From standard meta-analysis to network meta-analysis: the example of antidepressants, antimanics and antipsychotics

Andrea Cipriani

Oxford, June 29th 2016

Conflict of interest: None

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

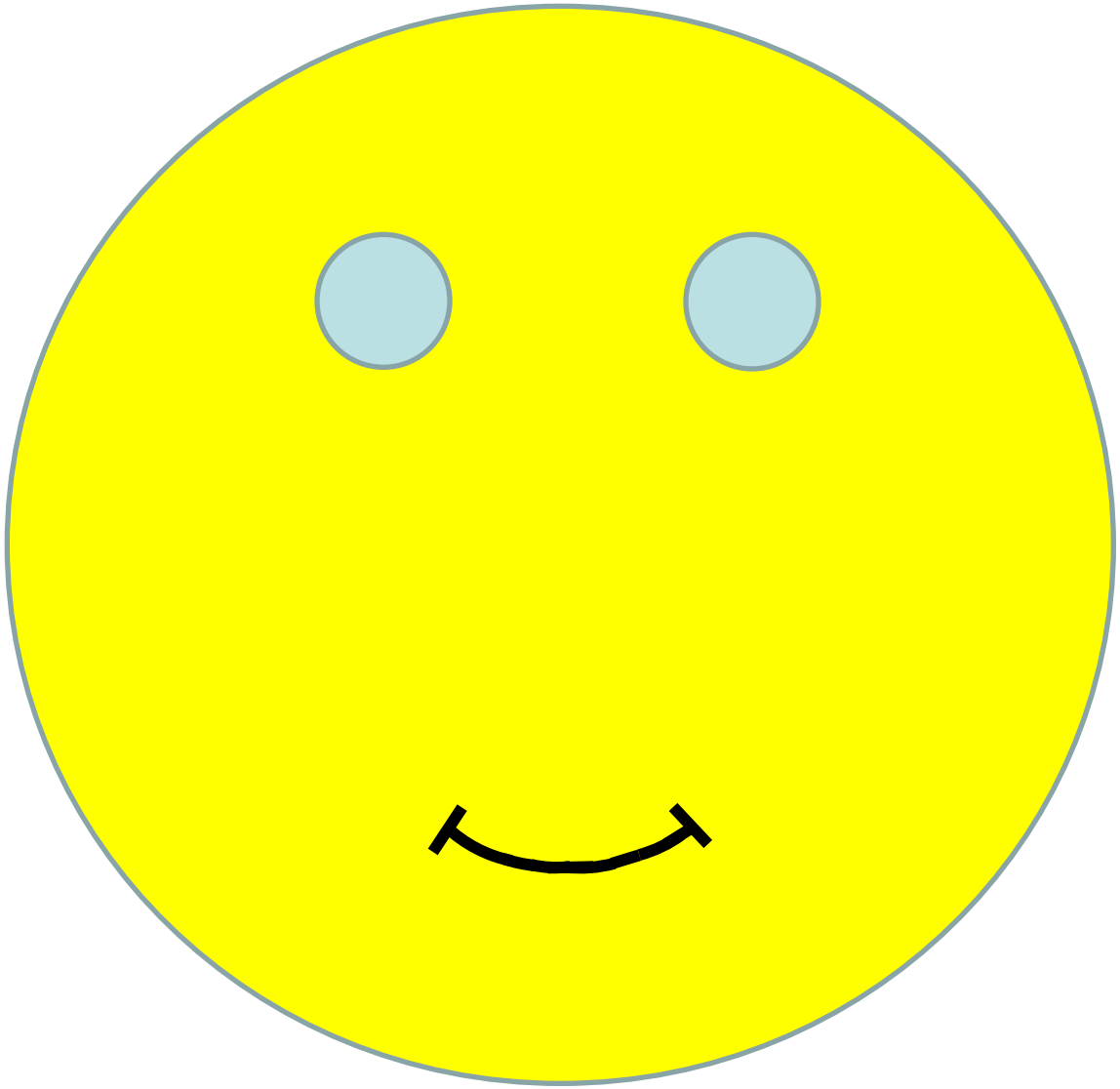
Results:

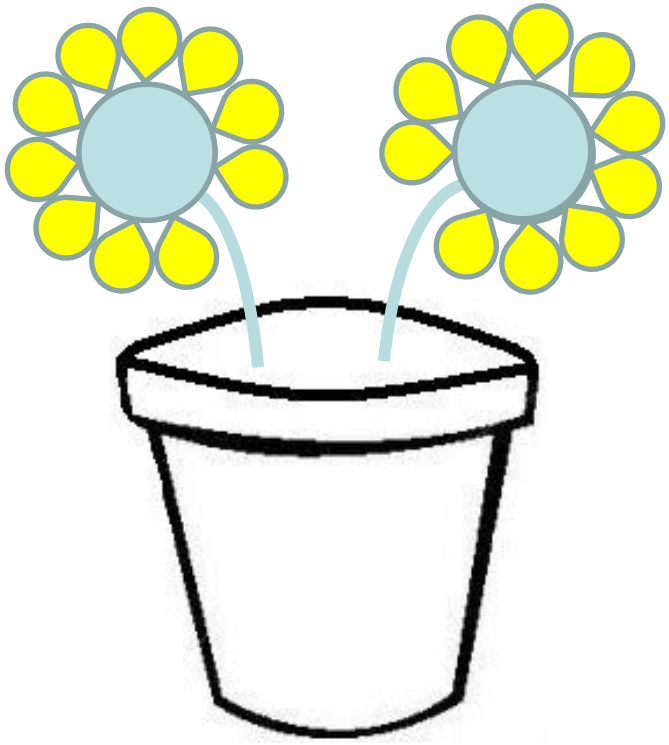
On measures of wax clearance Cerumol, sodium bicarbonate, olive oil and water are all more effective than no treatment; triethanolamine polypeptide (TP) is better than olive oil; wet irrigation is better than dry irrigation; sodium bicarbonate drops followed by irrigation by nurse is more effective than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

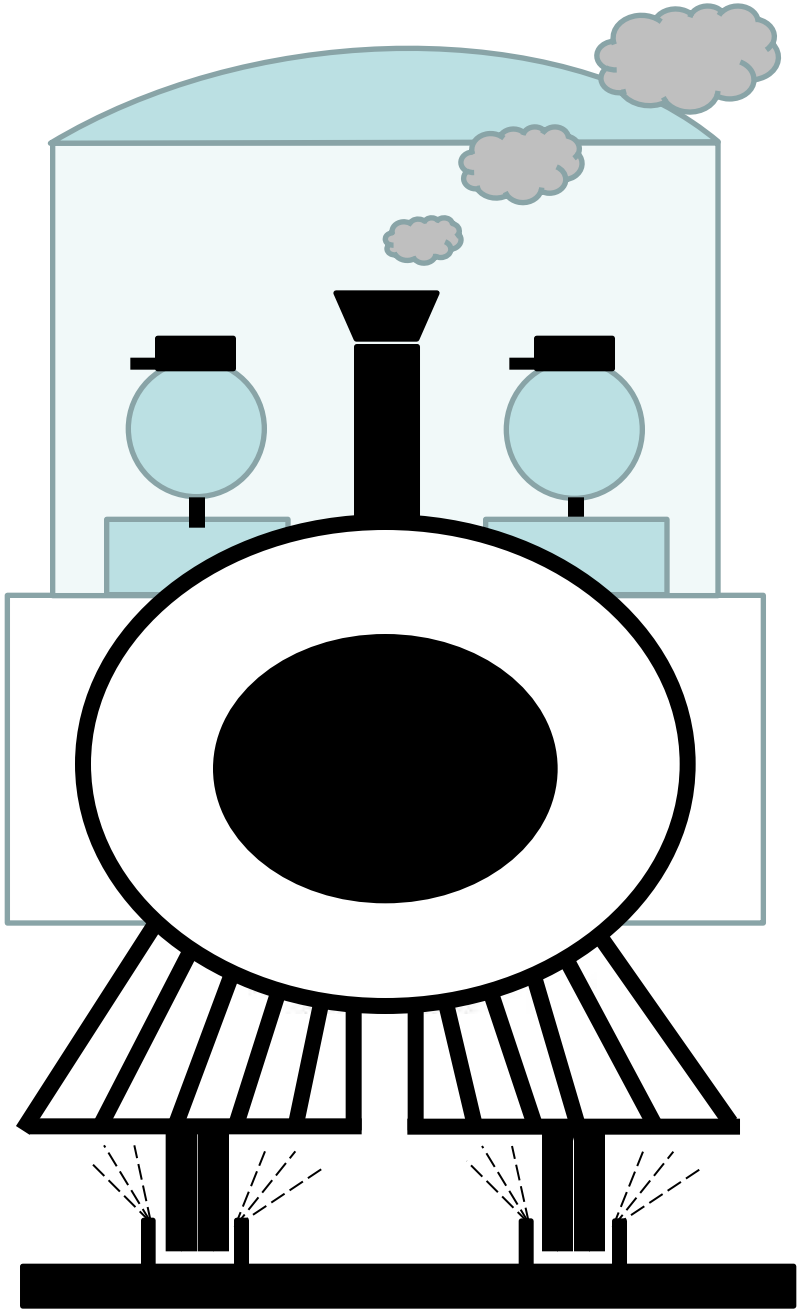
Ideally, clinicians would like to know how all the different options rank against each other and how big the differences are between all the available options.

