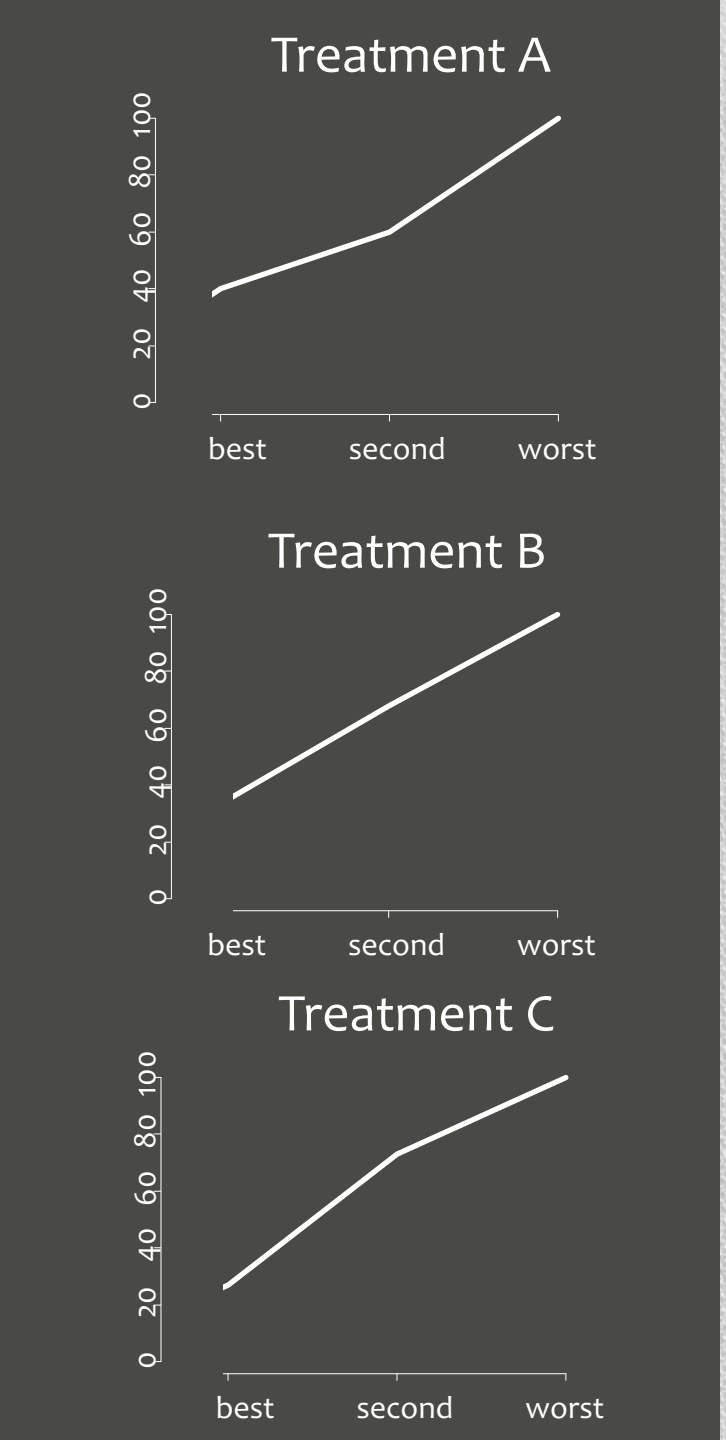


Answering complex hierarchy questions in network meta-analysis

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

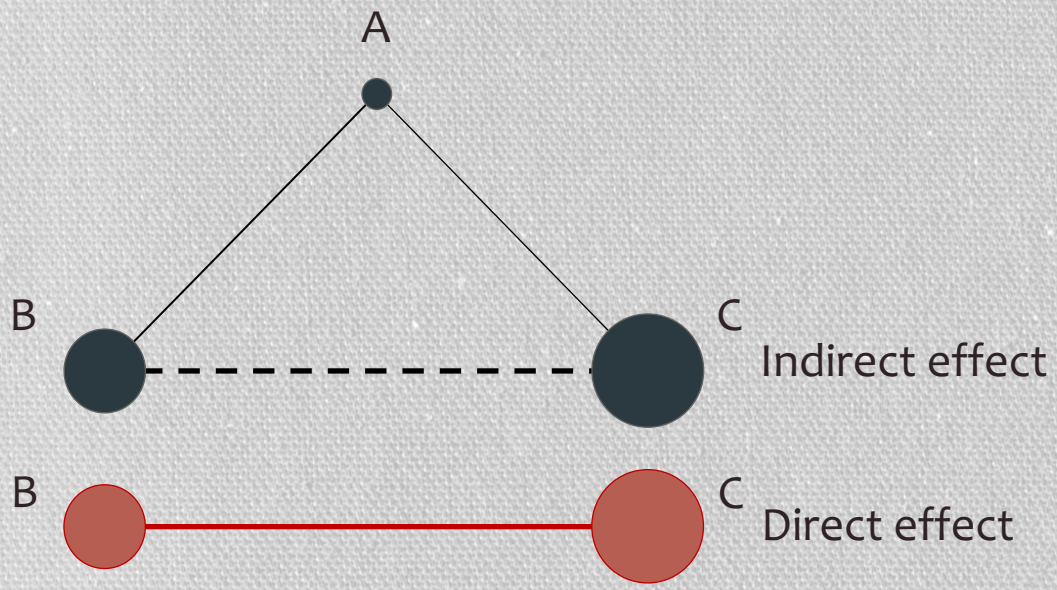
15 November 2022

Q1

Network meta-analysis (NMA)

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

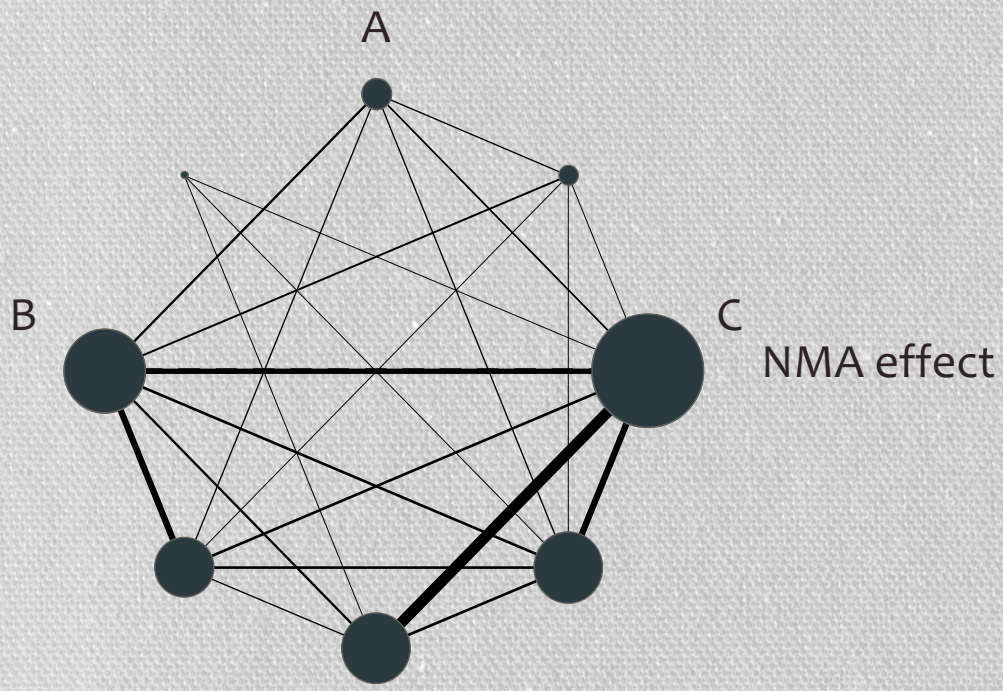
can give valuable insight into the **comparative benefits and harms** of multiple alternative treatment options



Network meta-analysis (NMA)

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

can give valuable insight into the **comparative benefits and harms** of multiple alternative treatment options



