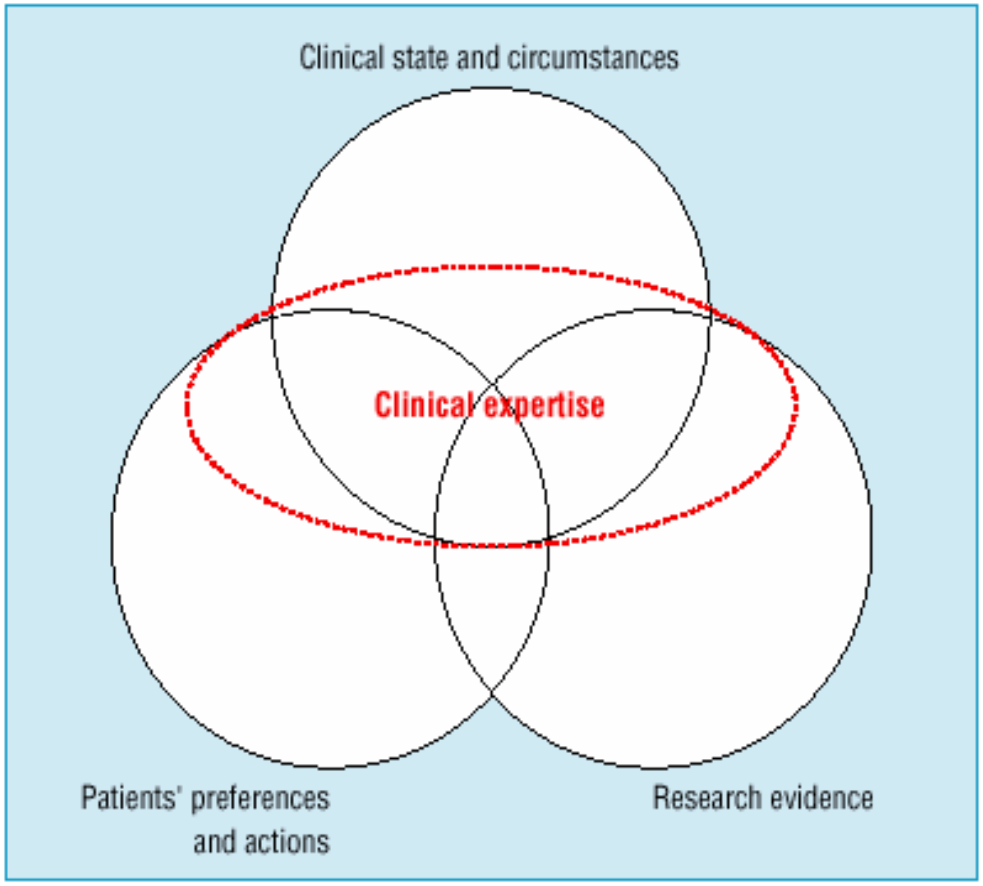
UNIVERSITY OF OXFORD

NHS National Institute for Health Research

Innovative methods of evidence synthesis: a primer on network meta-analysis

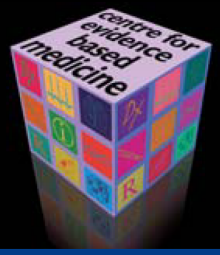
Andrea Cipriani

Programme in the Methods of Health Economics – Vienna, January 26th 2021



An updated model for evidence based clinical decisions[1]

BMJ 2002;324:1350



CEBM
CENTRE FOR EVIDENCE BASED MEDICINE

UNIVERSITY OF OXFORD

Levels of Evidence (March 2009)

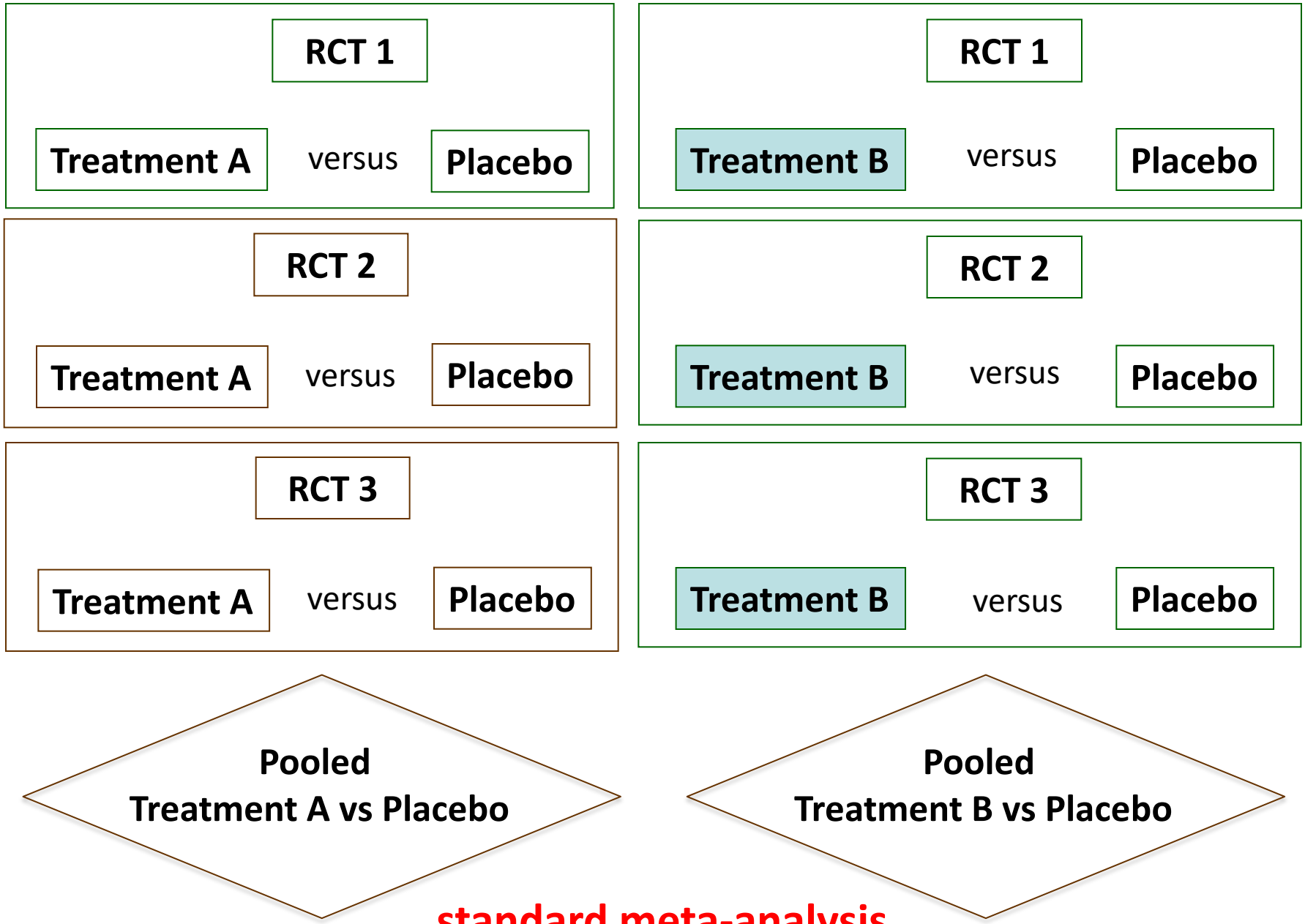
www.cebm.net

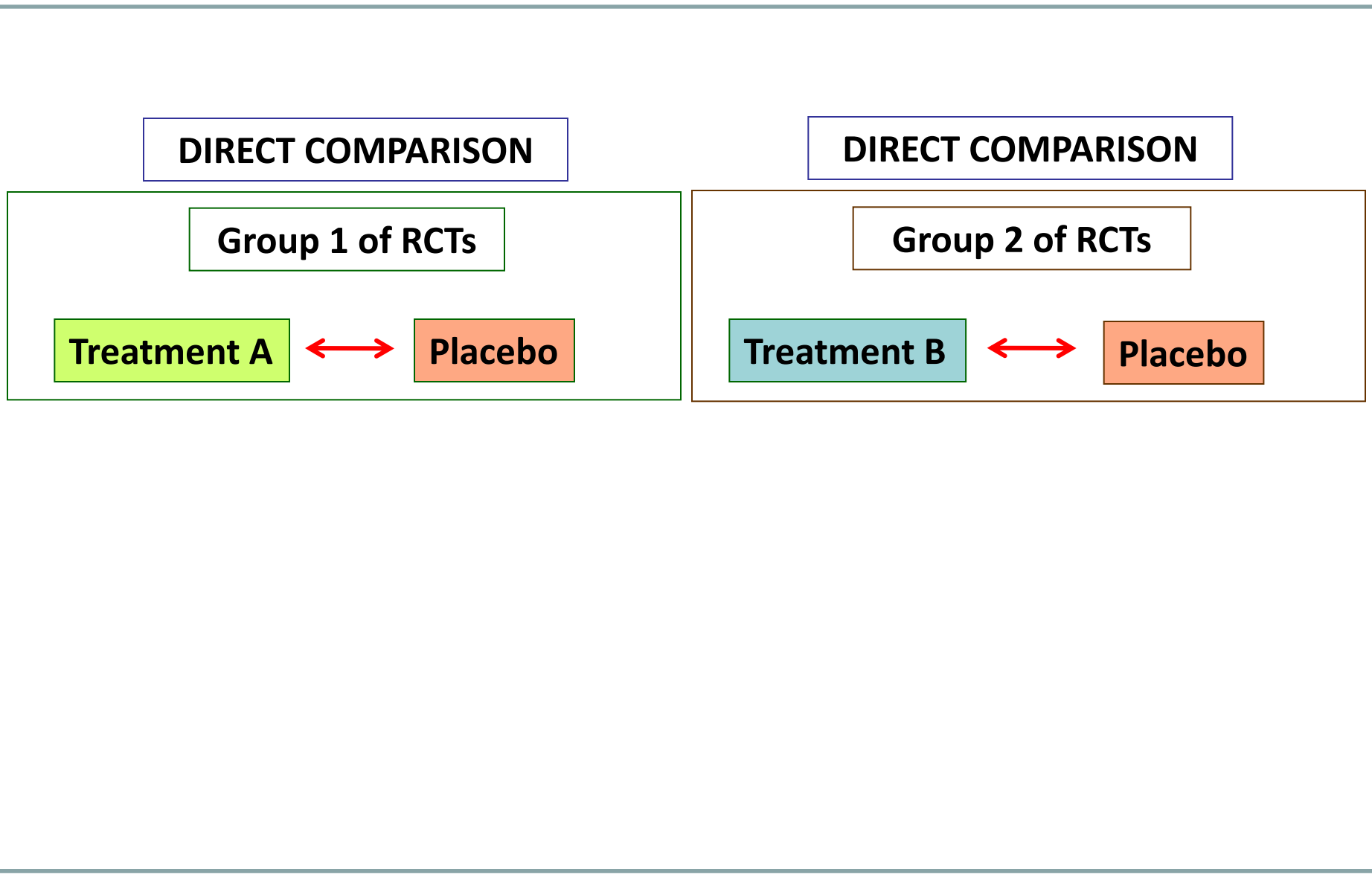
Level 1A	**Therapy/Prevention, Aetiology/Harm** **Prognosis** **Diagnosis** **Differential diag/symptom prevalence** **Economic and decision analyses**	1a SR (with homogeneity*)of RCTs SR (withhomogeneity*) of inception cohort studies; CDR† validated in different populations SR (with homogeneity*) of Level 1 diagnostic studies; CDR† with 1b studies from different clinical centres SR (with homogeneity*) of prospective cohort studies SR (with homogeneity*) of Level 1 economic studies
Level 1b	**Therapy/Prevention, Aetiology/Harm** **Prognosis** **Diagnosis** **Differential diag/symptom prevalence** **Economic and decision analyses**	Individual RCT (with narrow Confidence Interval‡) Individual inception cohort study with > 80% follow-up; CDR† validated in asingle population Validating** cohort study with good††† reference standards; or CDR† tested within one clinical centre Prospective cohort study with good follow-up**** Analysis based on clinically sensible costs or alternatives; systematic review(s) of the evidence; and including multi-way sensitivity analyses
Level 2a	**Therapy/Prevention, Aetiology/Harm** **Prognosis** **Diagnosis** **Differential diag/symptom prevalence** **Economic and decision analyses**	SR (with homogeneity*) of cohort studies SR (withhomogeneity*) of either retrospective cohort studies or untreated control groups in RCTs SR (with homogeneity*) of Level >2 diagnostic studies SR (with homogeneity*) of 2b and better studies SR (withhomogeneity*) of Level >2 economic studies
Level 2b	**Therapy/Prevention, Aetiology/Harm** **Prognosis** **Diagnosis** **Differential diag/symptom prevalence** **Economic and decision analyses**	Individual cohort study (including low quality RCT; e.g., <80% followup) Retrospective cohort study or follow-up of untreated control patients in an RCT; Derivation of CDR† or validated on split sample §§§ only Exploratory** cohort study with good††† reference standards; CDR† after derivation, or validated only on split-sample§§§ or databases Retrospective cohort study, or poor follow-up Analysis based on clinically sensible costs or alternatives; limited review(s) of the evidence, or single studies; and including multi-way sensitivity analyses
Level 3a	**Therapy/Prevention, Aetiology/Harm** **Prognosis** **Diagnosis** **Differential diag/symptom prevalence** **Economic and decision analyses**	SR (with homogeneity*) of case-control studies SR (with homogeneity*) of 3b and better studies SR (with homogeneity*) of 3b and better studies SR (with homogeneity*) of 3b And better studies