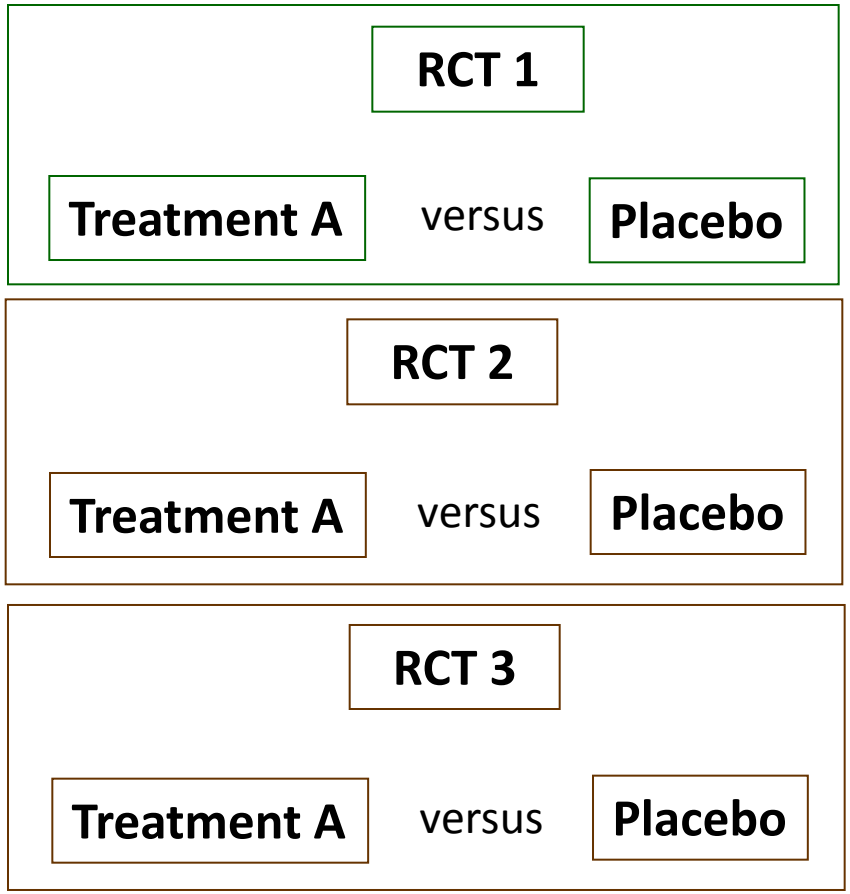
Need for a new robust framework which answers directly critical decision-making questions.

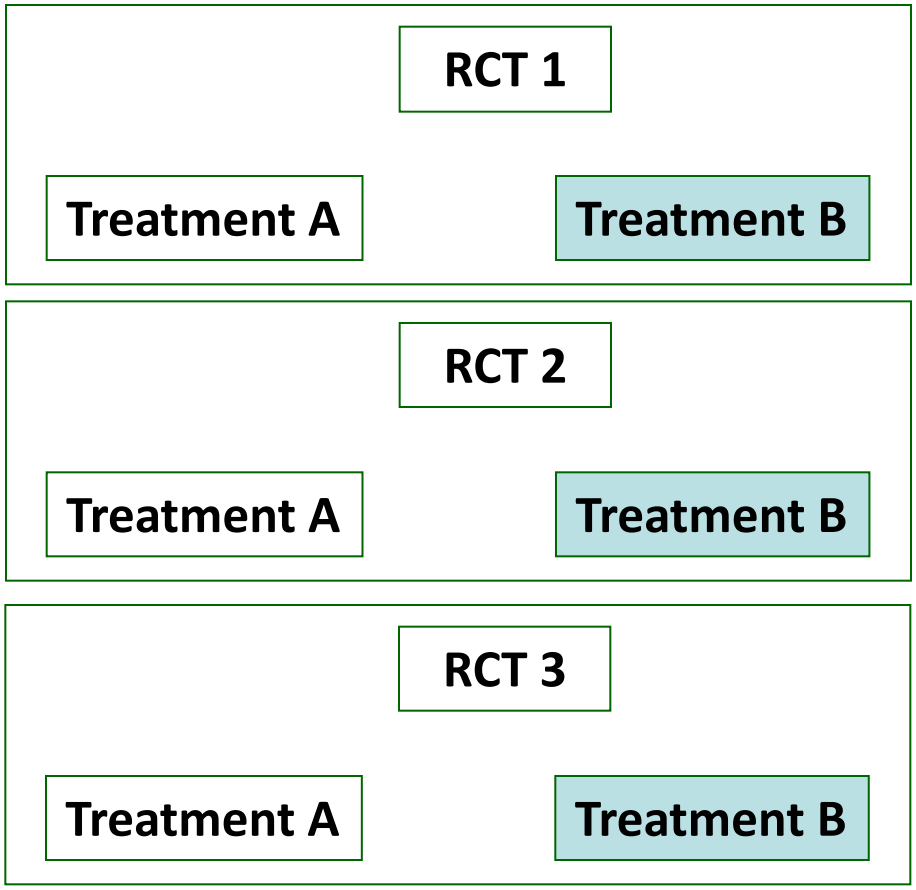
This emerging review format is not popular yet because of important difficulties with the necessary statistical component - the **multiple treatments meta-analysis** or **network meta-analysis**.