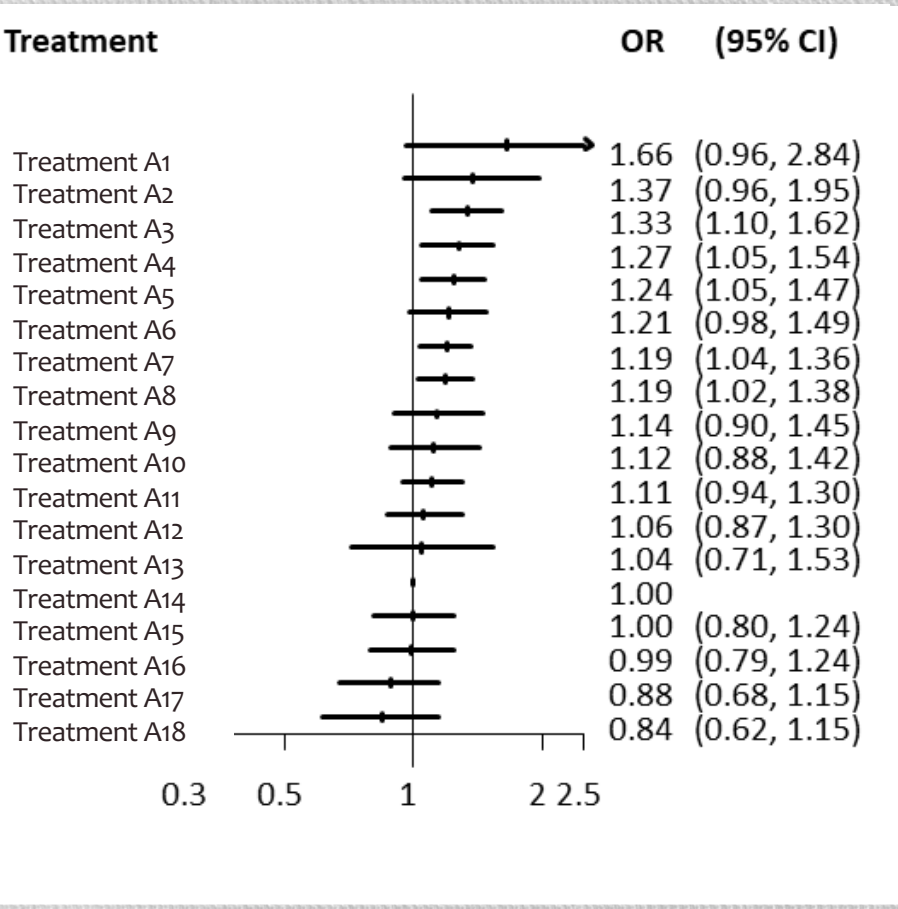
NMA effect

NMA output

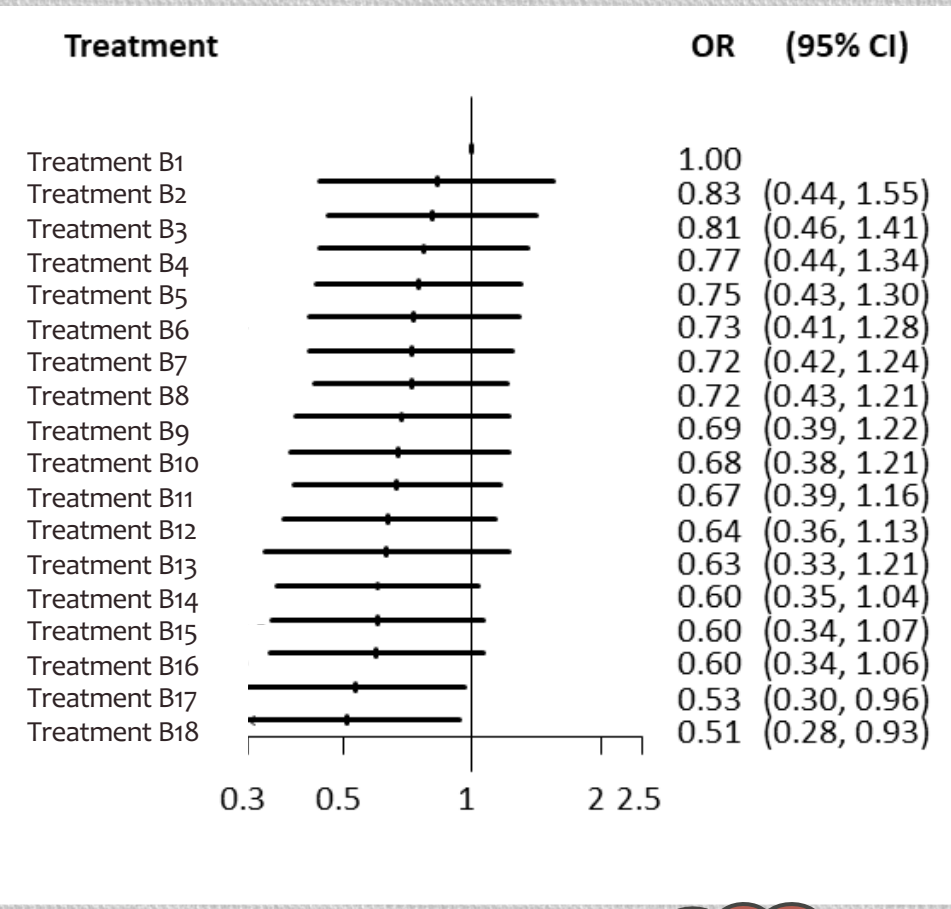
- All relative treatment effects
- A treatment hierarchy

Presentation of NMA treatment effects

NMA A



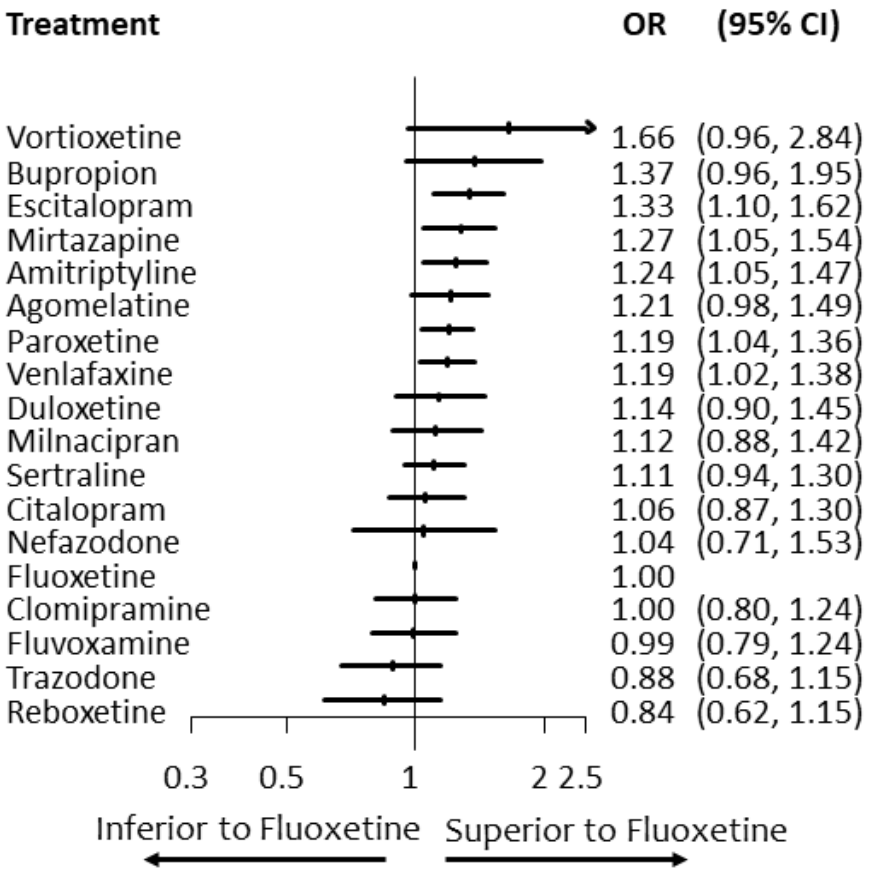
NMA B

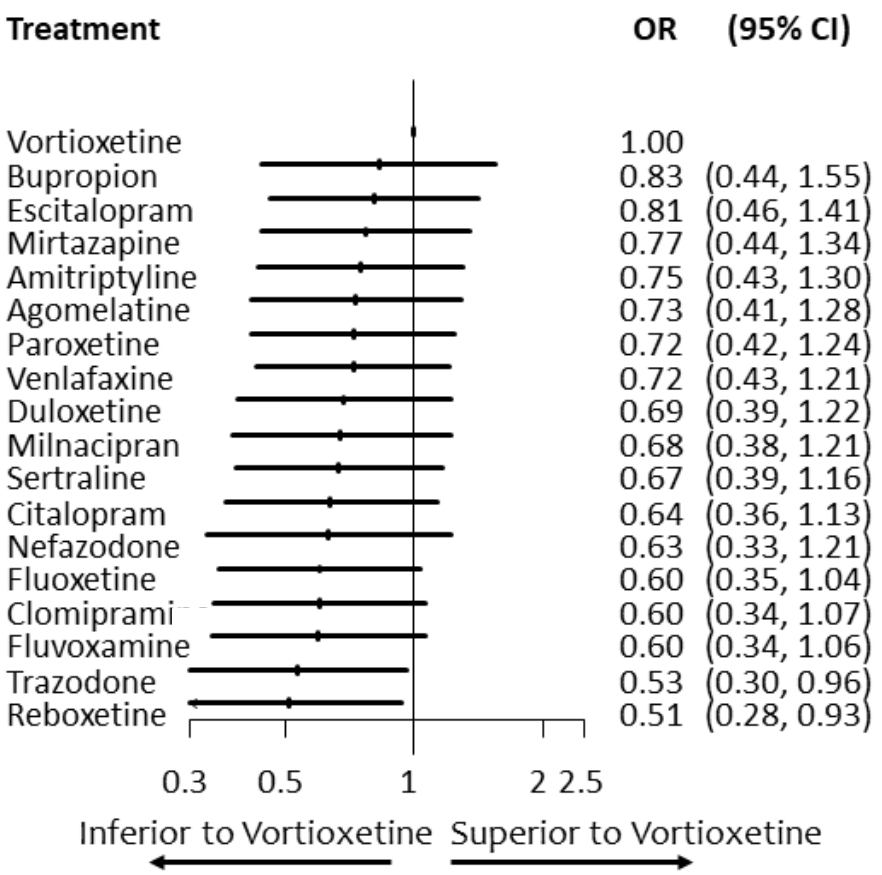


Q2

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 Salanti*, 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*