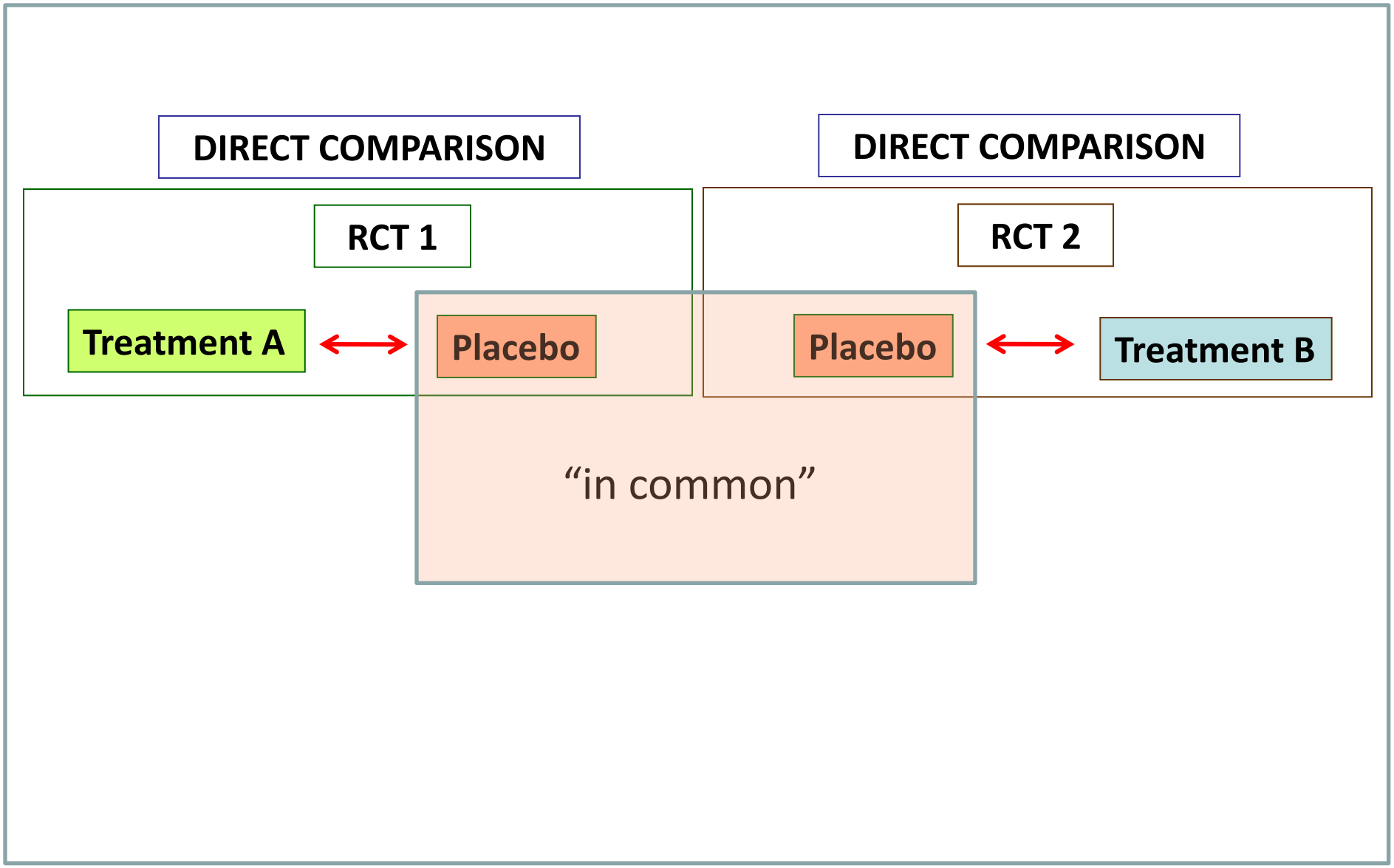
Results:

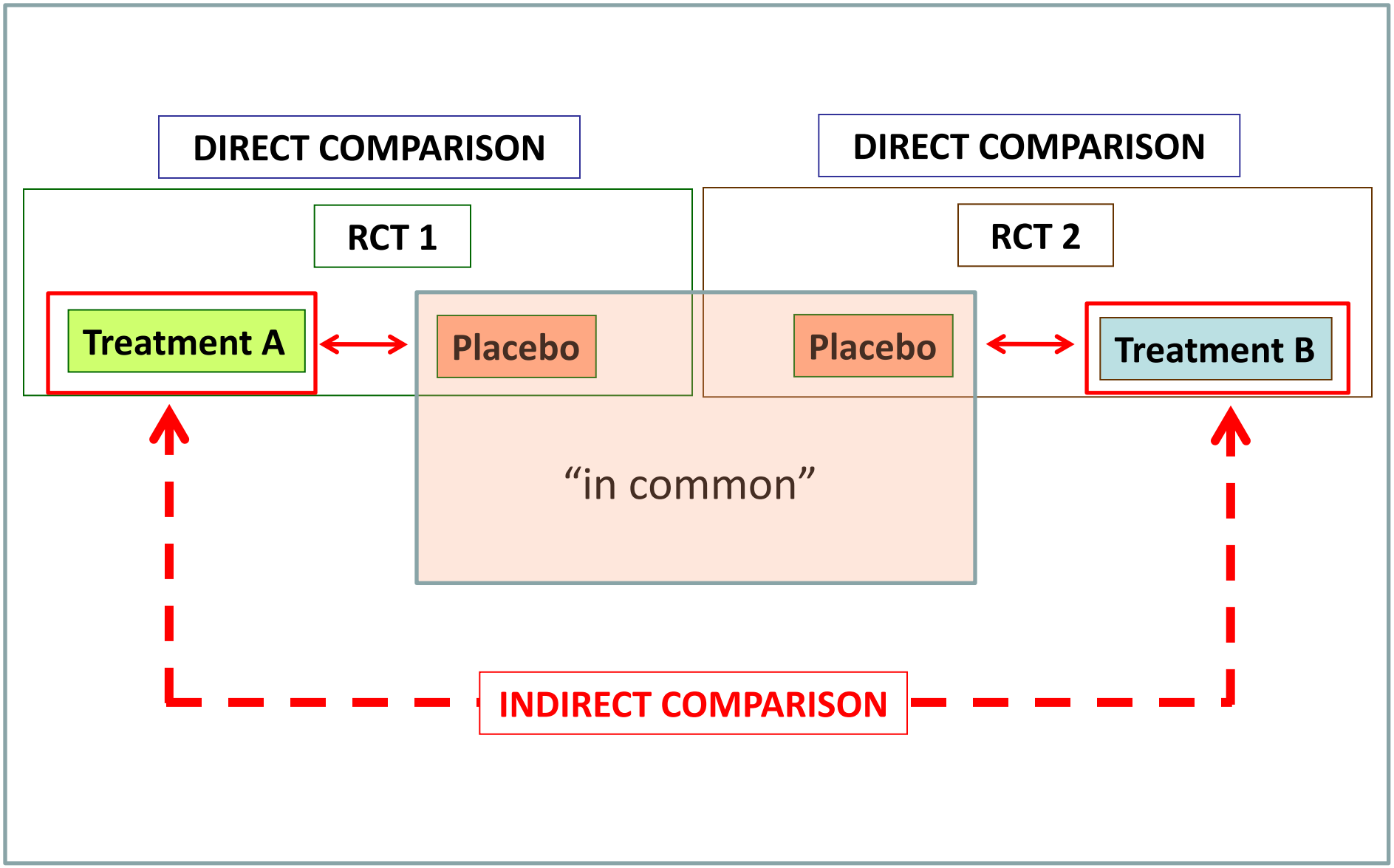
On measures of wax clearance Cerumol, sodium bicarbonate, olive oil and water are all more effective than no treatment; triethanolamine polypeptide (TP) is better than olive oil; wet irrigation is better than dry irrigation; sodium bicarbonate drops followed by irrigation by nurse is more effective than sodium bicarbonate drops followed by self-irrigation; softening with TP and self-irrigation is more effective than self-irrigation only; and endoscopic de-waxing is better than microscopic de-waxing.

Ideally, clinicians would like to know how all the different options rank against each other and how big the differences are between all the available options.

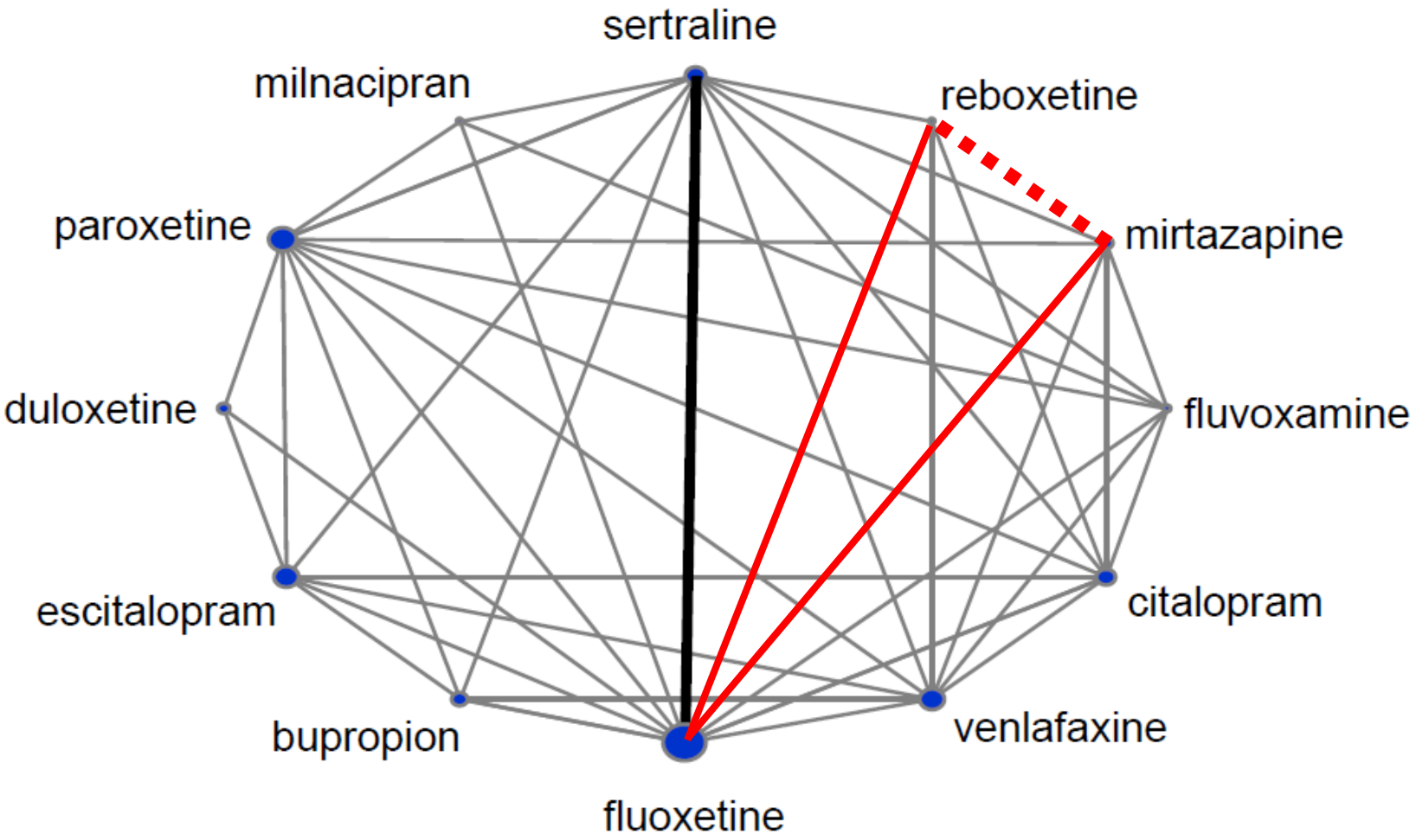

RCT 1
Treatment A
versus
Placebo
RCT 2
Treatment A
versus
Placebo
RCT 3
Treatment A
versus
Placebo
RCT 1
Treatment B
versus
Placebo
RCT 2
Treatment B
versus
Placebo
RCT 3
Treatment B
versus
Placebo
Pooled
Treatment A vs Placebo
Pooled
Treatment B vs Placebo
standard meta-analysis


DIRECT COMPARISON
Group 1 of RCTs
Treatment A
Placebo
DIRECT COMPARISON
Group 2 of RCTs
Treatment B
Placebo

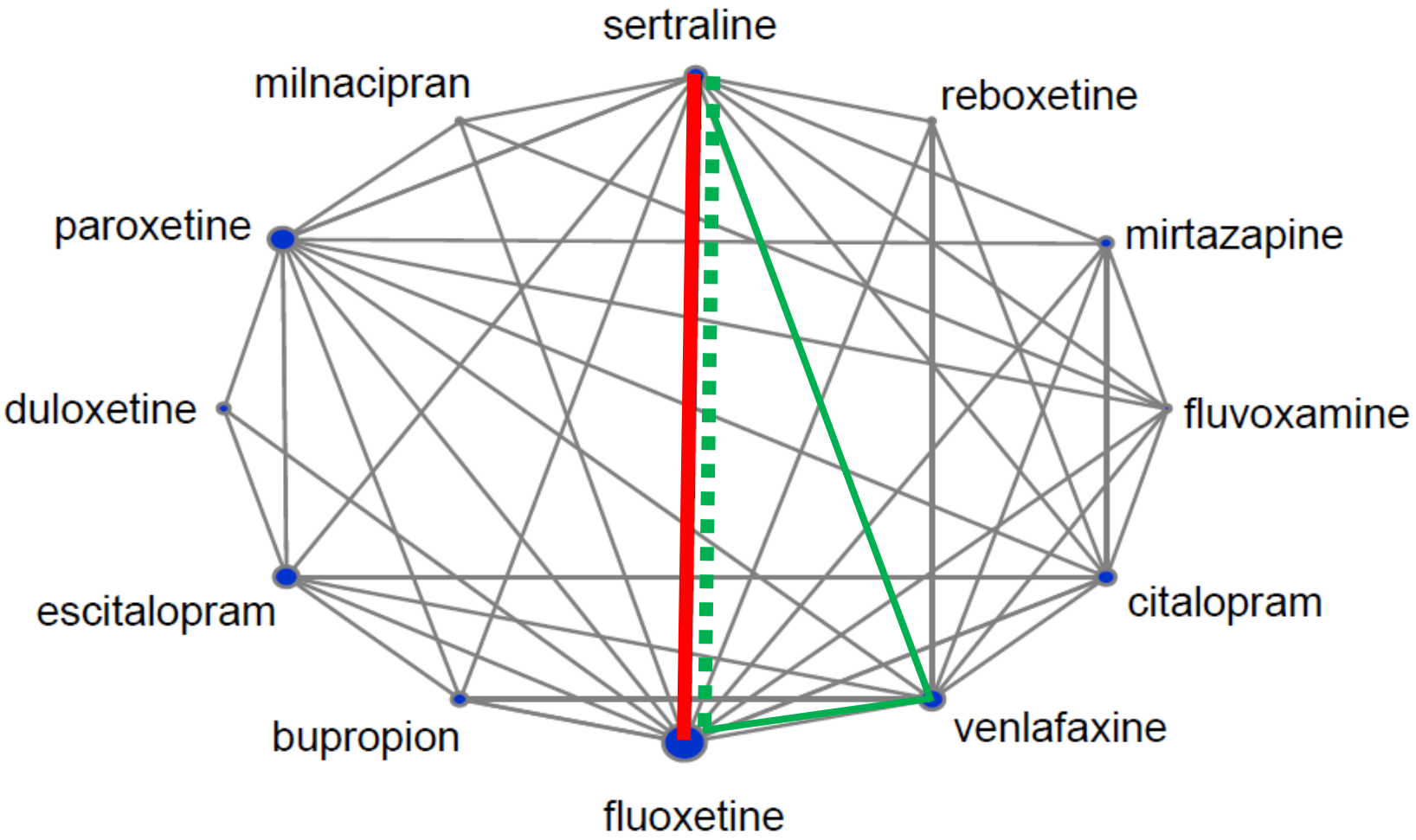

DIRECT COMPARISON
DIRECT COMPARISON
RCT 1
RCT 2
Treatment A
Placebo
Placebo
Treatment B
“in common”