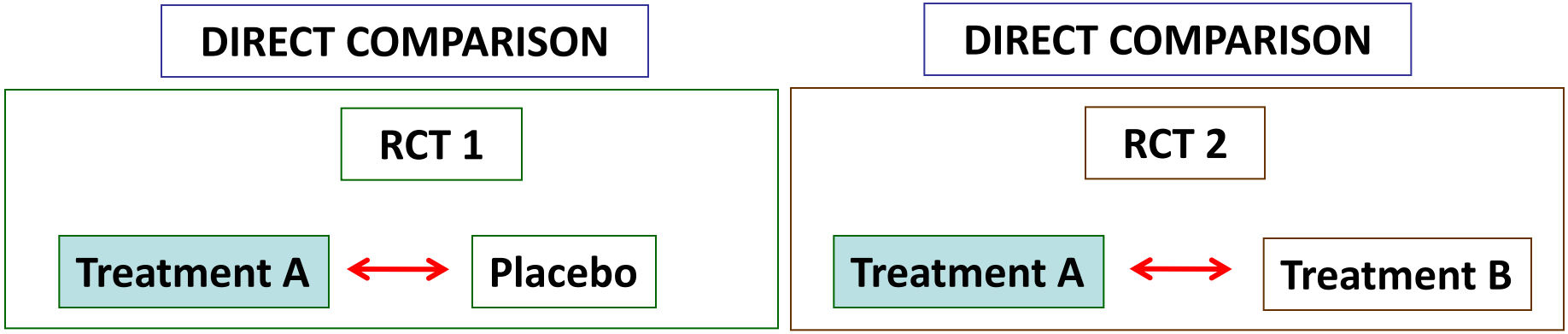


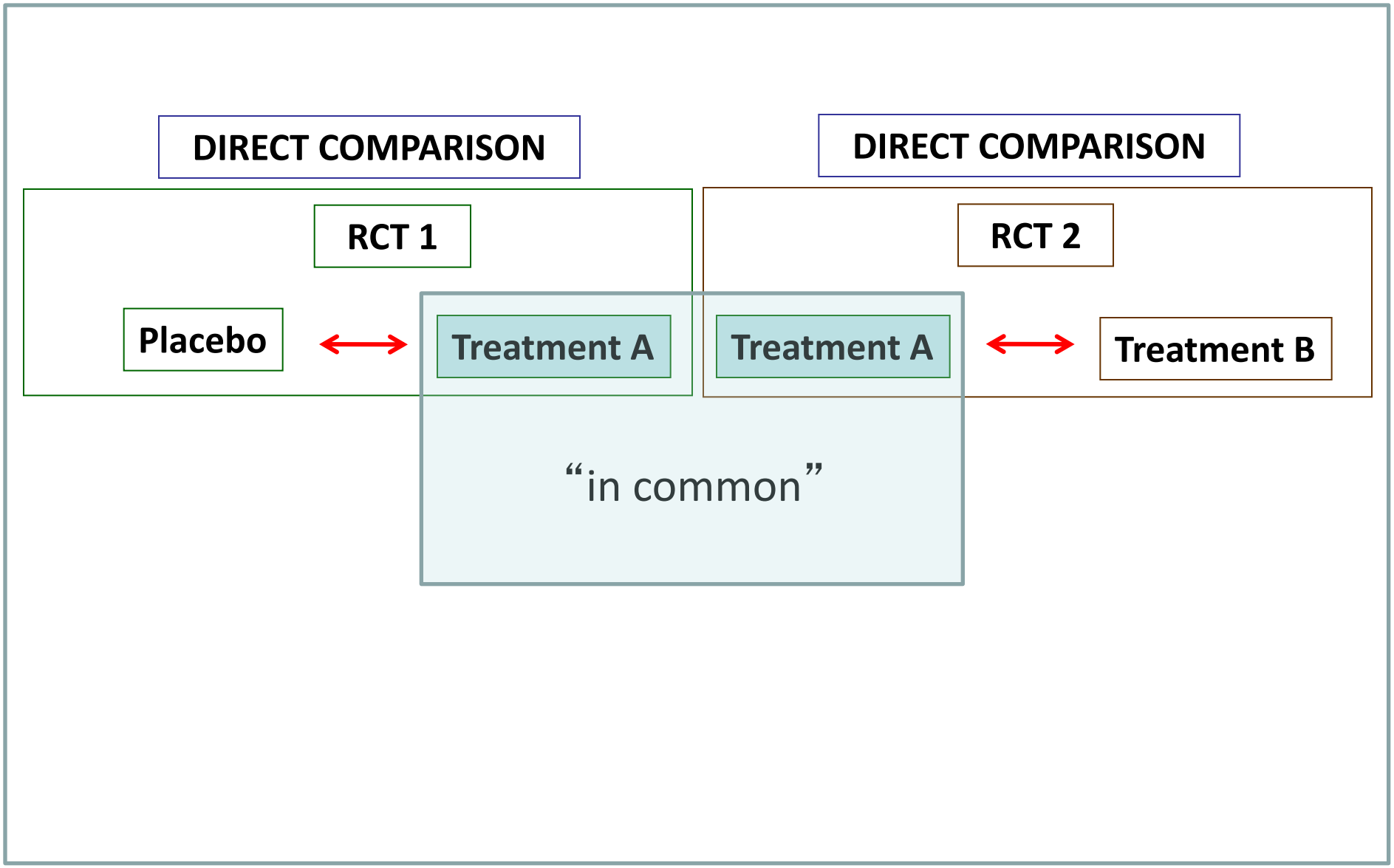


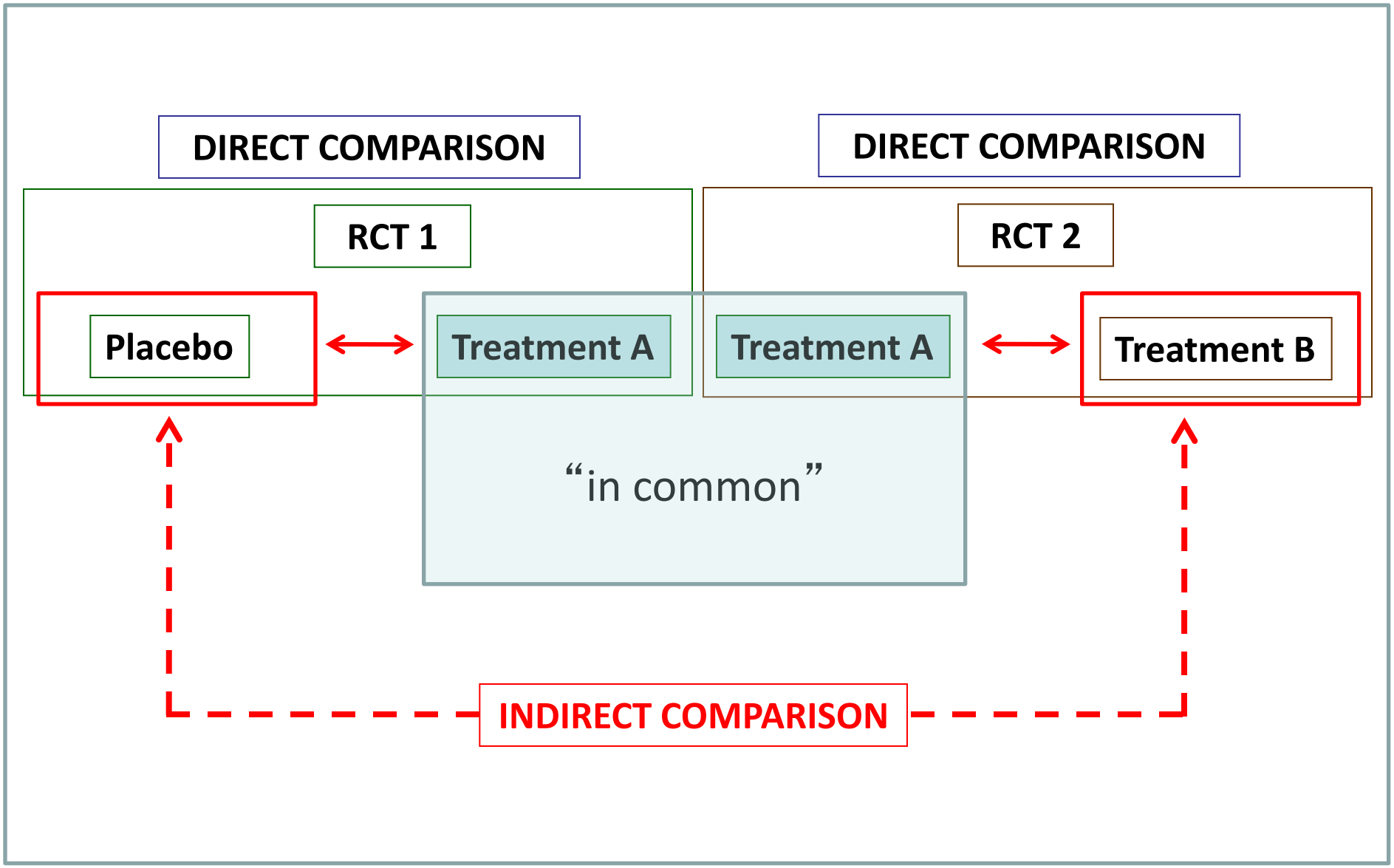


= standard meta-analysis

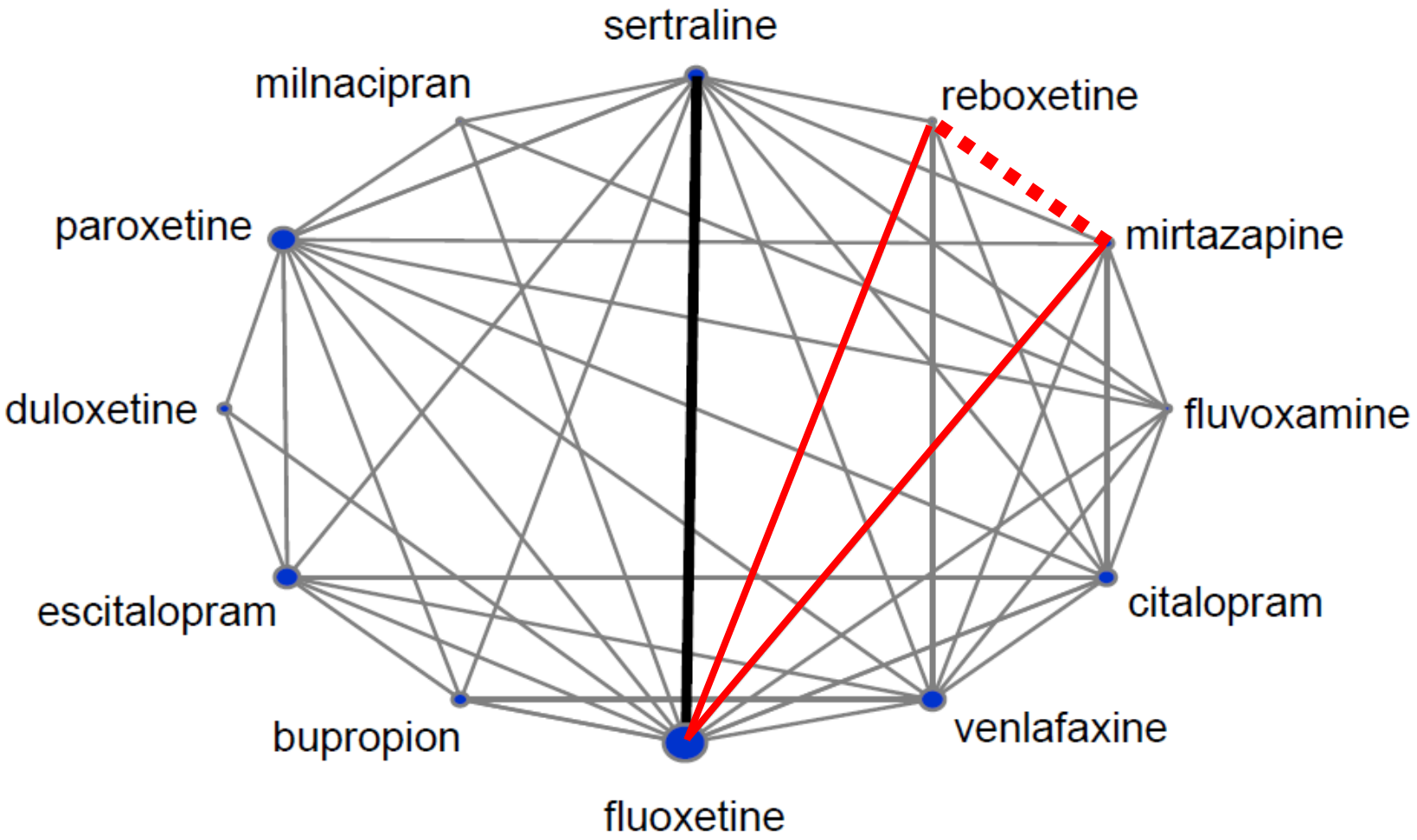
= standard meta-analysis


DIRECT COMPARISON
RCT 1
Treatment A
Placebo
DIRECT COMPARISON
RCT 2
Treatment A
Treatment B

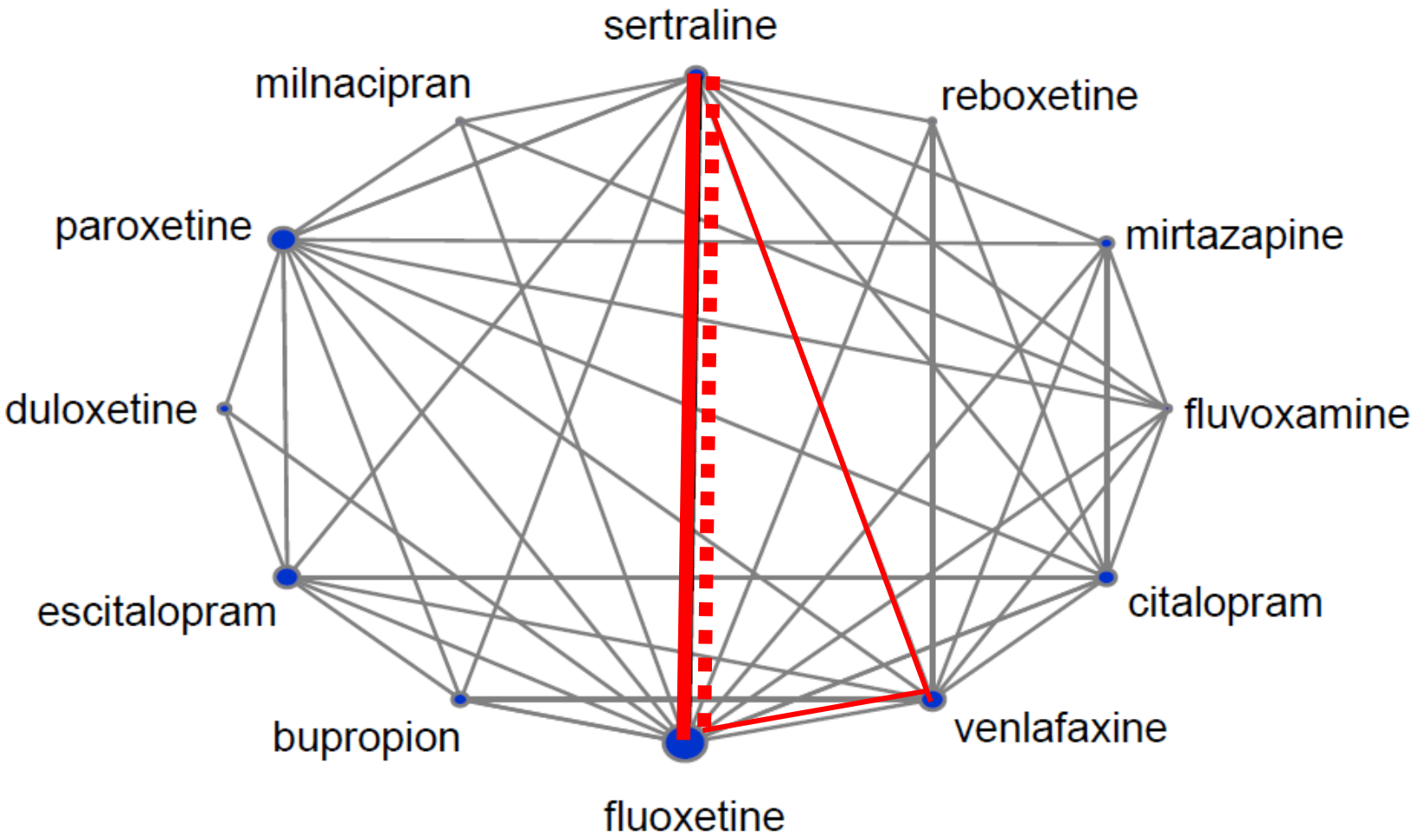

DIRECT COMPARISON
DIRECT COMPARISON