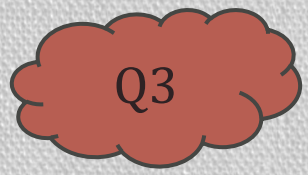
Looking at all treatment effects is recommended

Efficacy (response rate) | Comparison | Acceptability (dropout rate)

Agom	0·72* (0·55–0·92)	0·80* (0·54–1·15)	0·89* (0·66–1·19)	0·57* (0·42–0·77)	0·62† (0·47–0·82)	0·97* (0·74–1·27)	0·85† (0·68–1·05)	0·69† (0·51–0·97)	0·79* (0·58–1·09)	0·81* (0·61–1·05)	0·70* (0·44–1·14)	0·81* (0·65–1·00)	0·53* (0·36–0·80)	0·86* (0·66–1·13)	0·69* (0·48–0·98)	0·74† (0·58–0·92)	1·24† (0·71–2·19)
0·96* (0·76–1·24)	Amit	1·10‡ (0·78–1·58)	1·23* (0·94–1·64)	0·79† (0·60–1·05)	0·87† (0·66–1·15)	1·35* (1·05–1·74)	1·18† (0·99–1·42)	0·97† (0·74–1·24)	1·10† (0·84–1·45)	1·12* (0·89–1·42)	0·98‡ (0·62–1·55)	1·12† (0·95–1·34)	0·74† (0·51–1·10)	1·20* (0·97–1·47)	0·96‡ (0·70–1·31)	1·02† (0·83–1·26)	1·72† (1·00–3·05)
0·87† (0·59–1·30)	0·91‡ (0·62–1·31)	Bupr	1·11‡ (0·76–1·67)	0·71† (0·49–1·07)	0·78† (0·53–1·18)	1·23* (0·84–1·80)	1·07‡ (0·76–1·50)	0·87‡ (0·59–1·30)	1·00‡ (0·66–1·49)	1·01† (0·70–1·47)	0·89‡ (0·51–1·54)	1·02‡ (0·73–1·43)	0·67† (0·42–1·08)	1·08‡ (0·75–1·56)	0·87‡ (0·57–1·30)	0·92‡ (0·66–1·30)	1·55† (0·85–2·94)
1·13* (0·88–1·47)	1·18* (0·93–1·49)	1·30† (0·88–1·93)	Cita	0·64† (0·47–0·87)	0·70* (0·51–0·95)	1·09* (0·85–1·42)	0·96* (0·76–1·21)	0·78* (0·57–1·06)	0·89* (0·64–1·23)	0·91† (0·68–1·21)	0·79‡ (0·49–1·32)	0·91* (0·71–1·17)	0·60† (0·41–0·87)	0·97‡ (0·74–1·25)	0·77* (0·53–1·13)	0·83† (0·64–1·07)	1·40† (0·78–2·48)
1·20* (0·91–1·59)	1·24† (0·98–1·58)	1·37† (0·93–2·04)	1·06* (0·82–1·38)	Clom	1·10† (0·80–1·51)	1·71* (1·27–2·29)	1·49† (1·16–1·90)	1·22† (0·88–1·67)	1·40† (1·00–1·92)	1·41* (1·05–1·91)	1·24‡ (0·76–2·00)	1·42† (1·12–1·79)	0·94‡ (0·62–1·41)	1·51† (1·15–1·96)	1·21† (0·83–1·73)	1·29† (0·99–1·67)	2·20† (1·22–3·90)
1·06* (0·82–1·37)	1·10† (0·84–1·42)	1·21† (0·81–1·81)	0·93* (0·71–1·22)	0·88† (0·66–1·18)	Dulo	1·56* (1·19–2·01)	1·37* (1·06–1·73)	1·12* (0·80–1·53)	1·28† (0·91–1·75)	1·30* (0·96–1·72)	1·13‡ (0·69–1·83)	1·30* (1·02–1·63)	0·86‡ (0·57–1·29)	1·38† (1·04–1·80)	1·10† (0·76–1·59)	1·18‡ (0·92–1·49)	1·99† (1·13–3·52)
0·90* (0·71–1·14)	0·93* (0·74–1·17)	1·03† (0·70–1·51)	0·79* (0·65–0·97)	0·75* (0·58–0·97)	0·85* (0·67–1·08)	Esci	0·87* (0·70–1·09)	0·71* (0·53–0·96)	0·81* (0·60–1·11)	0·83* (0·63–1·08)	0·72† (0·45–1·18)	0·83* (0·67–1·03)	0·55* (0·37–0·81)	0·88* (0·69–1·12)	0·70* (0·49–1·00)	0·75* (0·60–0·94)	1·27‡ (0·73–2·25)
1·20* (0·99–1·48)	1·25† (1·06–1·48)	1·38† (0·97–1·97)	1·06* (0·87–1·29)	1·00‡ (0·81–1·24)	1·14* (0·91–1·44)	1·34* (1·11–1·61)	Fluo	0·82* (0·64–1·04)	0·94* (0·72–1·20)	0·95* (0·77–1·16)	0·83† (0·54–1·30)	0·95* (0·83–1·09)	0·63† (0·44–0·90)	1·01† (0·84–1·21)	0·81* (0·60–1·09)	0·87† (0·74–1·01)	1·46† (0·85–2·53)
1·20* (0·91–1·61)	1·25† (0·99–1·59)	1·38† (0·93–2·07)	1·06* (0·82–1·39)	1·01‡ (0·76–1·32)	1·14† (0·85–1·54)	1·34* (1·03–1·75)	1·00* (0·80–1·25)	Fluv	1·14† (0·84–1·56)	1·16* (0·89–1·52)	1·01‡ (0·62–1·71)	1·16* (0·90–1·49)	0·77† (0·51–1·17)	1·23* (0·94–1·63)	0·99‡ (0·69–1·42)	1·06* (0·80–1·38)	1·78‡ (1·00–3·24)
1·07* (0·80–1·44)	1·11† (0·86–1·43)	1·23† (0·81–1·85)	0·94† (0·71–1·26)	0·89† (0·67–1·19)	1·01‡ (0·74–1·38)	1·19* (0·90–1·58)	0·89* (0·70–1·13)	0·89† (0·67–1·17)	Miln	1·02† (0·75–1·37)	0·88‡ (0·54–1·44)	1·02‡ (0·80–1·31)	0·67† (0·45–1·03)	1·08* (0·82–1·44)	0·86* (0·60–1·25)	0·93* (0·71–1·22)	1·56† (0·89–2·84)
0·93* (0·72–1·21)	0·97* (0·77–1·21)	1·07† (0·73–1·57)	0·82* (0·65–1·05)	0·78* (0·60–1·01)	0·88* (0·67–1·16)	1·04* (0·82–1·32)	0·78* (0·64–0·94)	0·78* (0·60–0·99)	0·87* (0·66–1·15)	Mirt	0·87† (0·55–1·41)	1·00* (0·82–1·23)	0·66* (0·45–0·99)	1·06* (0·84–1·35)	0·85* (0·62–1·18)	0·91* (0·73–1·13)	1·53† (0·89–2·72)
1·15† (0·76–1·76)	1·19† (0·80–1·78)	1·32‡ (0·80–2·20)	1·01‡ (0·67–1·54)	0·96‡ (0·63–1·45)	1·09‡ (0·71–1·68)	1·28* (0·86–1·94)	0·96‡ (0·66–1·40)	0·95‡ (0·63–1·46)	1·07‡ (0·70–1·67)	1·23* (0·82–1·86)	Nefa	1·15‡ (0·74–1·78)	0·75‡ (0·43–1·32)	1·23† (0·77–1·90)	0·98‡ (0·57–1·64)	1·04‡ (0·66–1·65)	1·76† (0·90–3·56)
1·01* (0·82–1·24)	1·05† (0·89–1·23)	1·16† (0·81–1·64)	0·89* (0·72–1·09)	0·84† (0·68–1·03)	0·95† (0·76–1·19)	1·12* (0·93–1·35)	0·84* (0·73–0·95)	0·84* (0·67–1·04)	0·94† (0·75–1·18)	1·08* (0·89–1·30)	0·88‡ (0·60–1·27)	Paro	0·66† (0·46–0·94)	1·06* (0·88–1·28)	0·85† (0·63–1·15)	0·91* (0·77–1·07)	1·53† (0·90–2·66)
1·44* (1·02–2·04)	1·50† (1·07–2·07)	1·65† (1·05–2·60)	1·27† (0·92–1·75)	1·20† (0·84–1·70)	1·36† (0·95–1·95)	1·60* (1·14–2·23)	1·20† (0·88–1·62)	1·20† (0·83–1·71)	1·35† (0·92–1·95)	1·54* (1·09–2·17)	1·25‡ (0·77–2·01)	1·43† (1·05–1·94)	Rebo	1·61† (1·09–2·34)	1·29† (0·81–2·01)	1·38† (0·94–1·99)	2·32† (1·24–4·41)
1·07* (0·85–1·37)	1·11* (0·92–1·35)	1·23† (0·85–1·79)	0·95† (0·76–1·18)	0·90† (0·71–1·13)	1·02‡ (0·79–1·32)	1·20* (0·97–1·48)	0·89‡ (0·76–1·05)	0·89† (0·70–1·13)	1·00† (0·77–1·30)	1·15* (0·93–1·43)	0·93‡ (0·63–1·37)	1·07* (0·90–1·26)	0·75† (0·54–1·04)	Sert	0·80* (0·58–1·11)	0·86* (0·70–1·05)	1·45† (0·84–2·54)
1·36* (0·99–1·87)	1·41† (1·06–1·86)	1·56† (1·04–2·31)	1·20* (0·88–1·63)	1·13† (0·83–1·54)	1·28† (0·92–1·79)	1·51* (1·12–2·04)	1·13† (0·87–1·46)	1·13† (0·82–1·55)	1·27* (0·91–1·76)	1·45* (1·09–1·94)	1·18‡ (0·75–1·84)	1·35* (1·04–1·75)	0·94‡ (0·64–1·39)	1·26† (0·95–1·67)	Traz	1·07‡ (0·77–1·47)	1·80† (0·98–3·38)
1·01* (0·82–1·26)	1·05† (0·87–1·27)	1·16† (0·82–1·65)	0·90† (0·72–1·10)	0·85† (0·67–1·06)	0·96† (0·77–1·21)	1·13* (0·93–1·37)	0·84† (0·73–0·97)	0·84* (0·66–1·07)	0·95* (0·73–1·23)	1·09* (0·89–1·33)	0·88‡ (0·59–1·30)	1·01† (0·86–1·17)	0·70† (0·51–0·97)	0·94* (0·78–1·13)	0·75† (0·57–0·98)	Venl	1·69† (1·01–2·86)
0·73‡ (0·42–1·26)	0·76‡ (0·44–1·29)	0·83‡ (0·45–1·54)	0·64† (0·37–1·11)	0·61† (0·35–1·05)	0·69† (0·40–1·20)	0·81‡ (0·47–1·39)	0·60† (0·36–1·02)	0·60† (0·34–1·05)	0·68† (0·39–1·20)	0·78‡ (0·45–1·34)	0·63† (0·33–1·19)	0·72† (0·43–1·22)	0·51† (0·28–0·92)	0·68† (0·39–1·16)	0·54† (0·30–0·95)	0·72† (0·43–1·19)	Vort