DIRECT COMPARISON
DIRECT COMPARISON
RCT 1
RCT 2
Treatment A
Placebo
Placebo
Treatment B
"in common"
INDIRECT COMPARISON

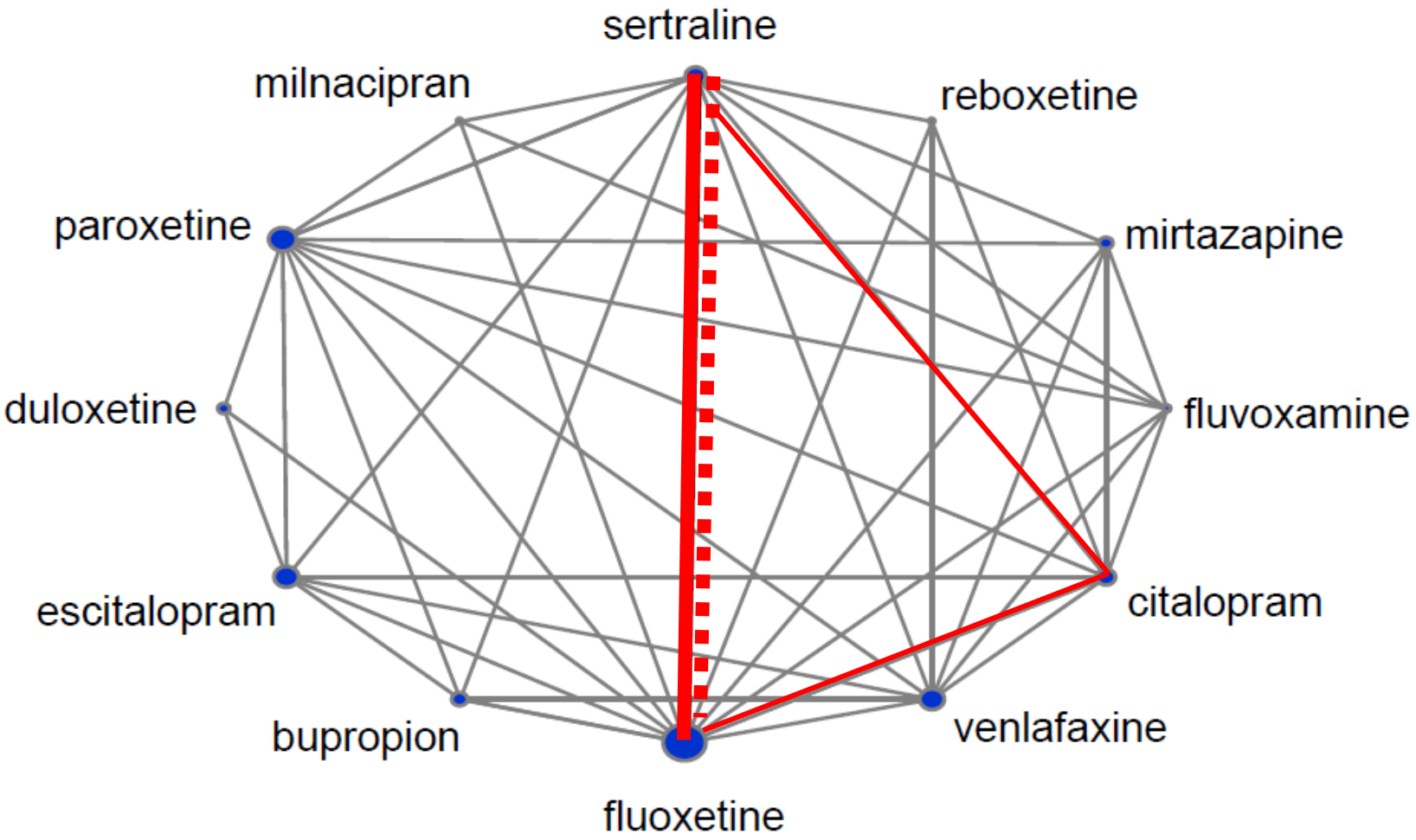
Network of experimental comparisons



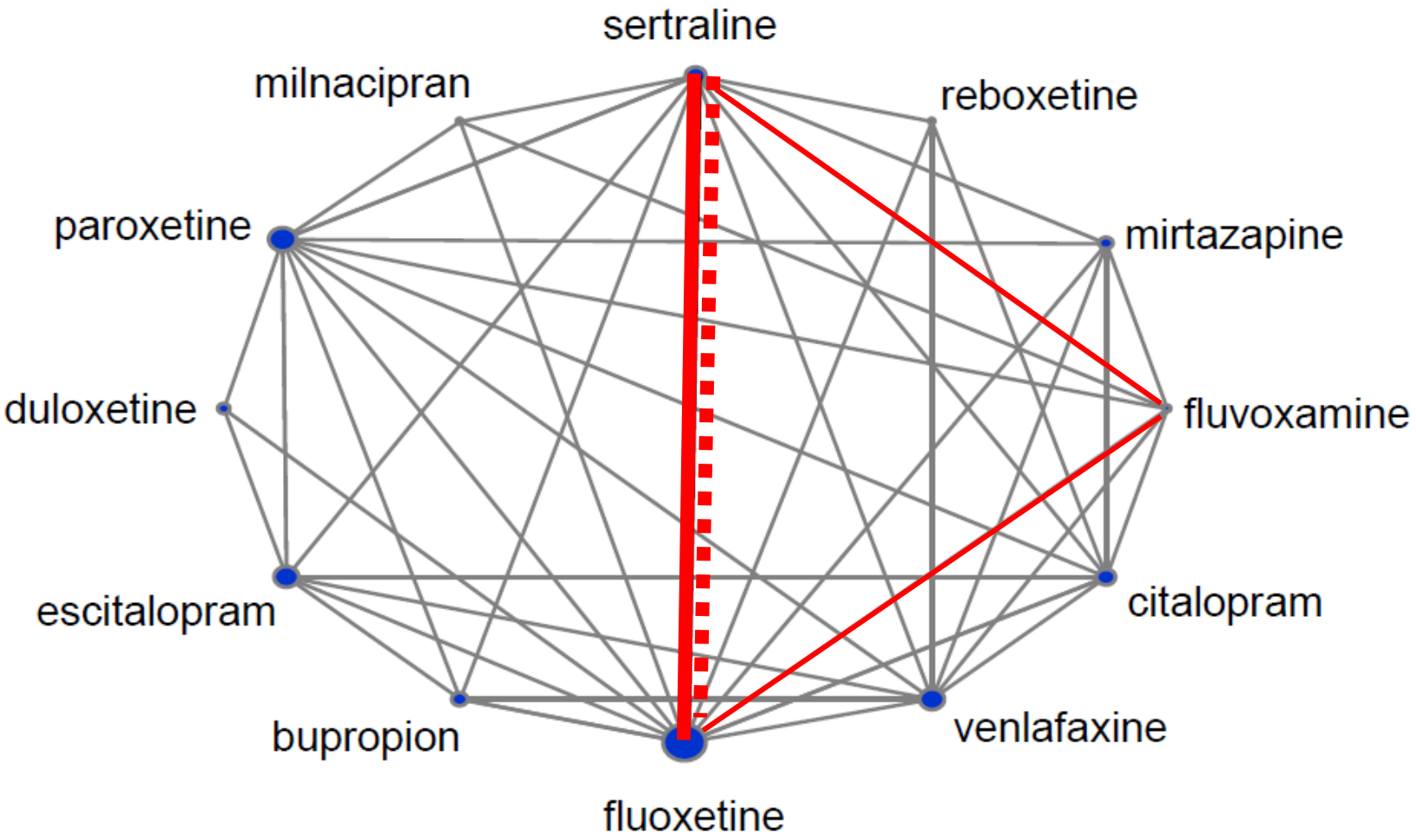
Network of experimental comparisons



Network of experimental comparisons



Network of experimental comparisons



Advantages of NMA

- Comprehensive use of all available data (direct evidence + indirect evidence)
- Comparison of interventions which haven't been directly compared in any trial
- Improved precision for each comparison
- Ranking of many treatments for the same condition

Ranking measures from MTM

- Estimate for each treatment the probability to be the best

% 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	0	0.75

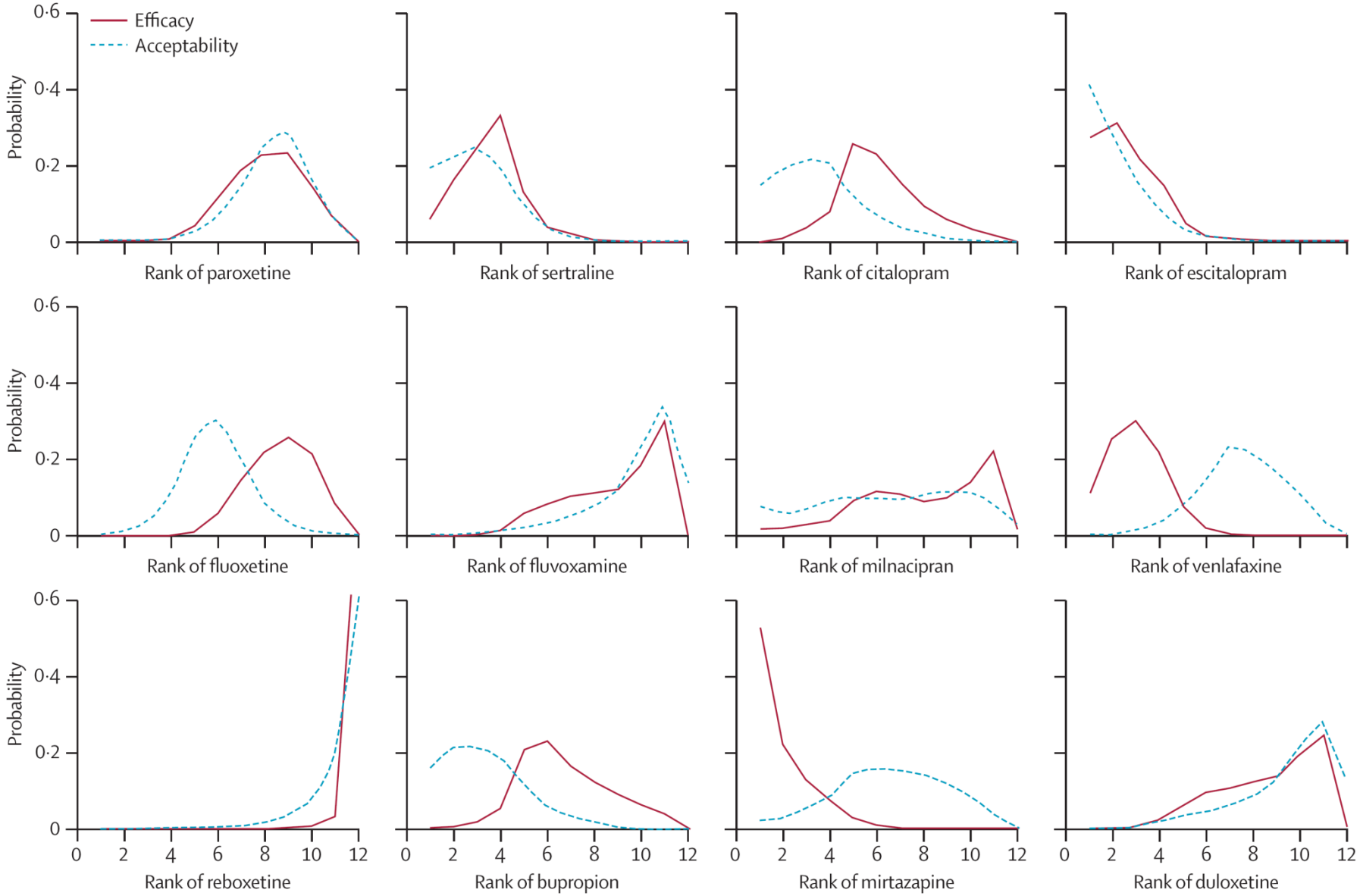


Figure 4: **Ranking for efficacy (solid line) and acceptability (dotted line)**
Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the 12 antidepressants.

- *Heterogeneity:* ‘excessive’ discrepancy among study-specific effects

- *Inconsistency:* it is the excessive discrepancy among source-specific effects (direct and indirect)

Heterogeneity?