RCT 1
RCT 2
Placebo
Treatment A
Treatment A
Treatment B
“in common”


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

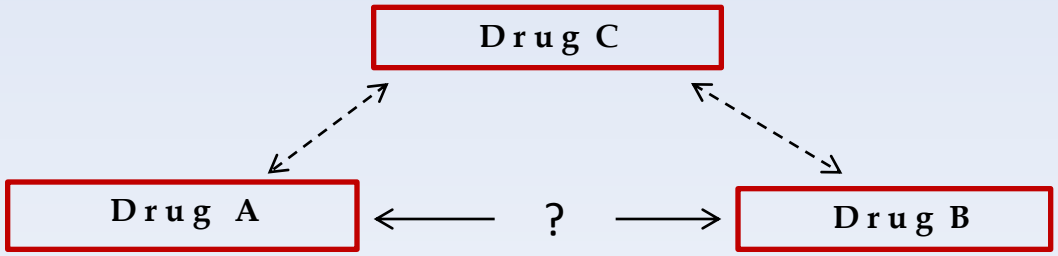
Network of experimental comparisons

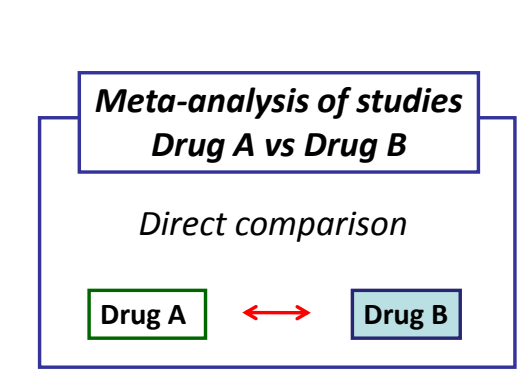


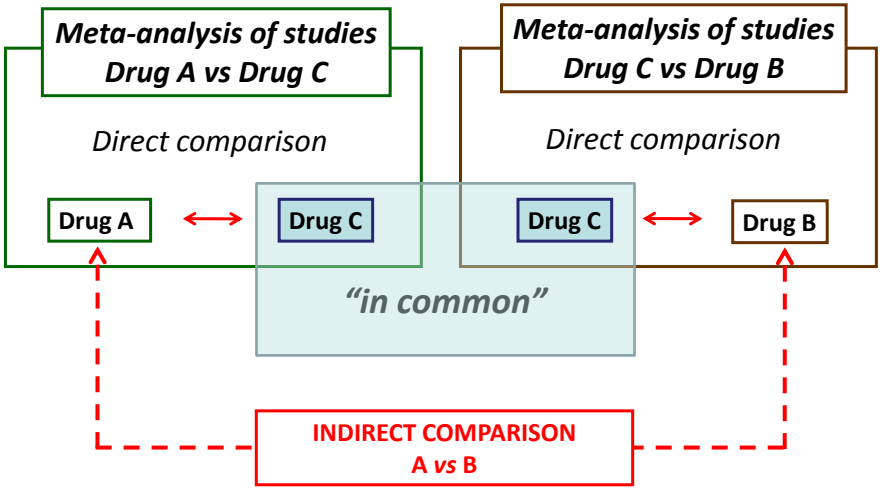
Network of experimental comparisons



How to compare treatment A vs B in a network of trials?