Motivation - Outline

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



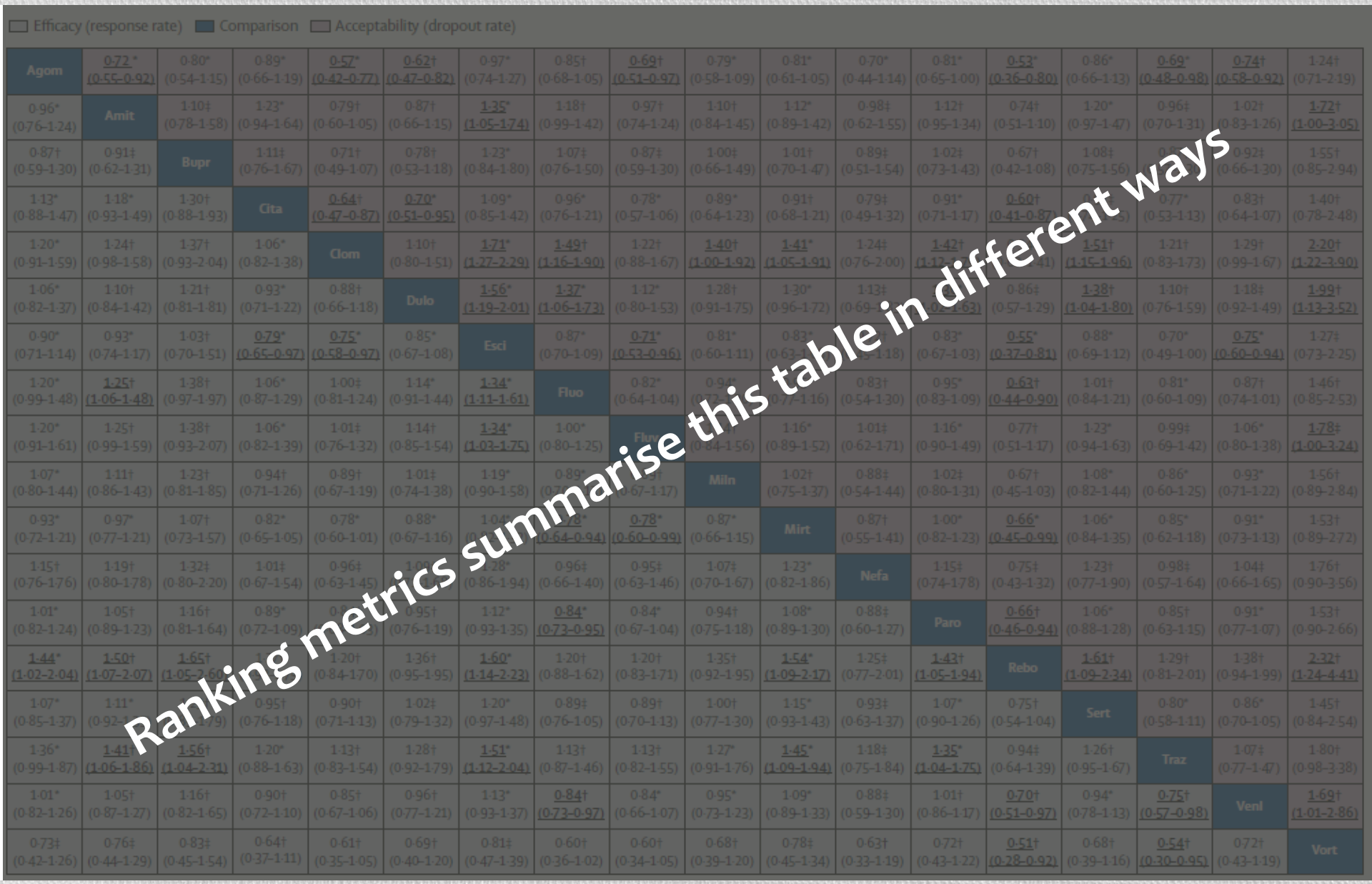
43% of published NMAs present some form of treatment hierarchy

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

Methodologists debate several issues underpinning the ranking metrics obtained from NMA