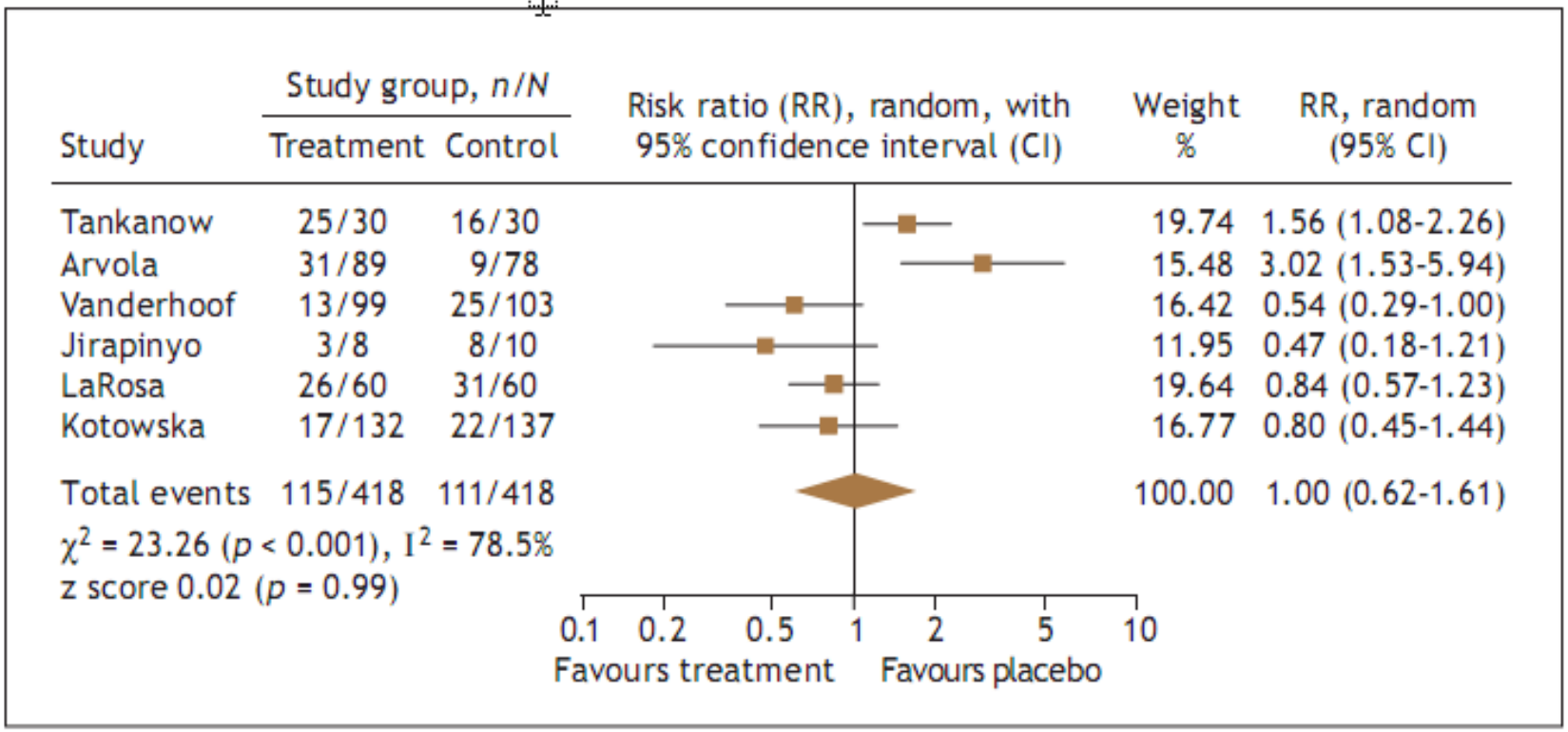
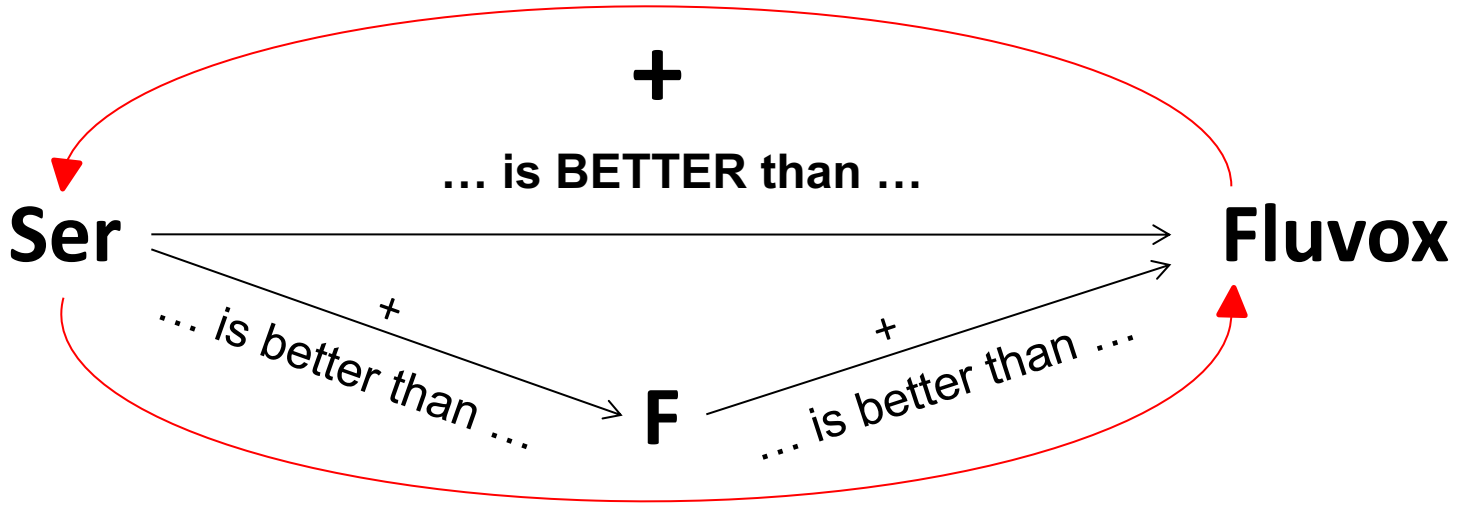
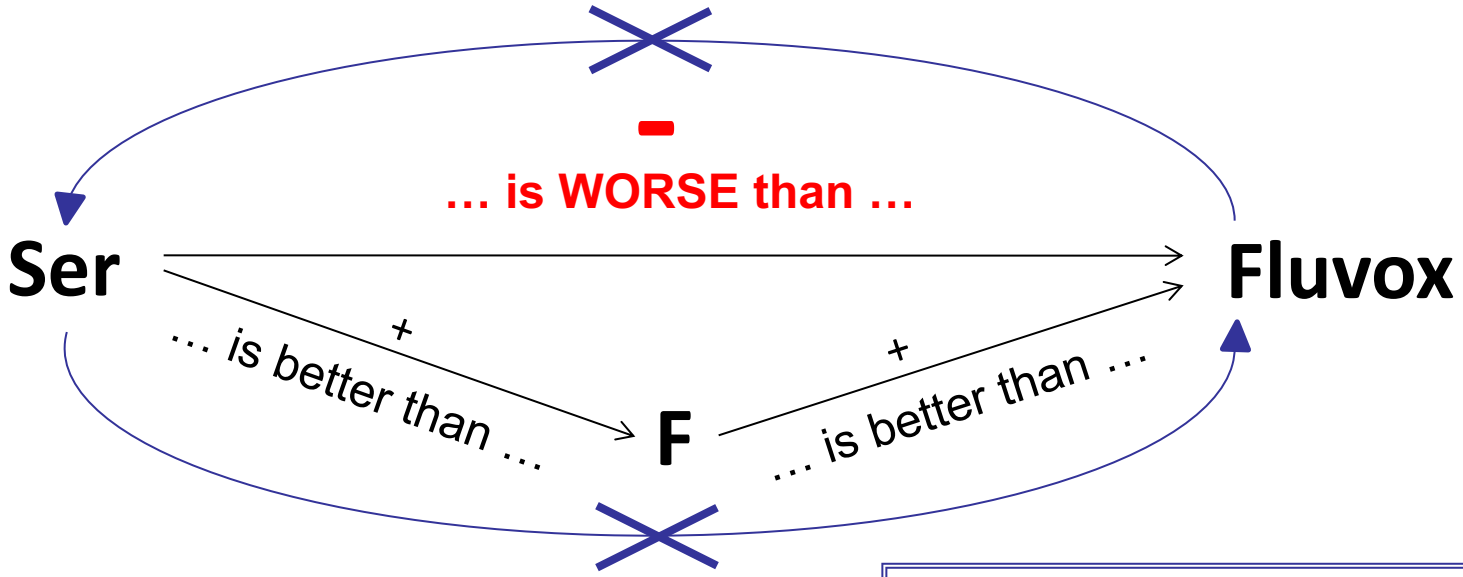


Fig 3: Incidence of antibiotic-associated diarrhea — intention-to-treat analysis. The analysis showed a nonsignificant difference between probiotics and placebo (z score) and statistically significant heterogeneity.

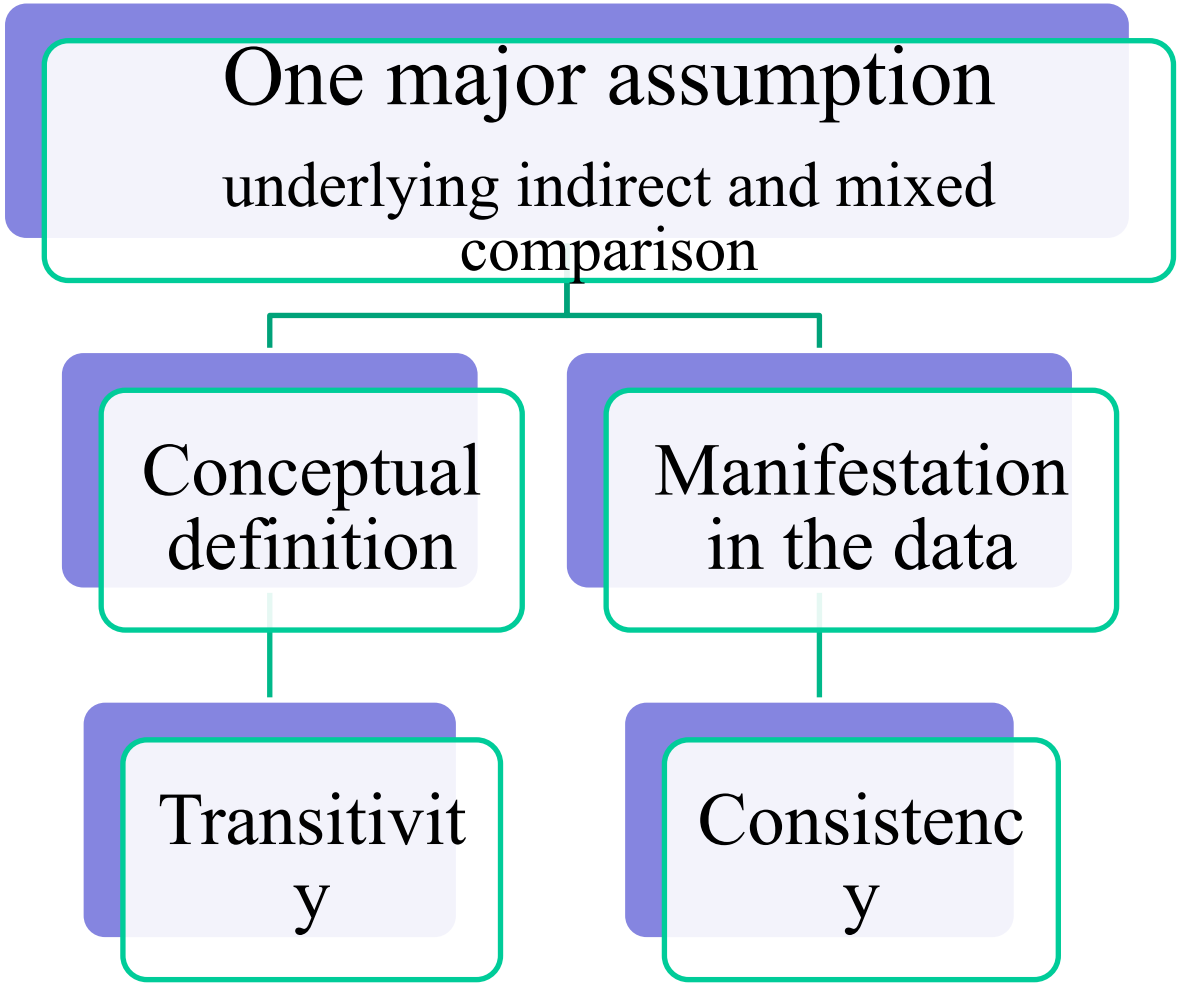


THE LOOP IS CONSISTENT



THE LOOP IS *INCONSISTENT*

Assumption underlying indirect/mixed comparison



Transitivity

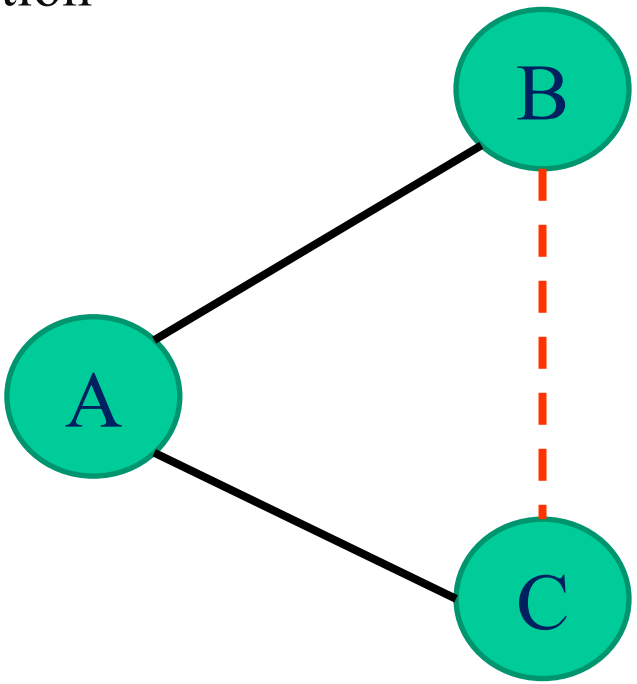
The underlying assumption when μ^{I}_{BC} is calculated is that **we can learn about B versus C via A**.

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

advantage of C over B =
advantage of C over A - advantage of B over A

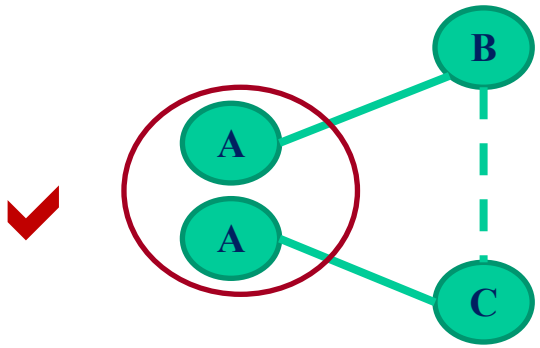
Transitivity

Sometime it is an untestable assumption

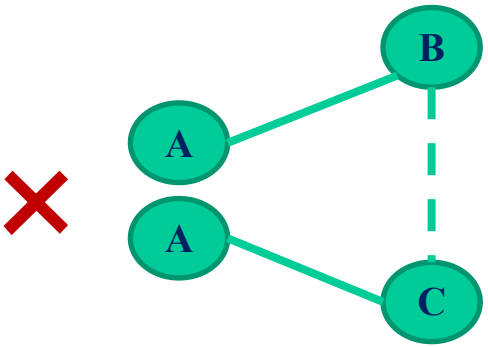


….but we can evaluate clinically and epidemiologically its plausibility

Ways of looking at transitivity... (1)