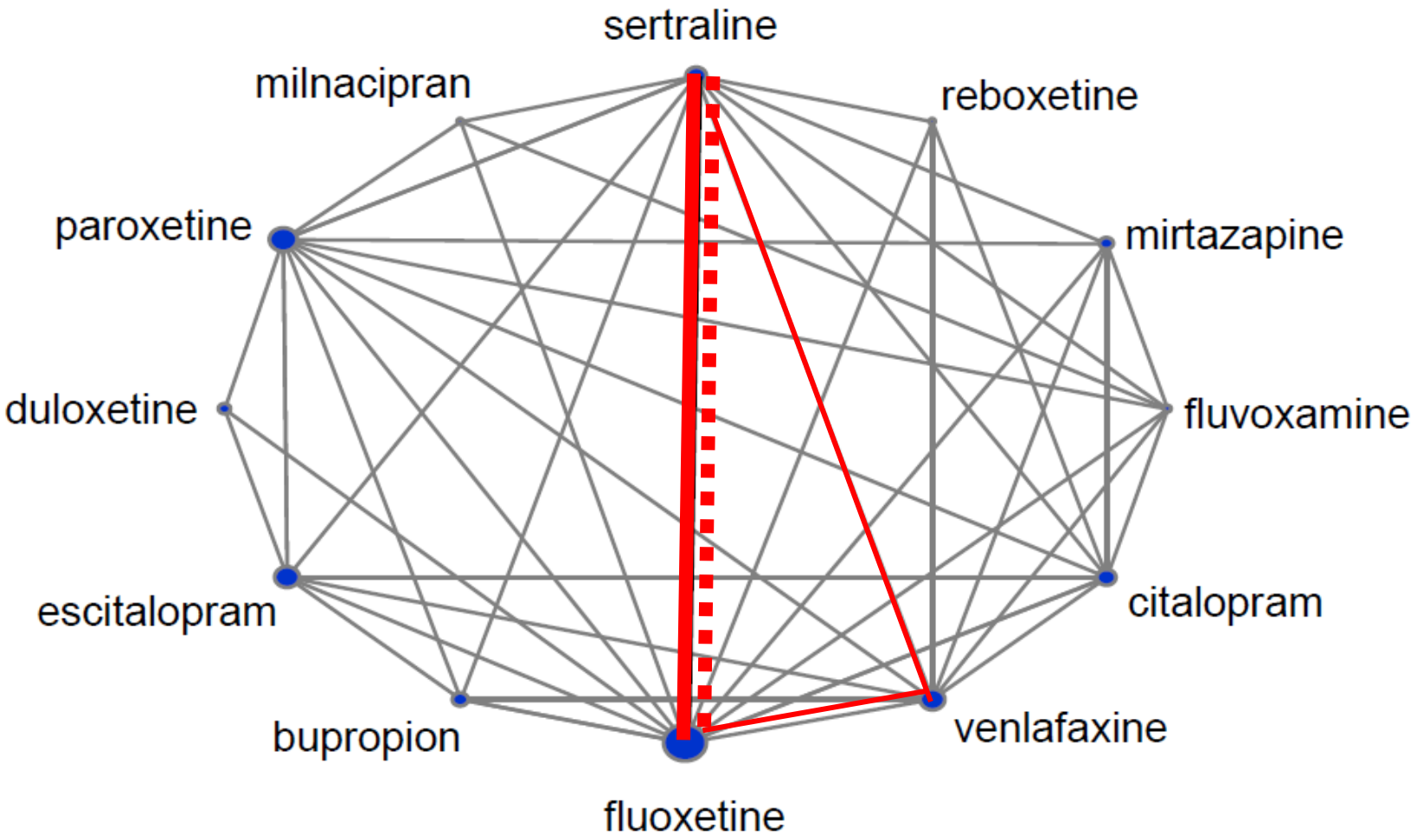
DIRECT ESTIMATE
A vs B

INDIRECT ESTIMATE
A vs B

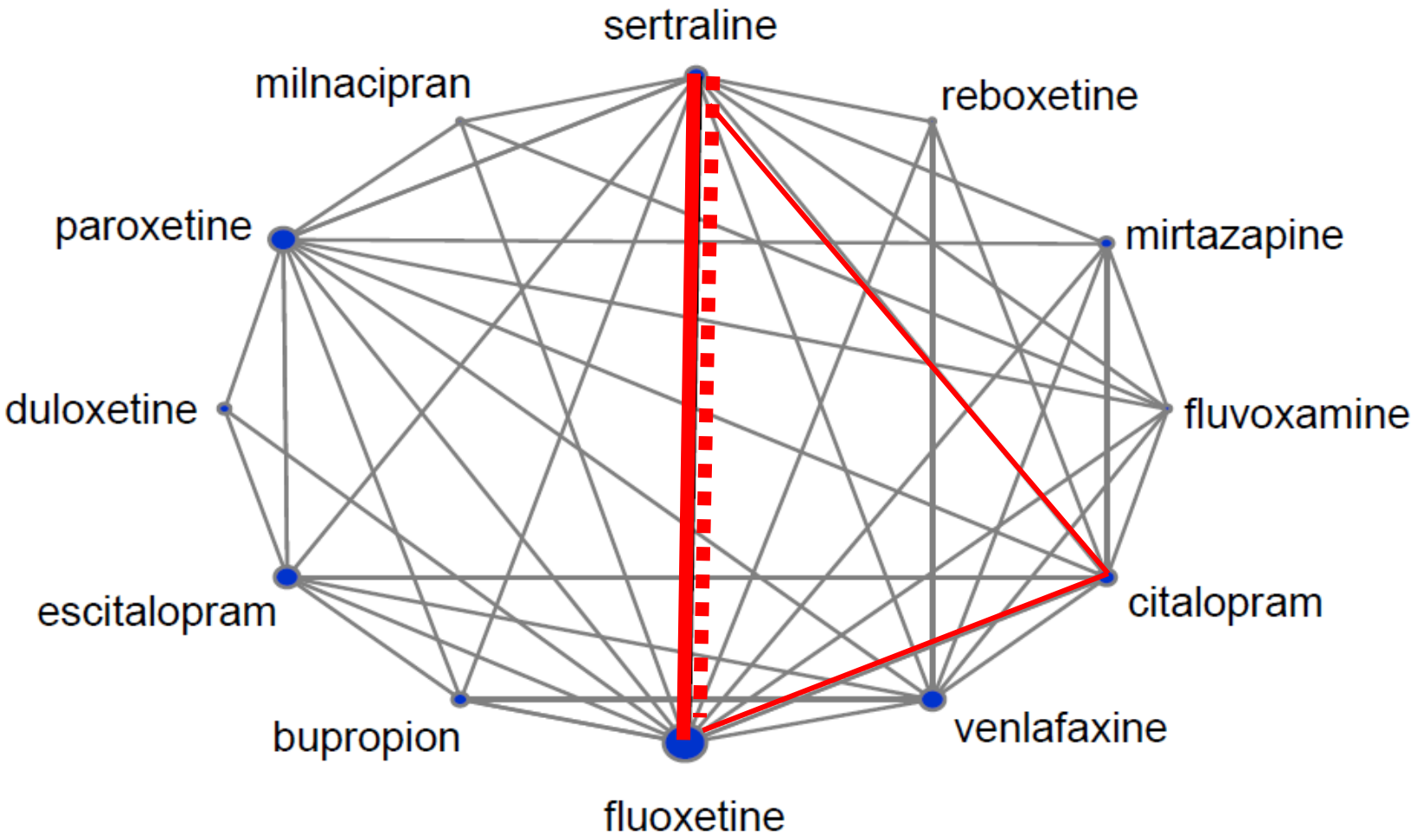
MIXED ESTIMATE (DIRECT AND INDIRECT)