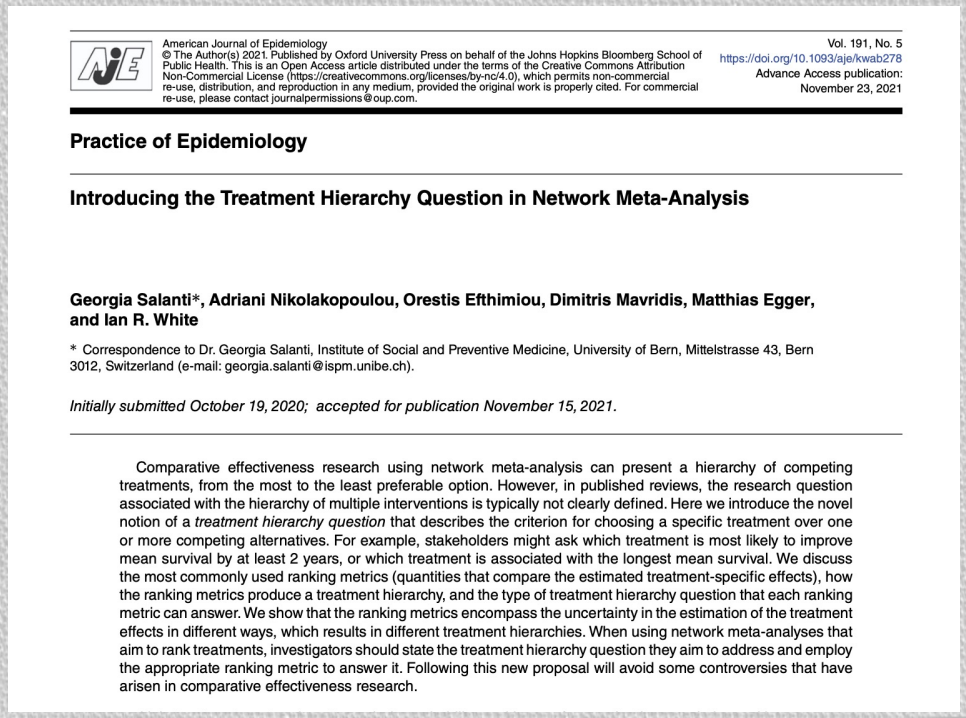
Main criticisms

- 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?



American Journal of Epidemiology

Vol. 191, No. 5
https://doi.org/10.1093/aje/kwab278
Advance Access publication:
November 23, 2021

Practice of Epidemiology

Introducing the Treatment Hierarchy Question in Network Meta-Analysis

Georgia Salanti*, Adriani Nikolakopoulou, Orestis Efthimiou, Dimitris Mavridis, Matthias Egger, and Ian R. White

* Correspondence to Dr. Georgia Salanti, Institute of Social and Preventive Medicine, University of Bern, Mittelstrasse 43, Bern 3012, Switzerland (e-mail: georgia.salanti@ispm.unibe.ch).

Initially submitted October 19, 2020; accepted for publication November 15, 2021.

Comparative effectiveness research using network meta-analysis can present a hierarchy of competing treatments, from the most to the least preferable option. However, in published reviews, the research question associated with the hierarchy of multiple interventions is typically not clearly defined. Here we introduce the novel notion of a *treatment hierarchy question* that describes the criterion for choosing a specific treatment over one or more competing alternatives. For example, stakeholders might ask which treatment is most likely to improve mean survival by at least 2 years, or which treatment is associated with the longest mean survival. We discuss the most commonly used ranking metrics (quantities that compare the estimated treatment-specific effects), how the ranking metrics produce a treatment hierarchy, and the type of treatment hierarchy question that each ranking metric can answer. We show that the ranking metrics encompass the uncertainty in the estimation of the treatment effects in different ways, which results in different treatment hierarchies. When using network meta-analyses that aim to rank treatments, investigators should state the treatment hierarchy question they aim to address and employ the appropriate ranking metric to answer it. Following this new proposal will avoid some controversies that have arisen in comparative effectiveness research.

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?

Q4

Cumulative probabilities of being at each rank

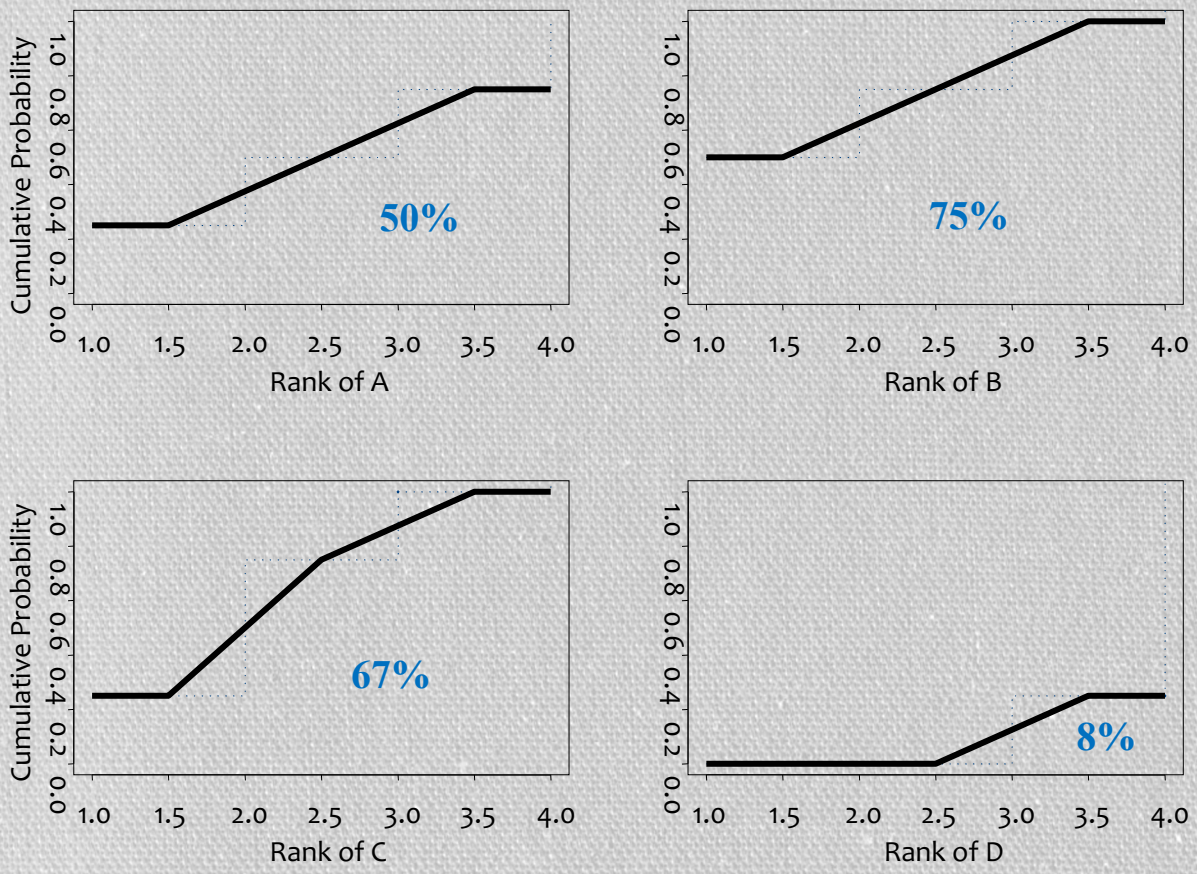
% cumulative probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00
j=2	0.50	0.75	0.75	0.00
j=3	0.75	1.00	1.00	0.25
j=4	1.00	1.00	1.00	1.00

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

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

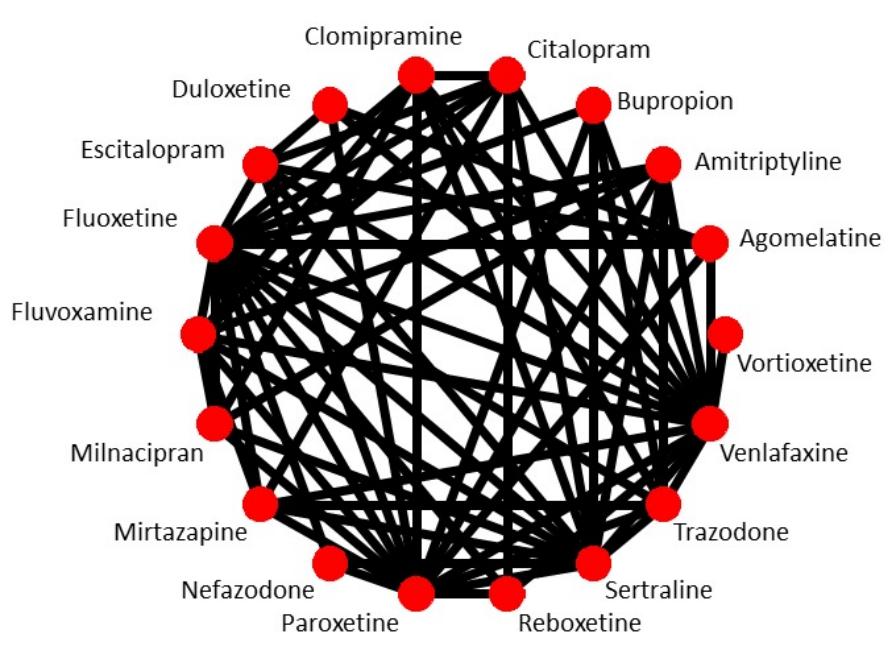
Treatment hierarchy question: Which treatment has the largest fraction of competitors that it beats?

