is rare in practice.

c) Incorporate benefit-harm considerations

d) Apply to a clinical example (either for one or multiple outcomes)
Other approaches

Brignardello-Petersen R, Johnston BC, Jadad AR, Tomlinson G. Using decision thresholds for ranking treatments in network meta-analysis results in more informative rankings. J Clin Epidemiol. 2018 Jun;98:62–9.

Mavridis D, Porcher R, Nikolakopoulou A, Salanti G, Ravaud P. Extensions of the probabilistic ranking metrics of competing treatments in network meta-analysis to reflect clinically important relative differences on many outcomes. Biom J Biom Z. 2019 Oct 29.

Tervonen T, Naci H, van Valkenhoef G, Ades AE, Angelis A, Hillege HL, et al. Applying Multiple Criteria Decision Analysis to Comparative Benefit-Risk Assessment: Choosing among Statins in Primary Prevention. Med Decis Mak Int J Soc Med Decis Mak. 2015 Oct;35(7):859–71.

van Valkenhoef, G., Tervonen, T., Zhao, J., de Brock, B., Hillege, H. L., & Postmus, D. (2012). Multicriteria benefit–risk assessment using network meta-analysis. Journal of Clinical Epidemiology, 65(4), 394–403.

Chaimani A, Porcher R, Sbidian E, Mavridis D. A Markov Chain approach for ranking treatments in network meta-analysis. Statistics in Medicine.

Rücker G, Schwarzer G. Resolve conflicting rankings of outcomes in network meta-analysis: Partial ordering of treatments. Res Synth Methods. 2017 Dec;8(4):526–36.

Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. J Clin Epidemiol 2016; 76: 193 - 199.
Thank you!