The combination of direct and indirect evidence into a single effect size for treatment A versus B (*mixed estimate*)

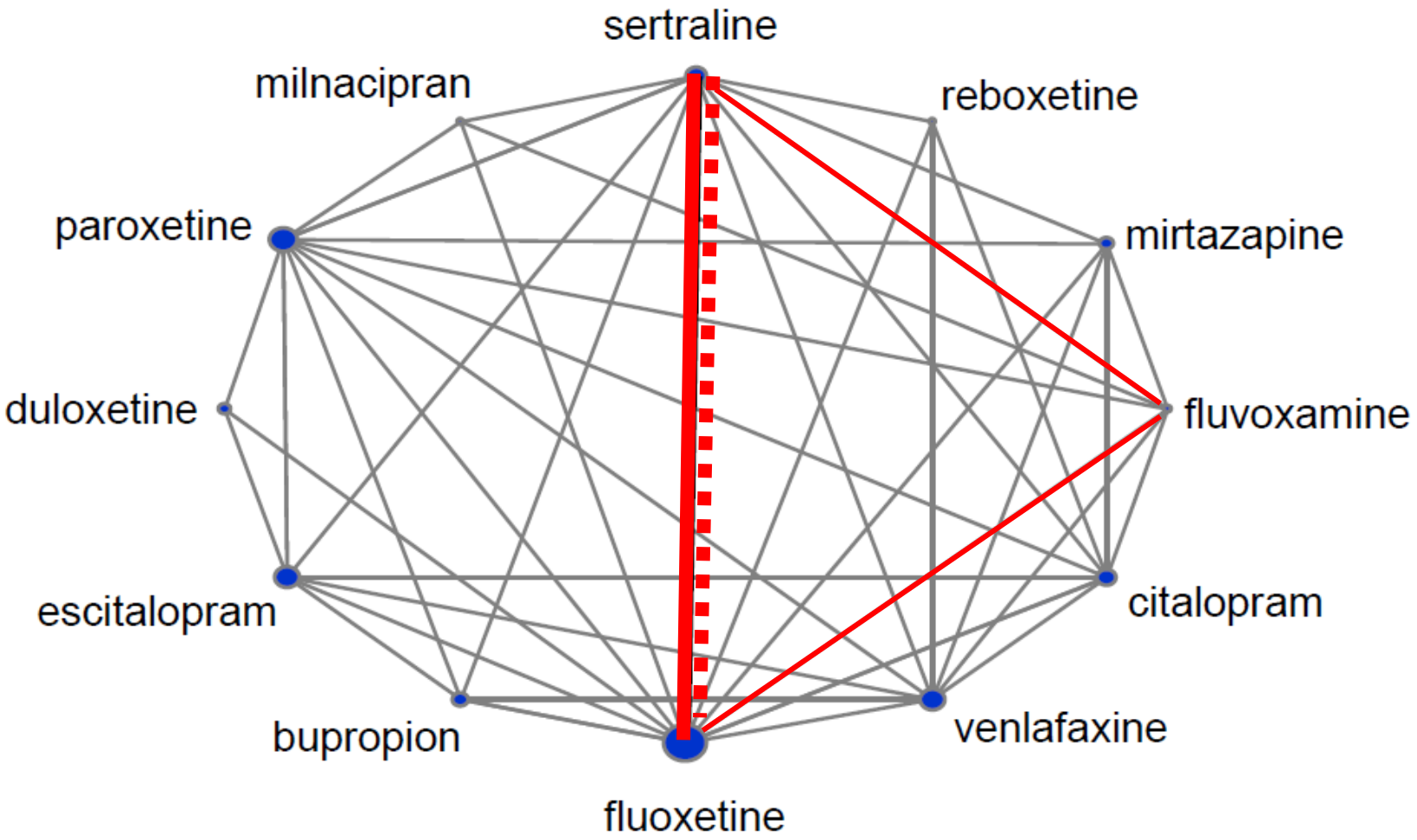
Network of experimental comparisons



Network of experimental comparisons



Network of experimental comparisons



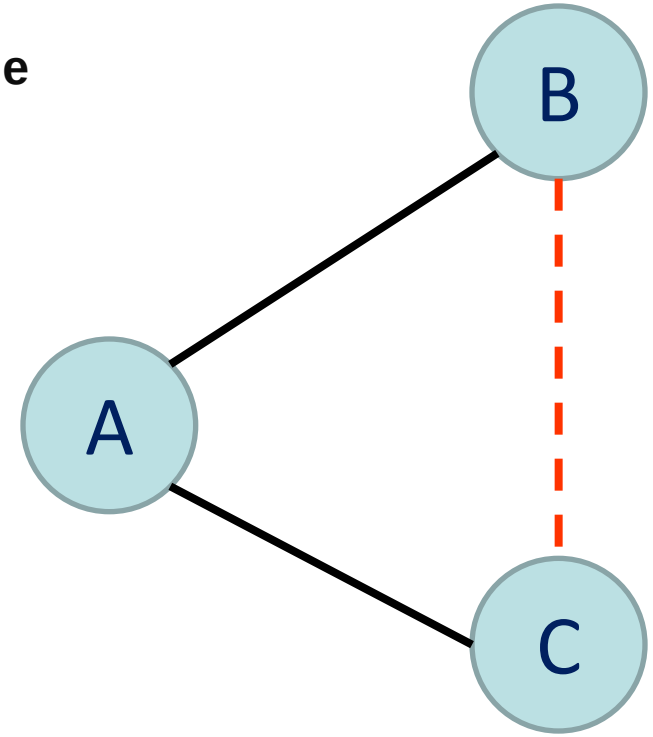
Advantages of MTM

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

Transitivity

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

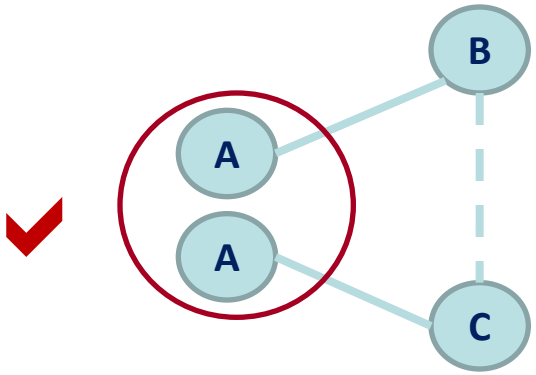
It is an **untestable** assumption



The anchor treatment A is ‘transitive’

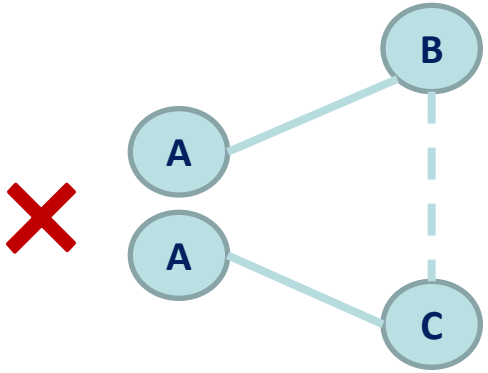
….but you can evaluate clinically, epidemiologically and statistically its plausibility

Transitivity means... (1)