Q5

What if other hierarchy questions are of interest?

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 Salanti*, 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*



Primary outcome: efficacy, defined as at least 50% reduction in the symptoms' scales between baseline and 8 weeks of follow up

Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ, Ogawa Y, et al. 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. Lancet 2018.

What if other hierarchy questions are of interest?

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 Salanti*, 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*

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

↓

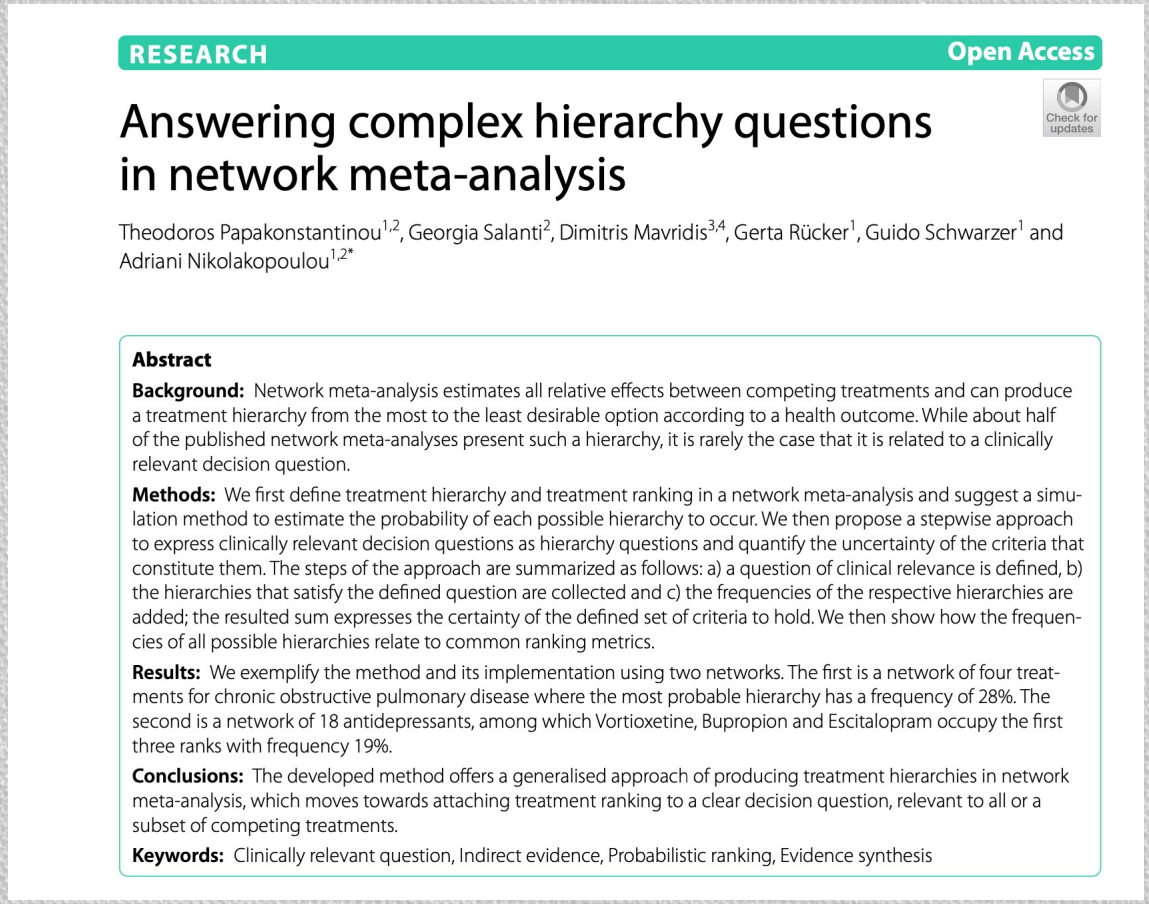
Set a number of criteria

↓

Which hierarchies satisfy these constraints?

↓

How certain we are about each one of these hierarchies?



RESEARCH Open Access

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 resulted 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.

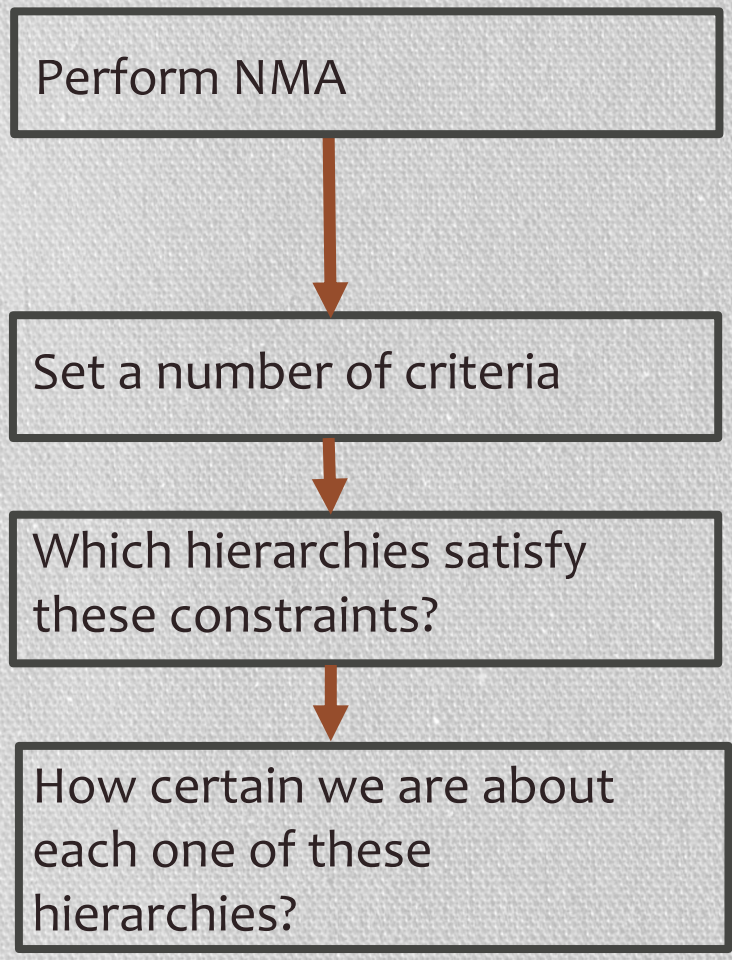
Results: We exemplify the method and its implementation using two networks. The first is a network of four treatments for chronic obstructive pulmonary disease where the most probable hierarchy has a frequency of 28%. The second is a network of 18 antidepressants, among which Vortioxetine, Bupropion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

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

Approach



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, tidyr
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 tpapak.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.tgz, r-oldrel (arm64): nmarank_0.2-3.tgz, r-release (x86_64): nmarank_0.2-3.tgz, r-oldrel (x86_64): nmarank_0.2-3.tgz

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

The nmarank package

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 '...': 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 '...': (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 '...': (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 '...': 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 '...': (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% (opposite(C) %XOR% D))
```

Example 1: network of 21 antidepressants

```
1  rm(list=ls())
2  library(devtools)
3  load_all()
4
5  #example 2: depression
6  # Load data
7  data("depression")
8
9  # Prepare data
10 p1=pairwise(treat=drug_name,event=Responders
11             ,n=Ntotal,data=depression,studlab = studyID, sm="OR")
12
13 # Conduct network meta-analysis
14 netp1=netmeta(p1)
15
16 # Calculate probabilities of hierarchies that satisfy a set of criteria
17
18 #criterion A and its probability
19 sel1 = condition("specificPosition", "Vortioxetine", 1)
20 sel2 = condition("specificPosition", "Bupropion", 2)
21 sel3 = condition("specificPosition", "Escitalopram", 3)
22 criterionA = (sel1 %AND% (sel2 %AND% sel3))
23
24 ranksA = nmarank(netp1,nsim=10000,small.values = "bad", condition=criterionA)
25 print(c("probability of Vor in 1st, Bup in 2nd and Esci in 3"
26         ,ranksA$probabilityOfSelection))
```

42:13 (Top Level)

Console ~/