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

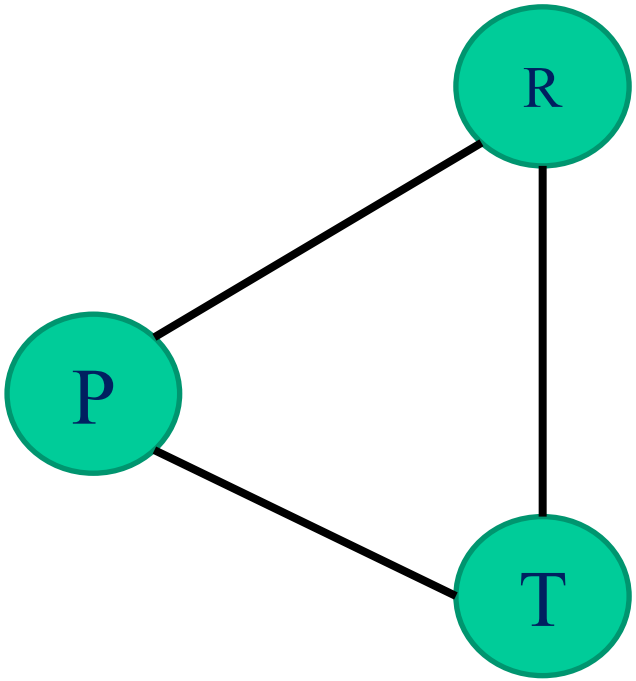


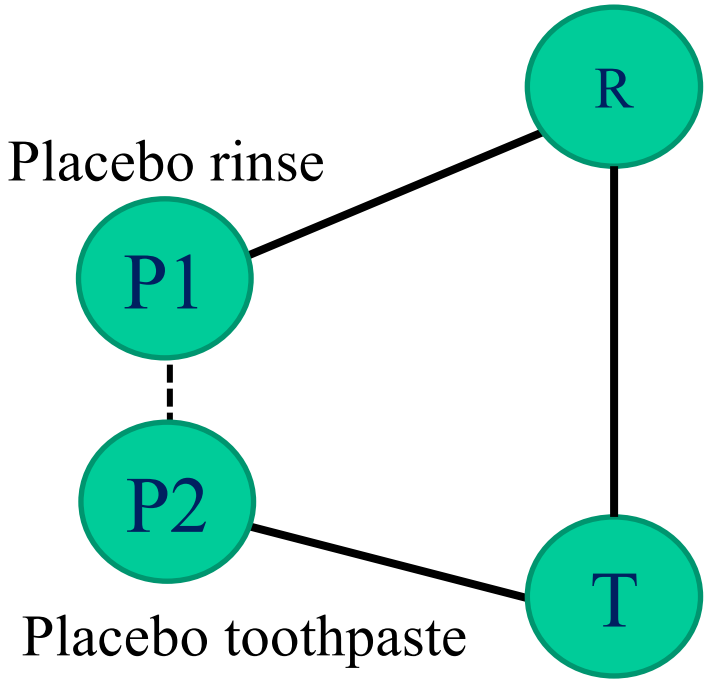
Is it plausible when A is placebo given in different forms (e.g. injection versus pill/different doses/different durations etc.)?

Ways of looking at transitivity... (1)

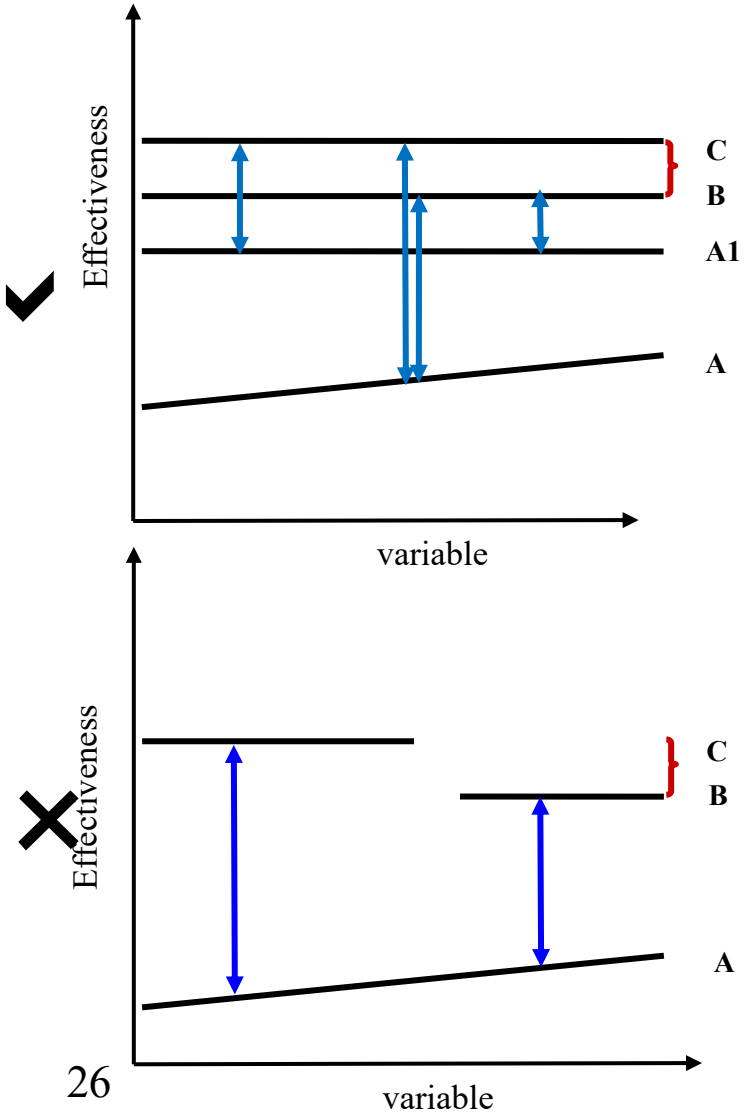
- Note that transitivity is violated when the anchor treatment differs systematically between trials (not randomly)
- Random differences may lead to excess heterogeneity
- But systematic differences correspond to intransitivity
- Consequently, the definition of the nodes in the treatment network is a **challenging issue** with important implications for the joint analysis
- *Eg. should we include a single placebo node to the network, or a placebo toothpaste and a placebo rinse?*

Should I split the placebo node?





Ways of looking at transitivity... (3)



Make sure that AC and AB trials do not differ with respect to the distribution of effect modifiers

But

- Difficult to defend when you have older and newer treatments
- It is not always possible to know whether a variable is a prognostic factor, an effect modifier or neither
- Variables are often unobserved

Ways of looking at transitivity... (3)

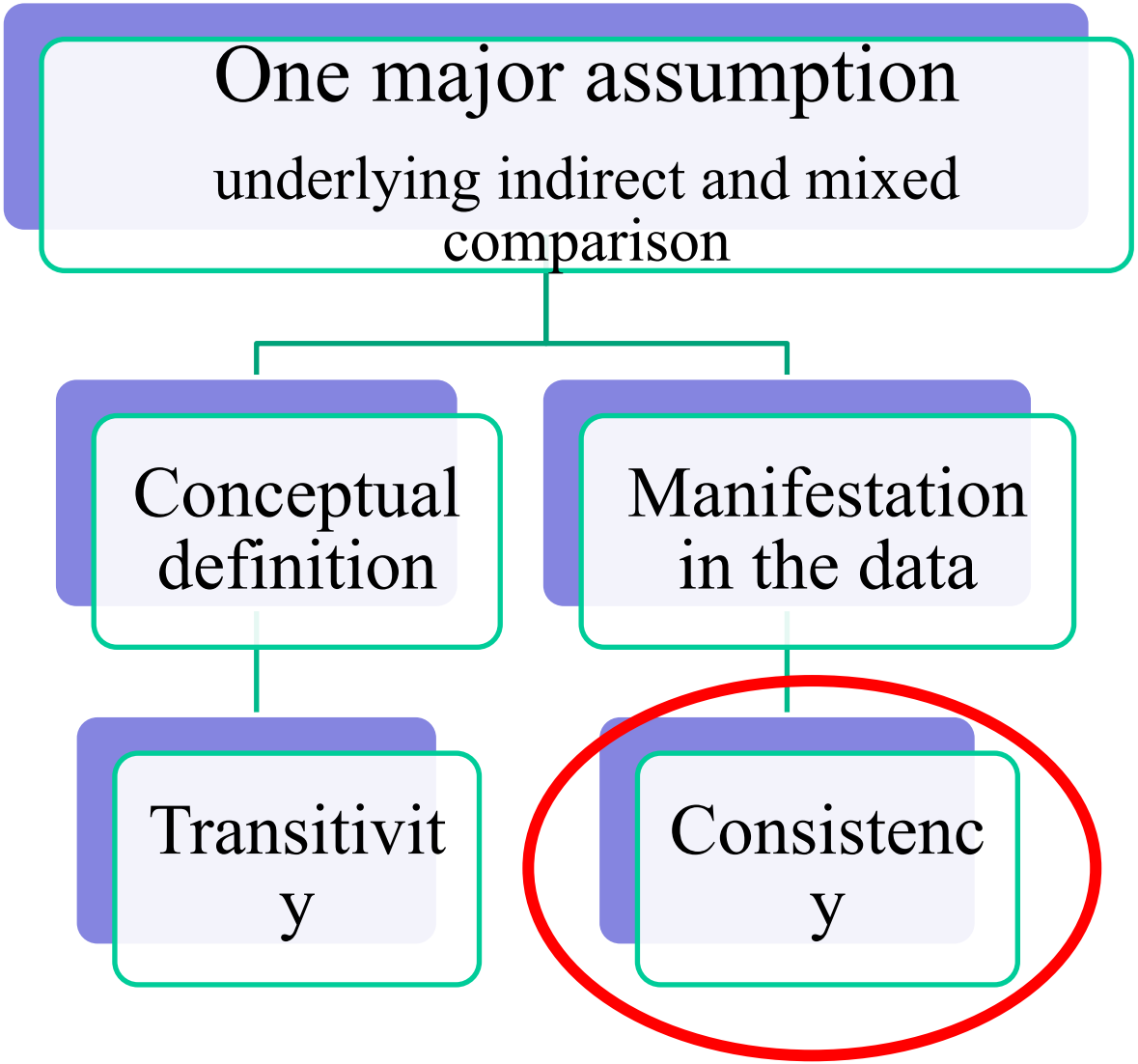
- This formulation facilitates evaluation of the transitivity assumption.
  - ✓ *Check distribution of effect modifiers of the relative treatment effects in AC and AB trials*
- Clinicians and methodologists that aim to synthesize evidence from many comparisons should identify **a priori** possible effect modifiers and compare their distributions across comparisons.

Ways of looking at transitivity... (4)

- All treatments are **"jointly randomizable"**
- One can think of a **mega-trial** including all treatments
- This consideration is a fundamental one and should be addressed when building the evidence network
- The assumption of transitivity could be violated if interventions have different indications.
  - *Ex: treatment A is a chemotherapy regimen administered as a second line treatment, whereas treatments B and C can be either as first or second line*
  - *we cannot assume that participants in a BC trial could have been randomized in an AC trial!*

Transitivity – key points

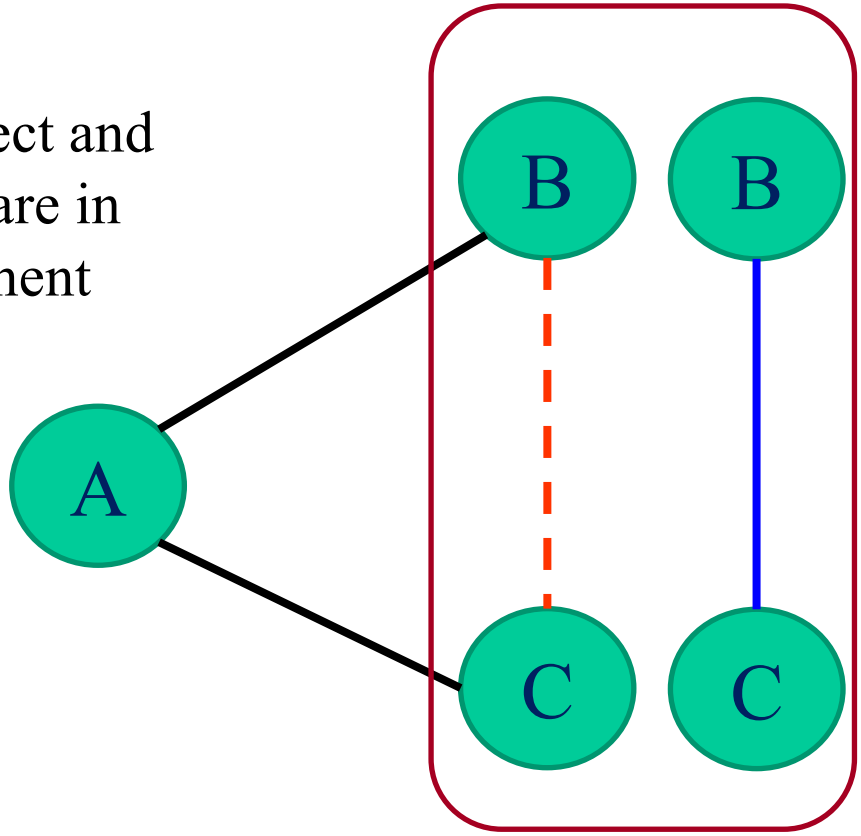
- Each treatment in the network pertains to a '**fixed' definition** independently of its comparator.
- **A mega-trial** could be performed
- All patients in the identified studies could in principle receive **any treatment**
- All sets of trials grouped by comparison are similar with respect to the **distribution of effect modifiers**


One major assumption
underlying indirect and mixed comparison
Conceptual definition
Manifestation in the data
Transitivity
Consistency

Checking for consistency provides a way for checking the transitivity assumption!