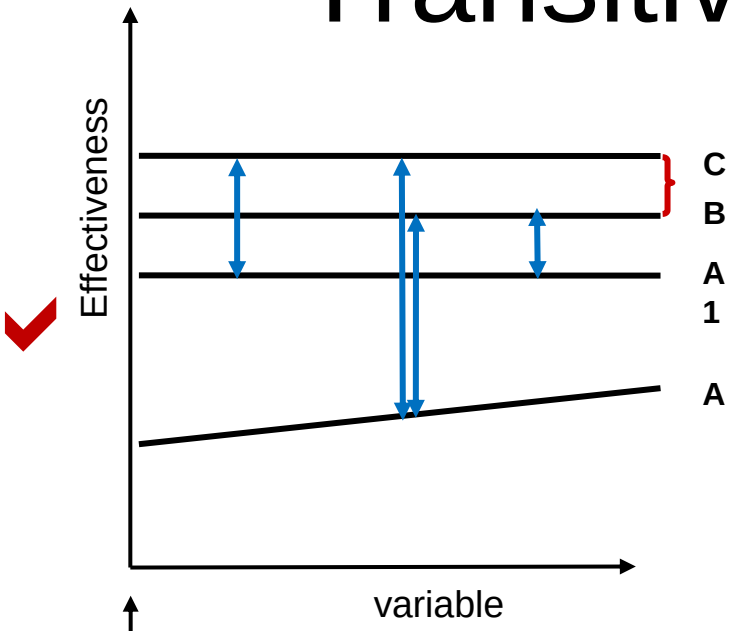


Treatment A is similar when it appears in AB and AC trials

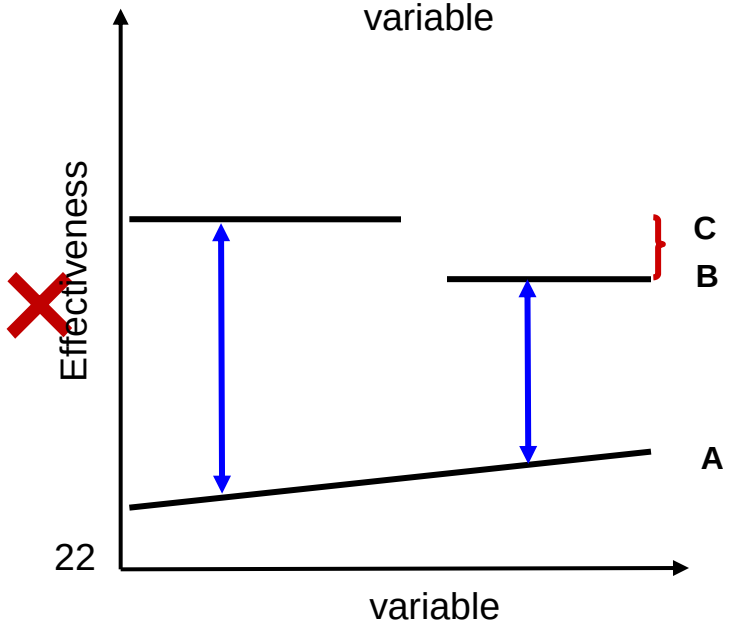
Plausible when A is placebo given in different forms?
(e.g. injection versus pill)?



Transitivity means... (3)

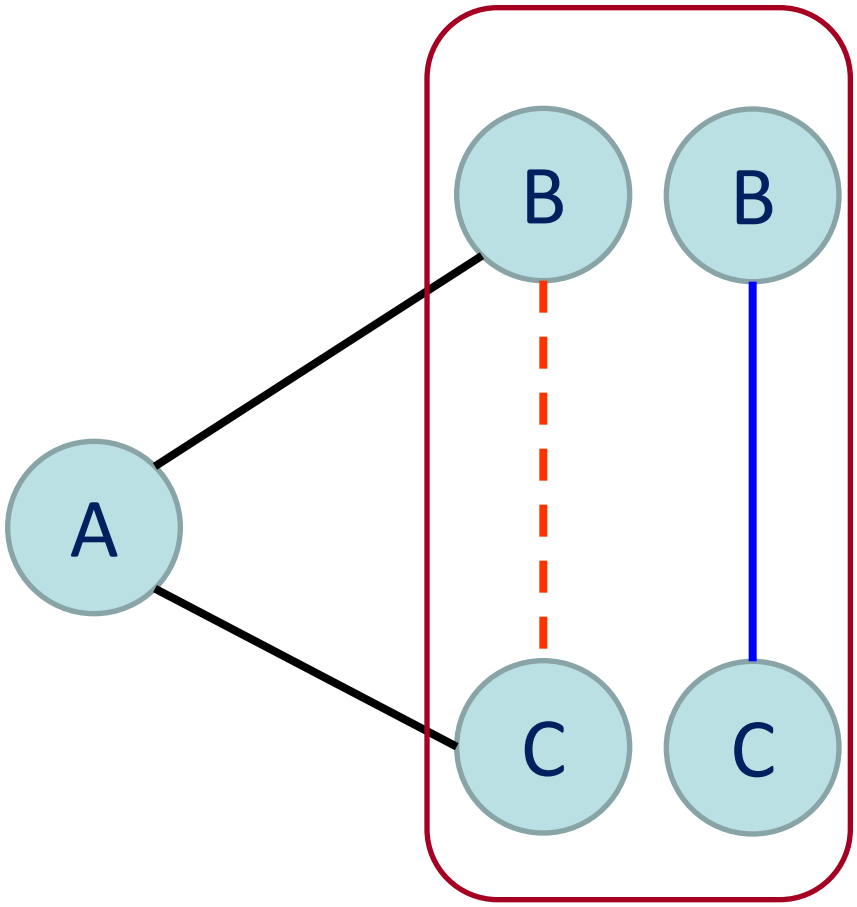


...that AC and AB trials do not differ with respect to the distribution of effect modifiers



Consistency

Direct and indirect evidence are in agreement



μ^{I}_{B} μ^{D}_{B} → μ^{M}_{B}

C C C

Heterogeneity?

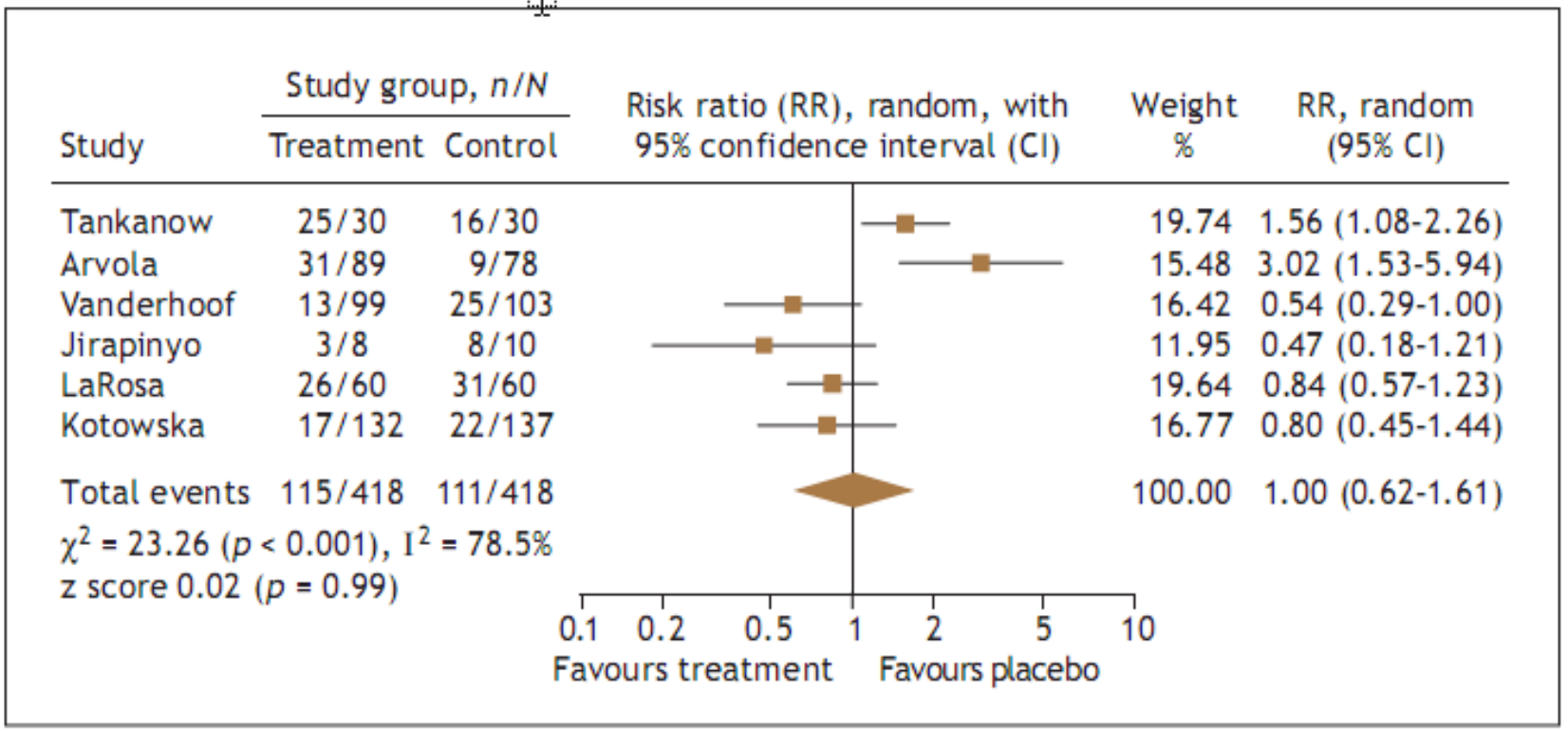
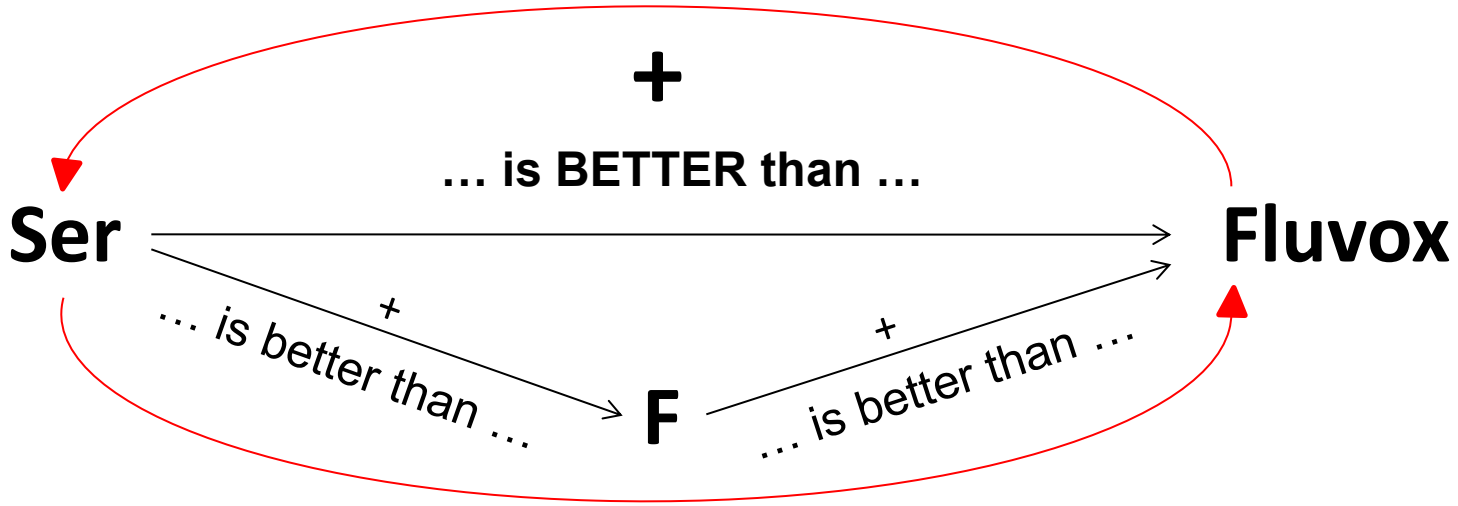