```
> sel3 = condition("specificPosition", "Escitalopram", 3)
> criterionA = (sel1 %AND% (sel2 %AND% sel3))
> ranksA = nmarank(netp1,nsim=10000,small.values = "bad", condition=criterionA)
> print(c("probability of Vor in 1st, Bup in 2nd and Esci in 3"
+         ,ranksA$probabilityOfSelection))
[1] "probability of Vor in 1st, Bup in 2nd and Esci in 3"
[2] "0.0939"
```

Perform NMA

Set a number of criteria

Which hierarchies satisfy these constraints?

How certain we are about each one of these hierarchies?

Example 1: network of 21 antidepressants

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 Salanti*, 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*

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%**

R shiny

nmarank | Data upload | Main

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

Browse... No file selected

Upload Variance-Covariance matrix txt file

Browse... No file selected

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

Download the example

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

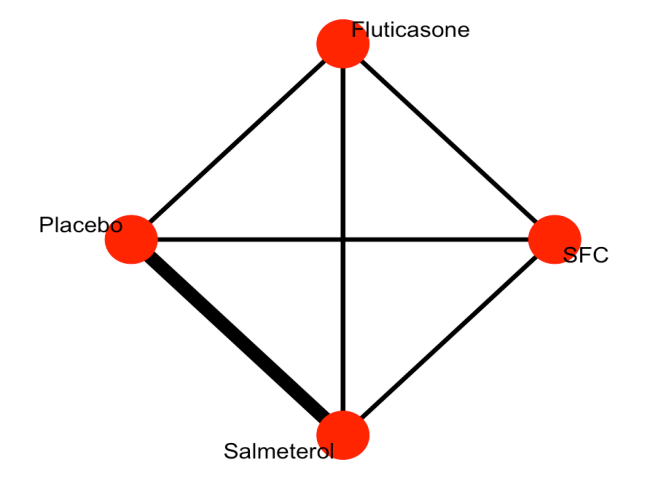
Variance-Covariance matrix

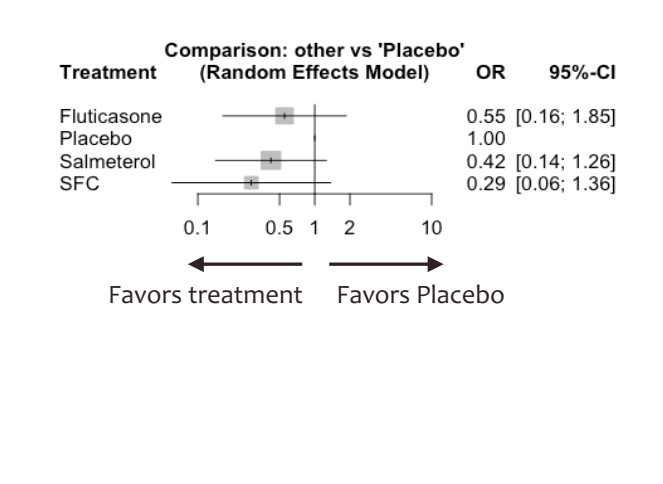
	Fluticasone:Placebo	Fluticasone:Salmeterol	Fluticasone:SFC	Placebo:Salmeterol	Placebo:SFC	Salmeterol:SFC
Fluticasone:Placebo	0.38	0.29	0.25	-0.09	-0.13	-0.04
Fluticasone:Salmeterol	0.29	0.50	0.25	0.22	-0.04	-0.25
Fluticasone:SFC	0.25	0.25	0.76	0.00	0.50	0.50
Placebo:Salmeterol	-0.09	0.22	0.00	0.31	0.09	-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

https://thodoris-papakonstantinou.shinyapps.io/nmarankshiny/

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 2nd
- Fluticasone is among the two best options

Ranking Metrics

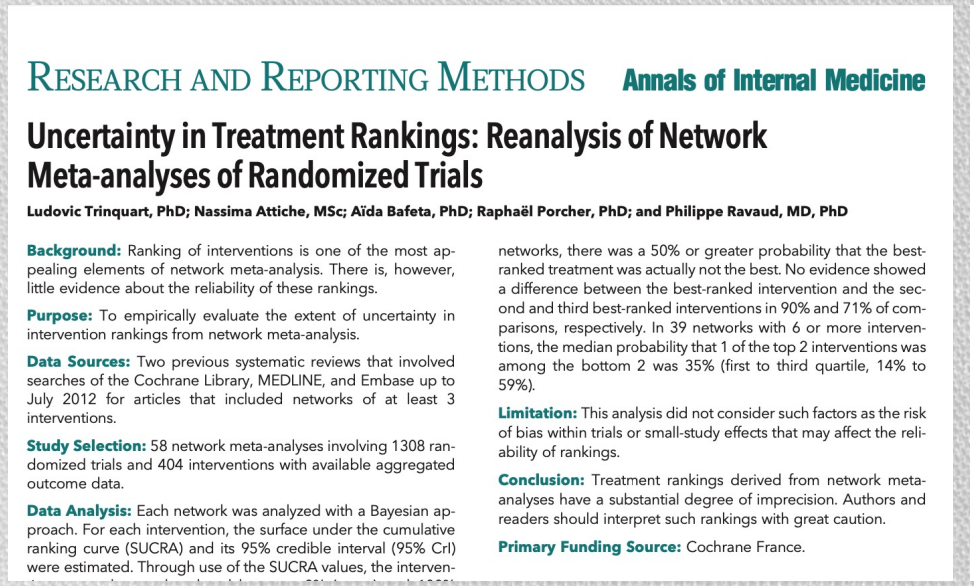
Methodologists debate several issues underpinning the ranking metrics obtained from NMA

Main criticisms

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

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 Attiche, 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.

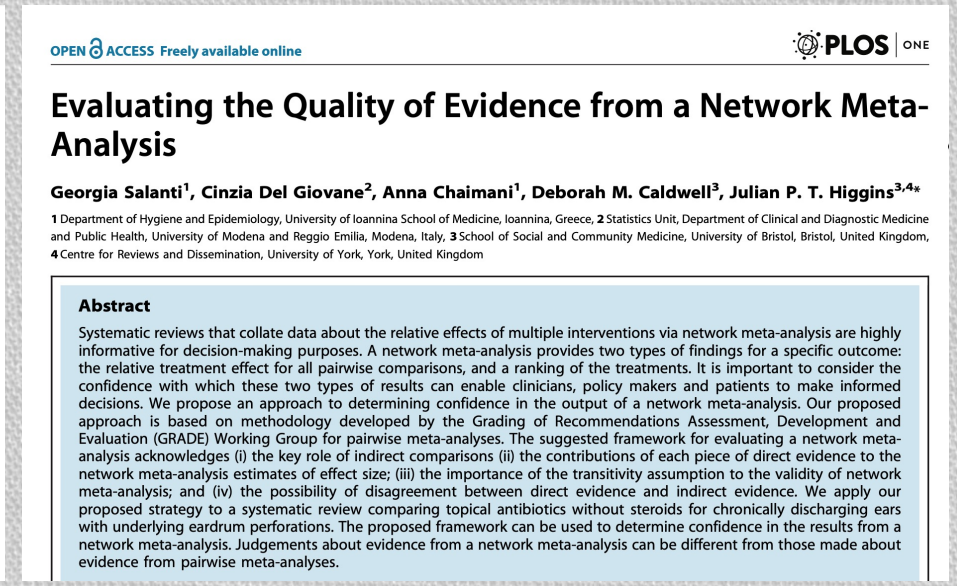
Data Analysis: Each network was analyzed with a Bayesian approach. For each intervention, the surface under the cumulative ranking curve (SUCRA) and its 95% credible interval (95% CrI) were estimated. Through use of the SUCRA values, the interven-

networks, there was a 50% or greater probability that the best-ranked treatment was actually not the best. No evidence showed a difference between the best-ranked intervention and the second and third best-ranked interventions in 90% and 71% of comparisons, respectively. In 39 networks with 6 or more interventions, the median probability that 1 of the top 2 interventions was among the bottom 2 was 35% (first to third quartile, 14% to 59%).

Limitation: This analysis did not consider such factors as the risk of bias within trials or small-study effects that may affect the reliability of rankings.