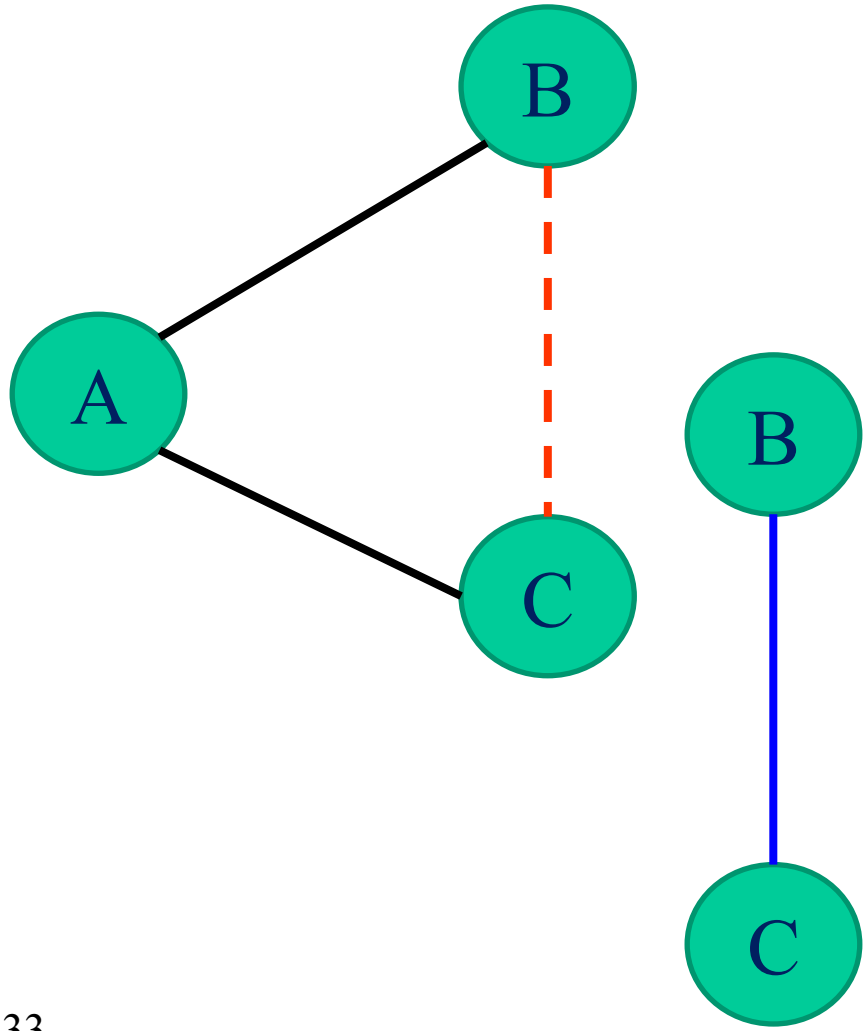
Consistency

Consistency = direct and indirect evidence are in (statistical) agreement



$$\mu^{D}_{AC} - \mu^{D}_{AB} = \mu^{I}_{BC} = \mu^{D}_{BC}$$

Inconsistency



Inonsistency = direct and indirect evidence are not in (statistical) agreement

Statistical consistency

- Consistency is a property of a '**closed loop**' (a path that starts and ends at the same node) or 'cycle' (as in graph theory)
- By definition, there can be no (in)consistency in **open** loops

- Global and local inconsistency

What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Check the data	Studies that 'stand out' in the forest plot are checked for data extraction errors	Using simple loop inconsistency you can identify studies with data extraction errors. Inconsistency in loops where a comparison is informed by a single study is particularly suspicious for data errors.
Try to bypass	There is empirical evidence that some measures are associated with larger heterogeneity than others (Deeks 2002; Friedrich et al. 2011)	Empirical evidence suggests that different effect measures of dichotomous outcomes does not impact on statistical inconsistency (Veroniki et al. 2013)

What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Resign to it	Investigators may decide not to undertake meta-analysis in the presence of excessive heterogeneity	Investigators may decide not to synthesize the network in the presence of excessive inconsistency
Encompass it	Apply random-effects meta-analysis	Apply models that relax the consistency assumption by adding an 'extra' loop-specific random effect (Higgins et al. 2012, Lu & Ades 2006)*.

*However, as random effects are not a remedy for excessive heterogeneity and should be applied only for unexplained heterogeneity, inconsistency models should be employed to reflect inconsistency in the results, not *to adjust* for it.

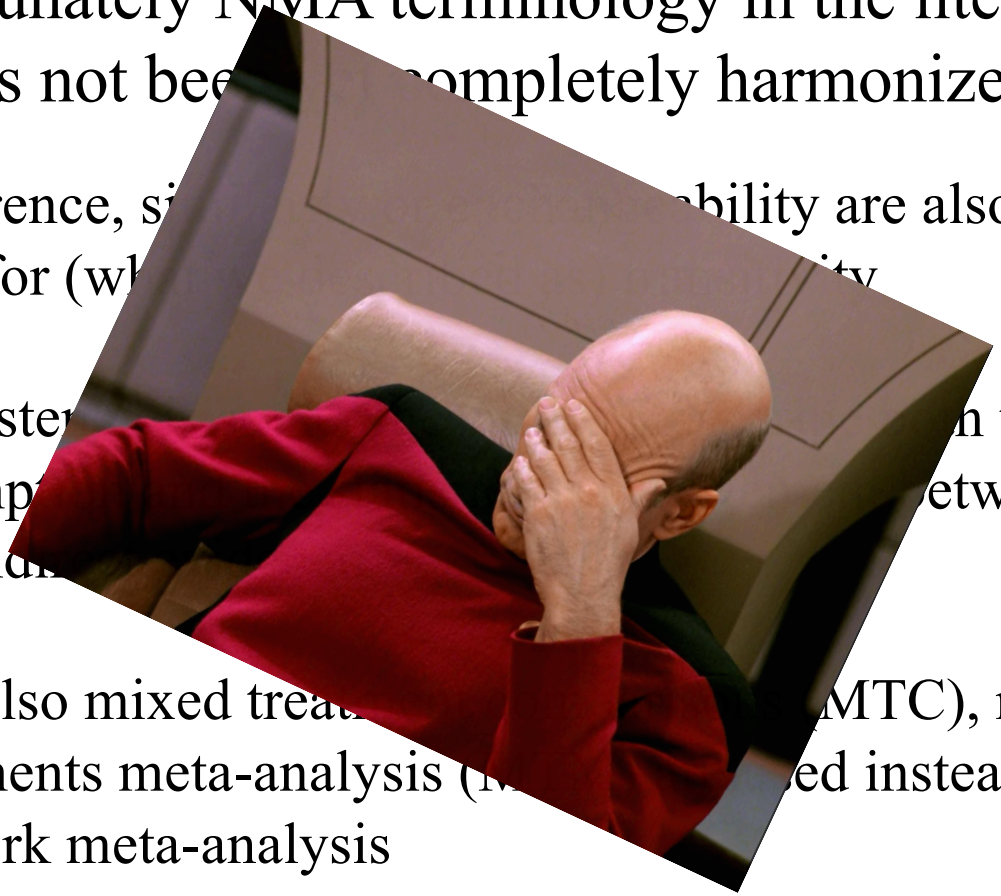
What to do when statistically significant inconsistency is found?

Action	Heterogeneity	Inconsistency
Explore it	Use pre-specified variables in a subgroup analysis or meta-regression	Split the network into subgroups or use network meta-regression to account for differences across studies and comparisons. Specify the variables in the protocol, including bias-related characteristics.

Beware of difference in terminology

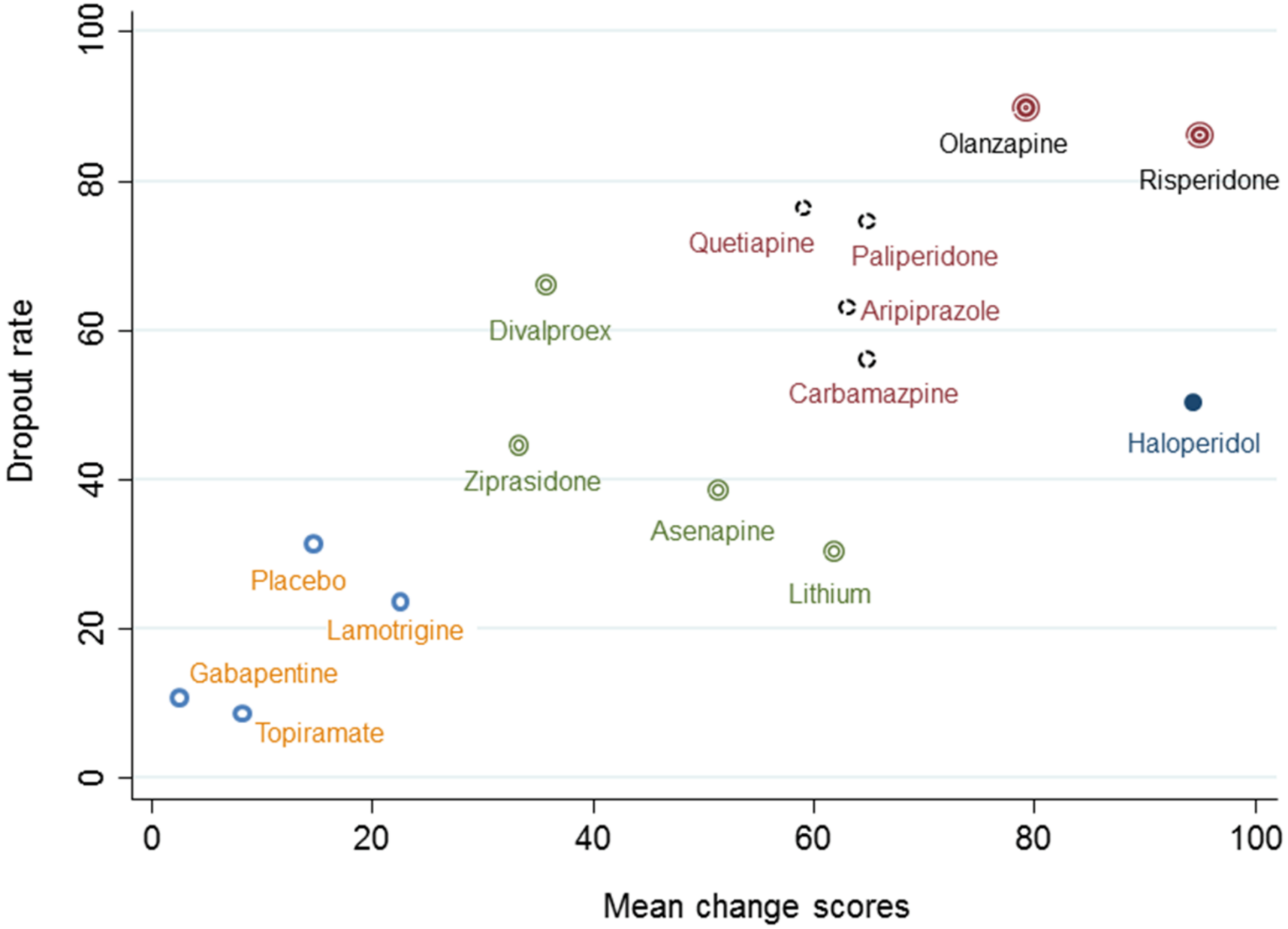
Unfortunately NMA terminology in the literature has not bee ompletely harmonized

- Coherence, si bility are also used as a term for (wh ity
- Consister n the assump etween direct and indi
- And also mixed trea MTC), multiple treatments meta-analysis (ed instead of network meta-analysis



Summary

- Transitivity refers to the validity of the indirect comparison and can be evaluated conceptually. It is a **key assumption underlying NMA**
- Statistical evaluation of the consistency can take place in a closed loop
- Care is needed when interpreting the results of a consistency test as issues of heterogeneity and power may limit its usefulness
- Conceptual evaluation of the transitivity assumption should include
  - ✓ Checking for effect modifiers that differ across comparisons
  - ✓ Checking the definition of each node/treatment
  - ✓ The concept of a mega-trial
  - ✓ Each patient can in principle receive every treatment in the network



Mavridis et al., *Evid Based Mental Health* 2015;18:40-46

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*

Summary

Background Major depressive disorder is one of the most common, burdensome, and costly psychiatric disorders worldwide in adults. Pharmacological and non-pharmacological treatments are available; however, because of inadequate resources, antidepressants are used more frequently than psychological interventions. Prescription of these agents should be informed by the best available evidence. Therefore, we aimed to update and expand our previous work to compare and rank antidepressants for the acute treatment of adults with unipolar major depressive disorder.

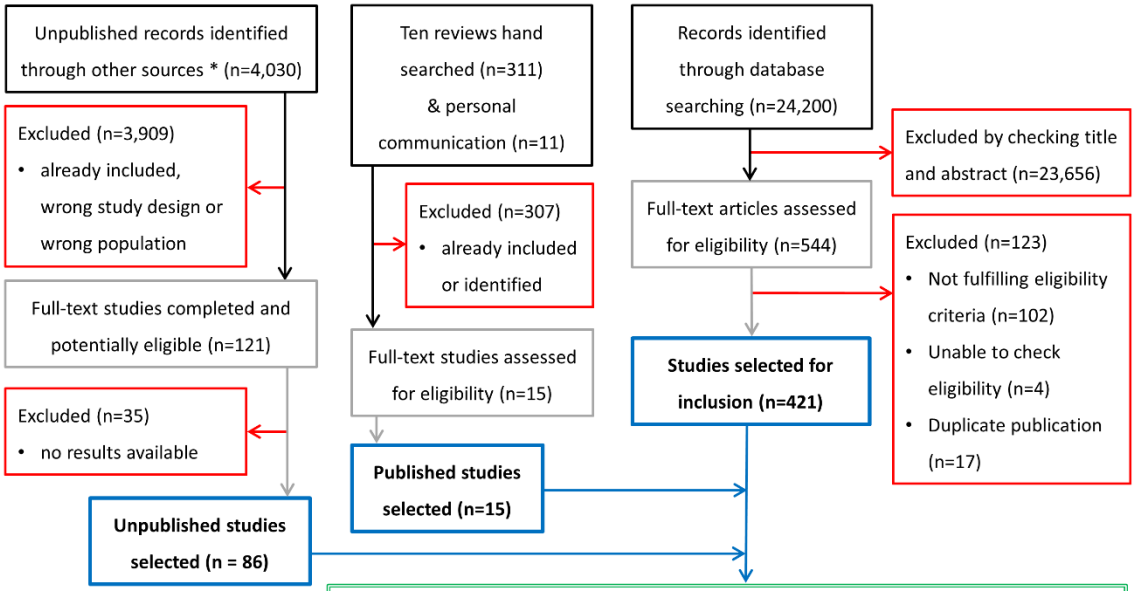
Methods We did a systematic review and network meta-analysis. We searched Cochrane Central Register of Controlled Trials, CINAHL, Embase, LILACS database, MEDLINE, MEDLINE In-Process, PsycINFO, the websites of regulatory agencies, and international registers for published and unpublished, double-blind, randomised controlled trials from their inception to Jan 8, 2016. We included placebo-controlled and head-to-head trials of 21 antidepressants used for the acute treatment of adults (≥18 years old and of both sexes) with major depressive disorder diagnosed according to standard operationalised criteria. We excluded quasi-randomised trials and trials that were incomplete or included 20% or more of participants with bipolar disorder, psychotic depression, or treatment-resistant depression; or patients with a serious concomitant medical illness. We extracted data following a predefined hierarchy. In network meta-analysis, we used group-level data. We assessed the studies' risk of bias in accordance to the Cochrane Handbook for Systematic Reviews of Interventions, and certainty of evidence using the Grading of Recommendations Assessment, Development and Evaluation framework. Primary outcomes were efficacy (response rate) and acceptability (treatment discontinuations due to any cause). We estimated summary

Published **Online** February 21, 2018 http://dx.doi.org/10.1016/S0140-6736(17)32802-7

See **Online/Comment** http://dx.doi.org/10.1016/S0140-6736(18)30421-5

*Joint first authors

Department of Psychiatry, University of Oxford, Oxford, UK (A Cipriani MD, L Z Atkinson MSc, H G Ruhe PhD, Prof J R Geddes MD); **Oxford Health NHS Foundation Trust** (A Cipriani, Prof J R Geddes) **and Oxford Centre for Human Brain Activity, Department of Psychiatry** (L Z Atkinson), **Warneford Hospital, Oxford, UK; Department of Health Promotion and Human**



Total number of DB RCTs included in the network meta-analysis (n=522, N=116,477)

- Agomelatine vs placebo or another active comparison (n = 23)
- Amitriptyline vs placebo or another active comparison (n = 96)
- Bupropion vs placebo or another active comparison (n = 33)
- Citalopram vs placebo or another active comparison (n = 38)
- Clomipramine vs placebo or another active comparison (n = 20)
- Desvenlafaxine vs placebo or another active comparison (n = 9)
- Duloxetine vs placebo or another active comparison (n = 30)
- Escitalopram vs placebo or another active comparison (n = 42)
- Fluoxetine vs placebo or another active comparison (n = 117)
- Fluvoxamine vs placebo or another active comparison (n = 32)
- Levomilnacipran vs placebo or another active comparison (n = 6)
- Milnacipran vs placebo or another active comparison (n = 10)
- Mirtazapine vs placebo or another active comparison (n = 34)
- Nefazodone vs placebo or another active comparison (n = 21)
- Paroxetine vs placebo or another active comparison (n = 114)
- Reboxetine vs placebo or another active comparison (n = 17)
- Sertraline vs placebo or another active comparison (n = 54)
- Trazodone vs placebo or another active comparison (n = 26)
- Venlafaxine vs placebo or another active comparison (n = 68)
- Vilazodone vs placebo or another active comparison (n = 9)
- Vortioxetine vs placebo or another active comparison (n = 15)

Figure 1: Selection of included and excluded studies (with reasons). Black boxes present screened references; red boxes present excluded studies (with reasons); blue boxes present selected studies, and green boxes present studies included in the network meta-analysis. DB: double blind; RCTs: randomized controlled trials. * Industry websites, contact with authors and trial registries. Clinicaltrials.gov was searched by 'drug name' AND 'major depressive disorder' as the major heading. The total number of unpublished records is the total number of results doing this for each drug and on each unpublished database source. The main reasons for exclusion included open label/single blind studies, studies including patients with comorbid disorders and combination therapy trials. Searches were only conducted on completed trials, which also removed many ongoing/terminated results, especially from clinicaltrials.gov.