Conclusion: Treatment rankings derived from network meta-analyses have a substantial degree of imprecision. Authors and readers should interpret such rankings with great caution.

Primary Funding Source: Cochrane France.



OPEN ACCESS Freely available online

PLOS ONE

Evaluating the Quality of Evidence from a Network Meta-Analysis

Georgia Salanti[1], Cinzia Del Giovane[2], Anna Chaimani[1], Deborah M. Caldwell[3], Julian P. T. Higgins[3,4]*

1 Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, **2** Statistics Unit, Department of Clinical and Diagnostic Medicine and Public Health, University of Modena and Reggio Emilia, Modena, Italy, **3** School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, **4** Centre for Reviews and Dissemination, University of York, York, United Kingdom

Abstract

Systematic reviews that collate data about the relative effects of multiple interventions via network meta-analysis are highly informative for decision-making purposes. A network meta-analysis provides two types of findings for a specific outcome: the relative treatment effect for all pairwise comparisons, and a ranking of the treatments. It is important to consider the confidence with which these two types of results can enable clinicians, policy makers and patients to make informed decisions. We propose an approach to determining confidence in the output of a network meta-analysis. Our proposed approach is based on methodology developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group for pairwise meta-analyses. The suggested framework for evaluating a network meta-analysis acknowledges (i) the key role of indirect comparisons (ii) the contributions of each piece of direct evidence to the network meta-analysis estimates of effect size; (iii) the importance of the transitivity assumption to the validity of network meta-analysis; and (iv) the possibility of disagreement between direct evidence and indirect evidence. We apply our proposed strategy to a systematic review comparing topical antibiotics without steroids for chronically discharging ears with underlying eardrum perforations. The proposed framework can be used to determine confidence in the results from a network meta-analysis. Judgements about evidence from a network meta-analysis can be different from those made about evidence from pairwise meta-analyses.

Uncertainty of the entire treatment hierarchy

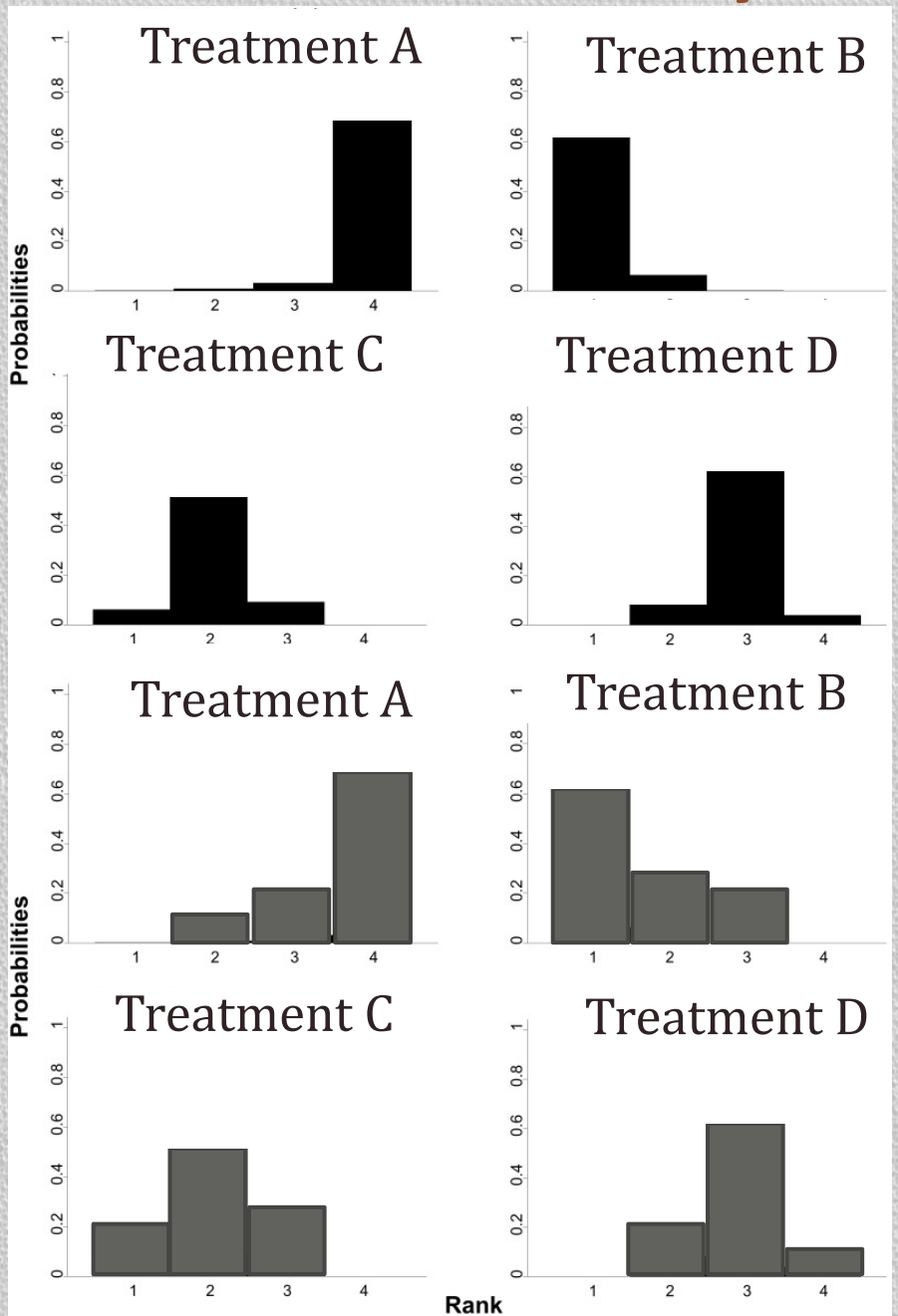
Preliminary suggestion:
look at the shape of rankograms

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

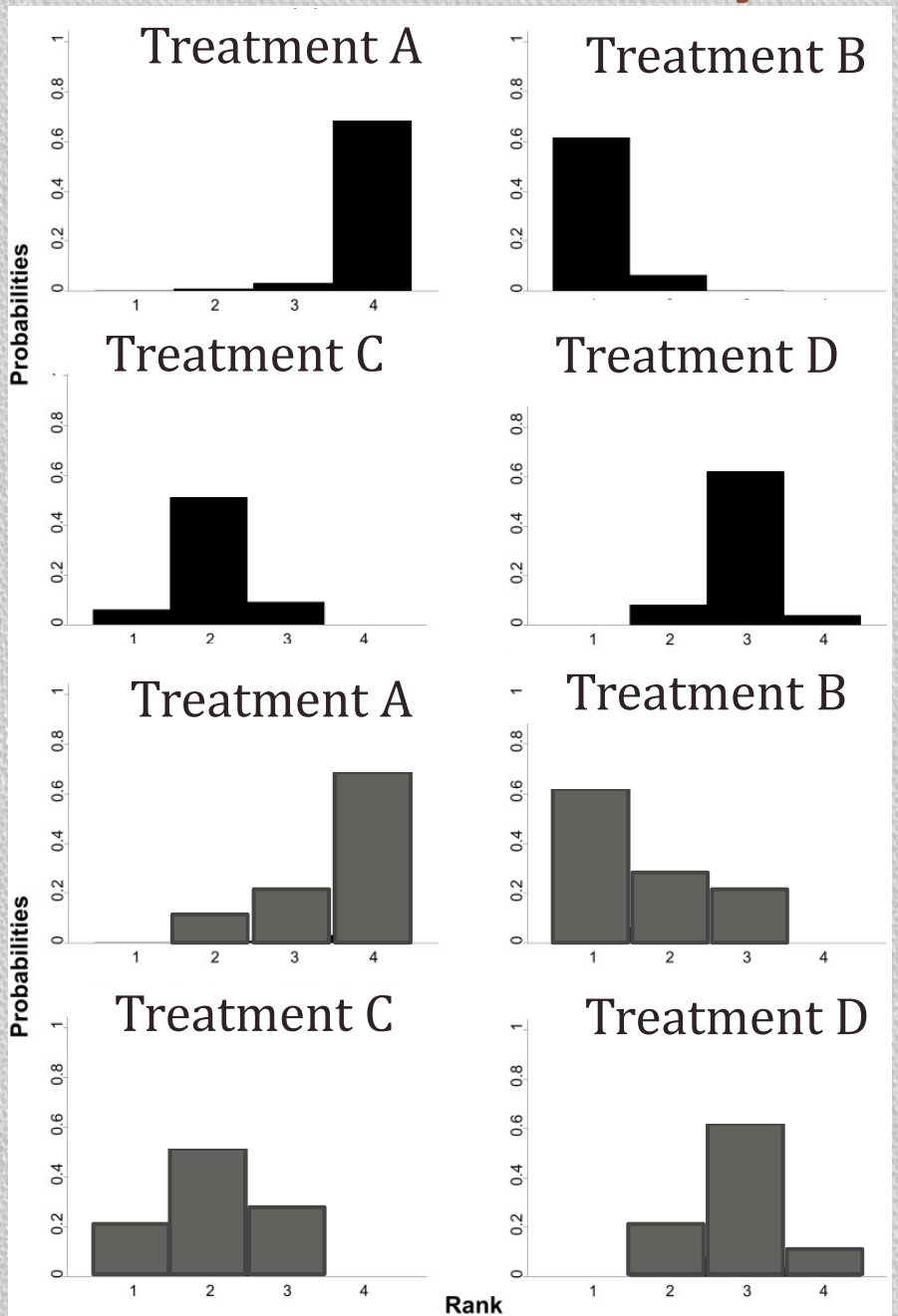


Uncertainty of the entire treatment hierarchy

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

$$P(A = 4 \cap B = 1 \cap C = 2 \cap D = 3)_{O1} \cap P(A = 4 \cap B = 1 \cap C = 2 \cap D = 3)_{O2}$$

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!