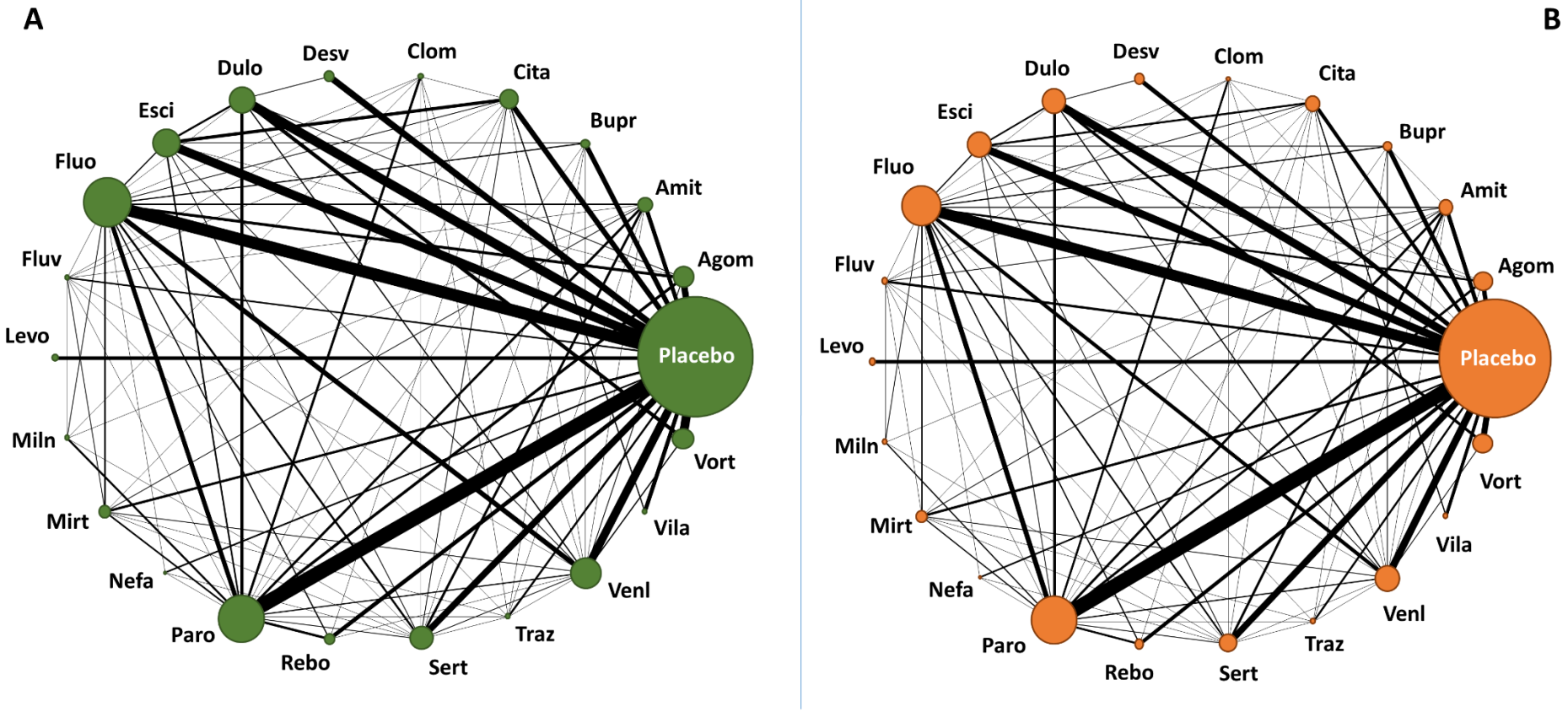
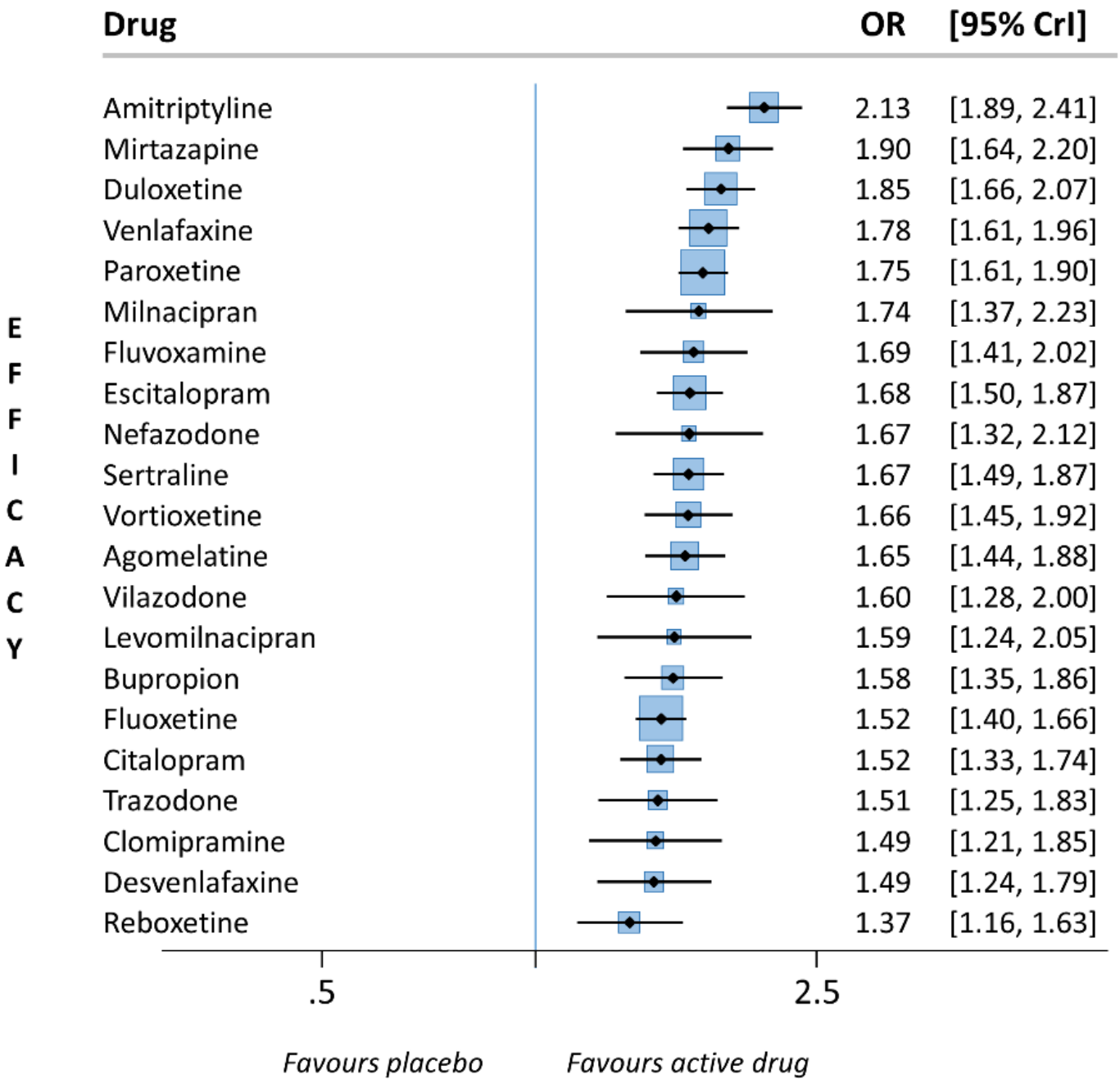
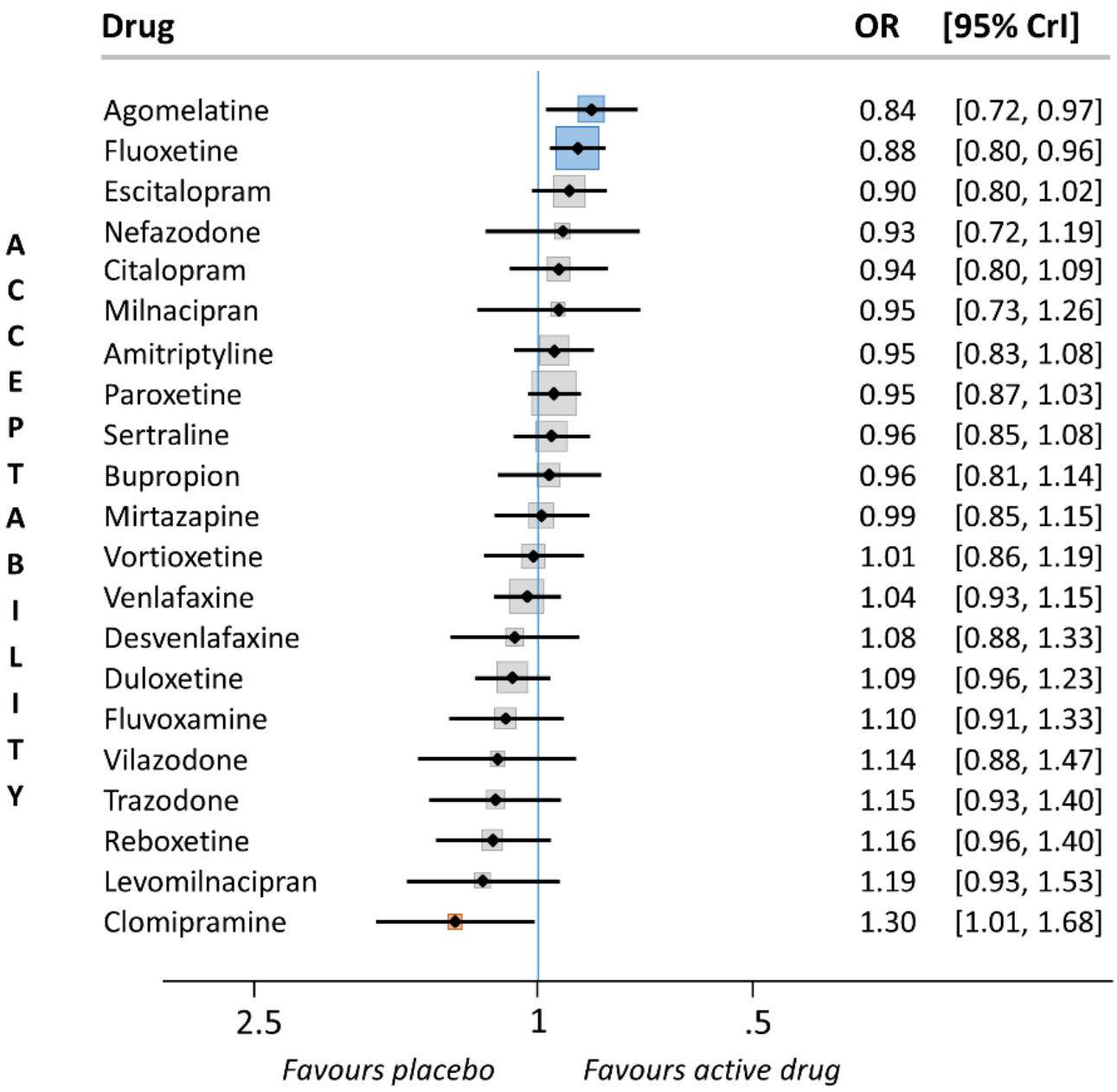
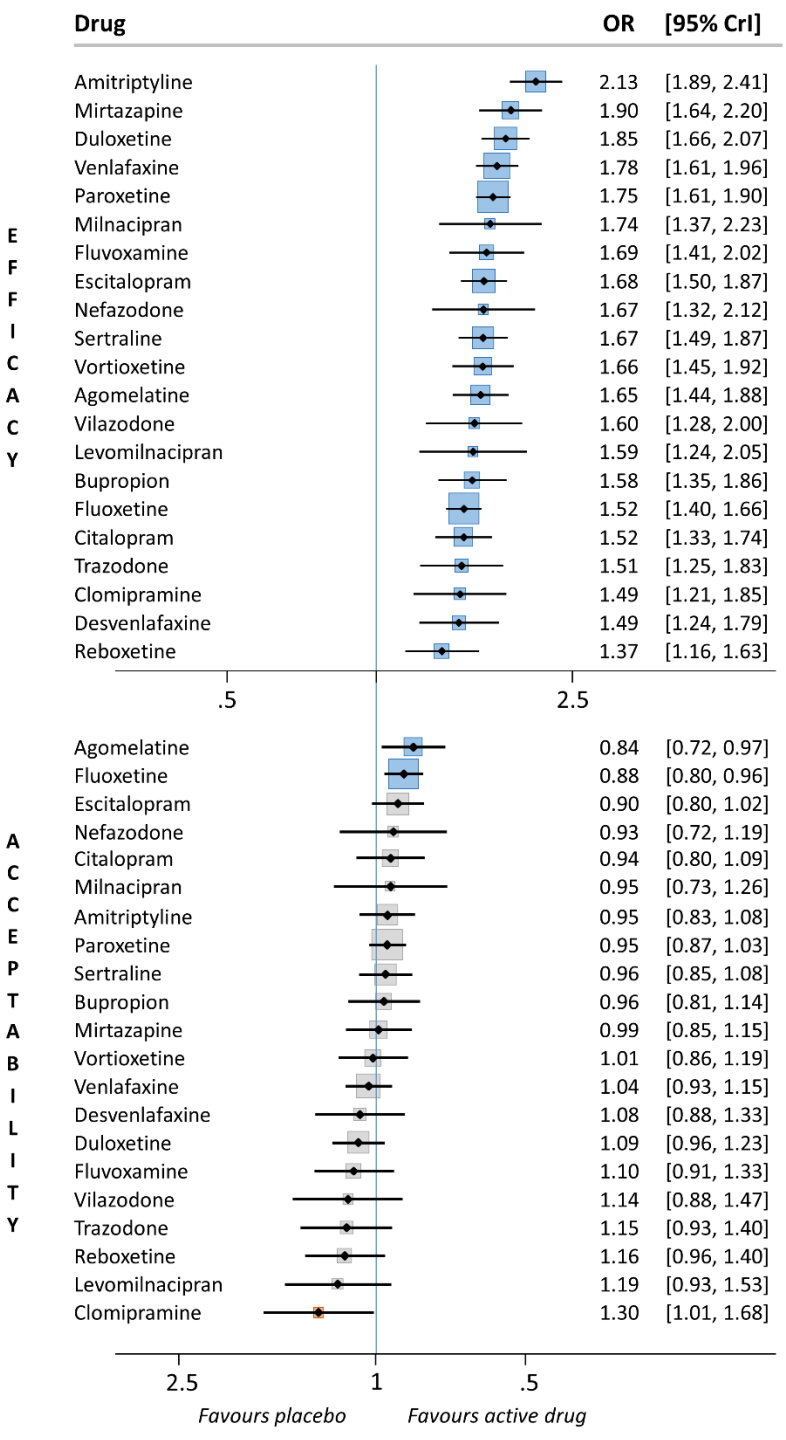
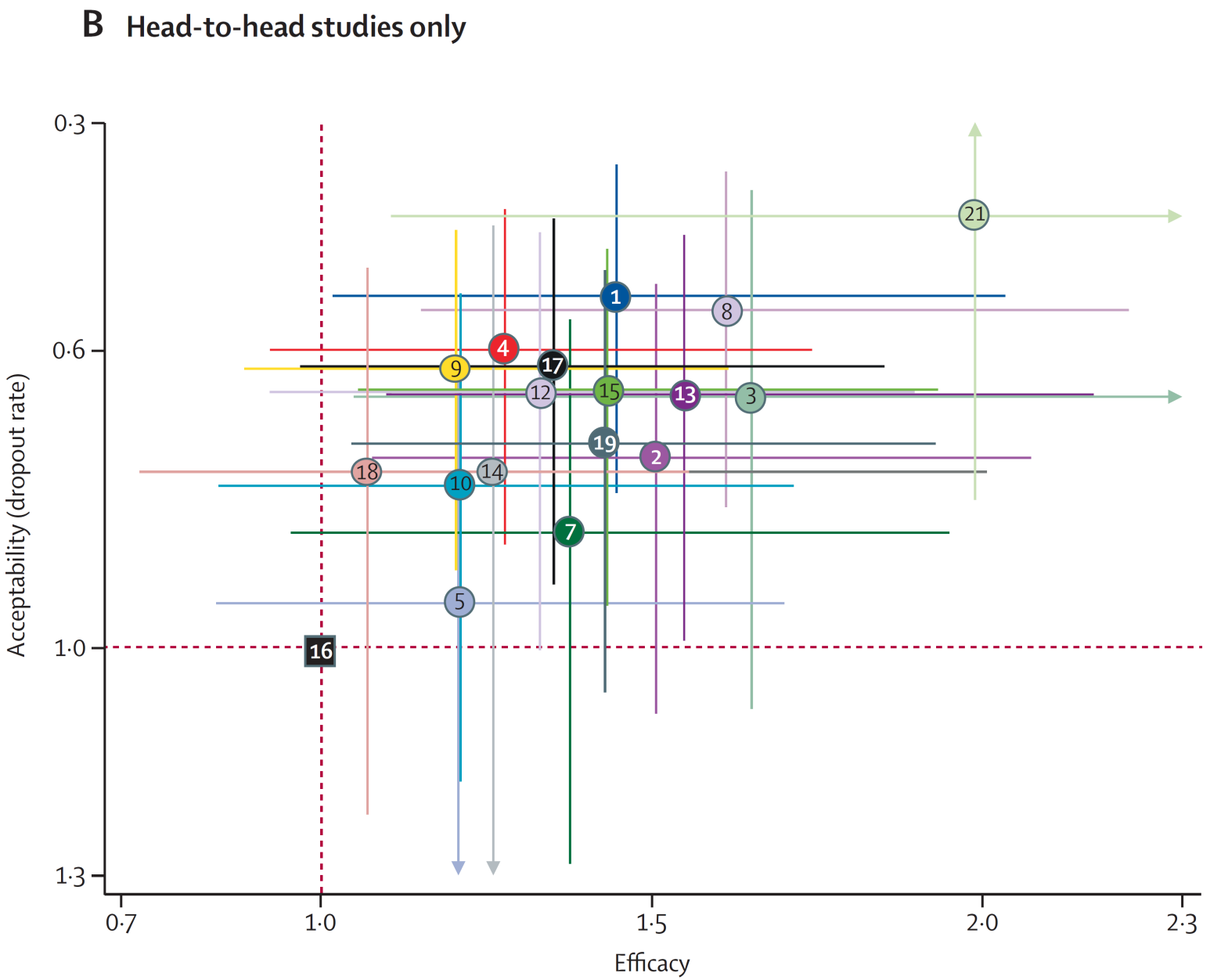


Figure 2: Network of eligible comparisons for efficacy (A) and acceptability (B). The width of the lines is proportional to the number of trials comparing every pair of treatments, and the size of every circle is proportional to the number of randomly assigned participants (sample size). Legend: Agom: agomelatine; Amit: amitriptyline; Bupr: bupropion; Cita: citalopram; Clom: clomipramine; Desv: desvenlafaxine; Dulo: duloxetine; Esci: escitalopram; Fluo: fluoxetine; Fluv: fluvoxamine; Levo: levomilnacipran; Miln: milnacipran; Mirt: mirtazapine; Nefa: nefazodone; Paro: paroxetine; Rebo: reboxetine; Sert: sertraline; Traz: trazodone; Venl: venlafaxine; Vila: vilazodone; Vort: vortioxetine.


Drug
OR
[95% CrI]
Amitriptyline 2.13 [1.89, 2.41]
Mirtazapine 1.90 [1.64, 2.20]
Duloxetine 1.85 [1.66, 2.07]
Venlafaxine 1.78 [1.61, 1.96]
Paroxetine 1.75 [1.61, 1.90]
Milnacipran 1.74 [1.37, 2.23]
Fluvoxamine 1.69 [1.41, 2.02]
Escitalopram 1.68 [1.50, 1.87]
Nefazodone 1.67 [1.32, 2.12]
Sertraline 1.67 [1.49, 1.87]
Vortioxetine 1.66 [1.45, 1.92]
Agomelatine 1.65 [1.44, 1.88]
Vilazodone 1.60 [1.28, 2.00]
Levomilnacipran 1.59 [1.24, 2.05]
Bupropion 1.58 [1.35, 1.86]
Fluoxetine 1.52 [1.40, 1.66]
Citalopram 1.52 [1.33, 1.74]
Trazodone 1.51 [1.25, 1.83]
Clomipramine 1.49 [1.21, 1.85]
Desvenlafaxine 1.49 [1.24, 1.79]
Reboxetine 1.37 [1.16, 1.63]
EFFICACY
.5
2.5
Favours placebo
Favours active drug


Drug
OR
[95% CrI]
ACCEPTABILITY
Agomelatine 0.84 [0.72, 0.97]
Fluoxetine 0.88 [0.80, 0.96]
Escitalopram 0.90 [0.80, 1.02]
Nefazodone 0.93 [0.72, 1.19]
Citalopram 0.94 [0.80, 1.09]
Milnacipran 0.95 [0.73, 1.26]
Amitriptyline 0.95 [0.83, 1.08]
Paroxetine 0.95 [0.87, 1.03]
Sertraline 0.96 [0.85, 1.08]
Bupropion 0.96 [0.81, 1.14]
Mirtazapine 0.99 [0.85, 1.15]
Vortioxetine 1.01 [0.86, 1.19]
Venlafaxine 1.04 [0.93, 1.15]
Desvenlafaxine 1.08 [0.88, 1.33]
Duloxetine 1.09 [0.96, 1.23]
Fluvoxamine 1.10 [0.91, 1.33]
Vilazodone 1.14 [0.88, 1.47]
Trazodone 1.15 [0.93, 1.40]
Reboxetine 1.16 [0.96, 1.40]
Levomilnacipran 1.19 [0.93, 1.53]
Clomipramine 1.30 [1.01, 1.68]
2.5
1
.5
Favours placebo
Favours active drug


Drug
OR
[95% CrI]
EFFICACY
Amitriptyline 2.13 [1.89, 2.41]
Mirtazapine 1.90 [1.64, 2.20]
Duloxetine 1.85 [1.66, 2.07]
Venlafaxine 1.78 [1.61, 1.96]
Paroxetine 1.75 [1.61, 1.90]
Milnacipran 1.74 [1.37, 2.23]
Fluvoxamine 1.69 [1.41, 2.02]
Escitalopram 1.68 [1.50, 1.87]
Nefazodone 1.67 [1.32, 2.12]
Sertraline 1.67 [1.49, 1.87]
Vortioxetine 1.66 [1.45, 1.92]
Agomelatine 1.65 [1.44, 1.88]
Vilazodone 1.60 [1.28, 2.00]
Levomilnacipran 1.59 [1.24, 2.05]
Bupropion 1.58 [1.35, 1.86]
Fluoxetine 1.52 [1.40, 1.66]
Citalopram 1.52 [1.33, 1.74]
Trazodone 1.51 [1.25, 1.83]
Clomipramine 1.49 [1.21, 1.85]
Desvenlafaxine 1.49 [1.24, 1.79]
Reboxetine 1.37 [1.16, 1.63]
.5
2.5
ACCEPTABILITY
Agomelatine 0.84 [0.72, 0.97]
Fluoxetine 0.88 [0.80, 0.96]
Escitalopram 0.90 [0.80, 1.02]
Nefazodone 0.93 [0.72, 1.19]
Citalopram 0.94 [0.80, 1.09]
Milnacipran 0.95 [0.73, 1.26]
Amitriptyline 0.95 [0.83, 1.08]
Paroxetine 0.95 [0.87, 1.03]
Sertraline 0.96 [0.85, 1.08]
Bupropion 0.96 [0.81, 1.14]
Mirtazapine 0.99 [0.85, 1.15]
Vortioxetine 1.01 [0.86, 1.19]
Venlafaxine 1.04 [0.93, 1.15]
Desvenlafaxine 1.08 [0.88, 1.33]
Duloxetine 1.09 [0.96, 1.23]
Fluvoxamine 1.10 [0.91, 1.33]
Vilazodone 1.14 [0.88, 1.47]
Trazodone 1.15 [0.93, 1.40]
Reboxetine 1.16 [0.96, 1.40]
Levomilnacipran 1.19 [0.93, 1.53]
Clomipramine 1.30 [1.01, 1.68]
2.5
1
.5
Favours placebo
Favours active drug


B Head-to-head studies only
0·3
0·6
1·0
1·3
Acceptability (dropout rate)
0·7
1·0
1·5
2·0
2·3
Efficacy
1
2
3
4
5
7
8
9
10
12
13
14
15
16
17
18
19
21

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

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

A primer on network meta-analysis with emphasis on mental health

Dimitris Mavridis,[1,2] Myrsini Giannatsi,[1] Andrea Cipriani,[3] Georgia Salanti[1]

[1]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; [2]Department of Primary Education, University of Ioannina, Ioannina, Greece; [3]Department of Psychiatry, University of Oxford, Oxford, UK
Correspondence to Dr Dimitris Mavridis, dimi.mavridis@googlemail.com

ABSTRACT

Objective A quantitative synthesis of evidence via standard pair-wise meta-analysis lies on the top of the hierarchy for evaluating the relative effectiveness or safety between two interventions. In most healthcare problems, however, there is a plethora of competing interventions. Network meta-analysis allows to rank competing interventions and evaluate their relative effectiveness even if they have not been compared in an individual trial. The aim of this paper is to explain and discuss the main features of this statistical technique.

Methods We present the key assumptions underlying network meta-analysis and the graphical methods to visualise results and information in the network. We used one illustrative example that compared the relative effectiveness of 15 antimanic drugs and placebo in acute mania.

Results A network plot allows to visualise how information flows in the network and reveals important information about network geometry. Discrepancies between direct and indirect evidence can be detected using inconsistency plots. Relative effectiveness or safety of competing interventions can be presented in a league table. A contribution plot reveals the contribution of each direct comparison to each network estimate. A comparison-adjusted funnel plot is an extension of simple funnel plot to network meta-analysis. A rank probability matrix can be estimated to present the probabilities of all interventions assuming each rank and can be represented using rankograms and cumulative probability plots.

Conclusions Network meta-analysis is very helpful in comparing the relative effectiveness and acceptability of competing treatments. Several issues, however, still need to be addressed when conducting a network meta-analysis for the results to be valid and correctly interpreted.

INTRODUCTION

Evidence-based practices are crucial in informing healthcare decisions as they provide evidence on the effectiveness and adverse effects of the available treatment options. A quantitative synthesis of research findings from randomised controlled trials (RCTs) via meta-analysis lies at the top of evidence based methods.[1] The benefits from meta-analysis are well established and include increased power, more precise effect estimates, and ability to generalise research findings and identify factors that modify the effect of an intervention (effect modifiers). In mental health, several meta-analyses have identified interventions assess the comparative efficacy and tolerability of competing treatments for various disorders.[11–14]

BASIC CONCEPTS AND ASSUMPTIONS IN NMA

A fundamental concept in NMA is that of an indirect comparison. If two treatments, A and B, have both been compared with a common treatment, say C, in two different sets of trials (A vs C and B vs C), then the relative effectiveness between A and B can be estimated indirectly via the common comparator C.[15] For illustrative purposes, we will consider three active antipsychotics, namely haloperidol (H), olanzapine

Thanks!

andrea.cipriani@psych.ox.ac.uk

@And_Cipriani