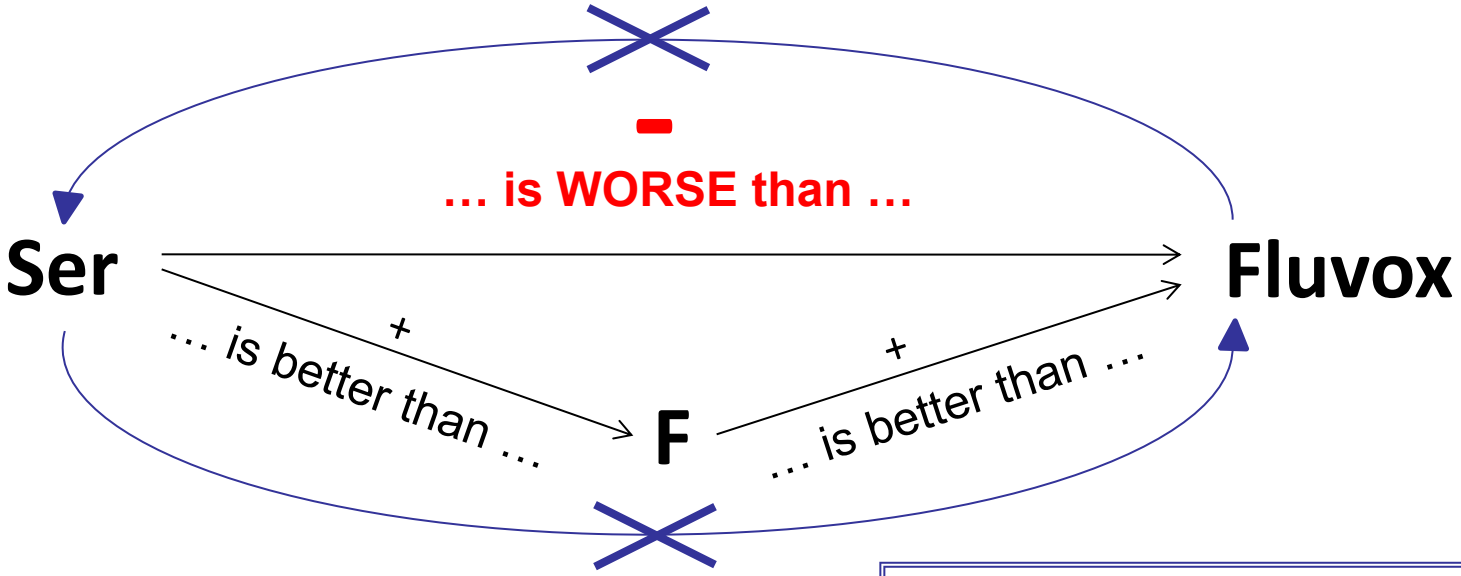
Fig 3: Incidence of antibiotic-associated diarrhea — intention-to-treat analysis. The analysis showed a nonsignificant difference between probiotics and placebo (z score) and statistically significant heterogeneity.

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

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



THE LOOP IS CONSISTENT



THE LOOP IS *INCONSISTENT*

Advantages of MTM

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

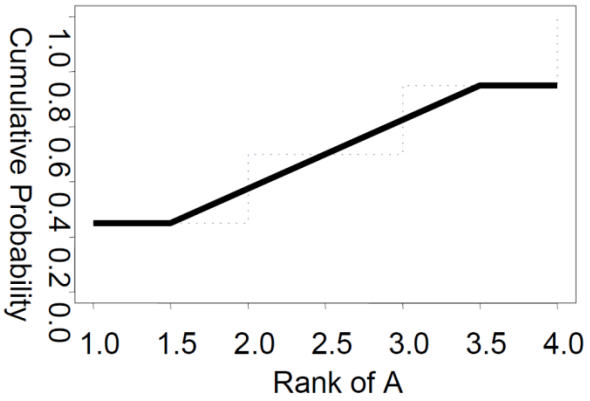
Ranking measures from MTM

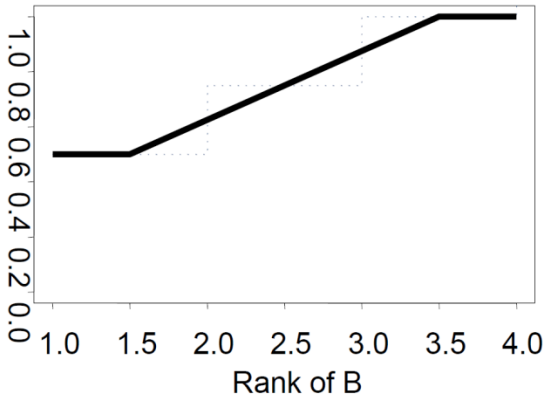
- Estimate for each treatment the probability to be the best

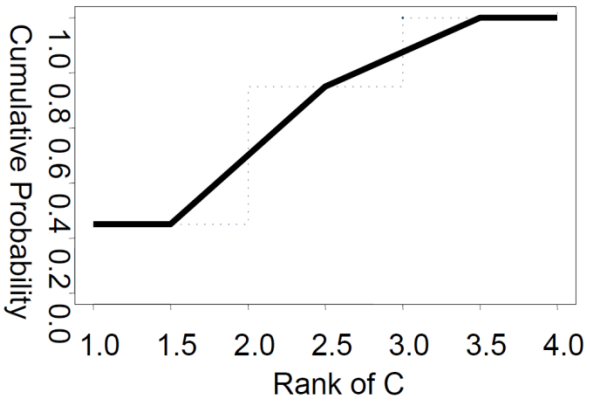
% probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00
j=2	0.25	0.25	0.50	0.00
j=3	0.25	0.25	0.25	0.25
j=4	0.25	0	0	0.75

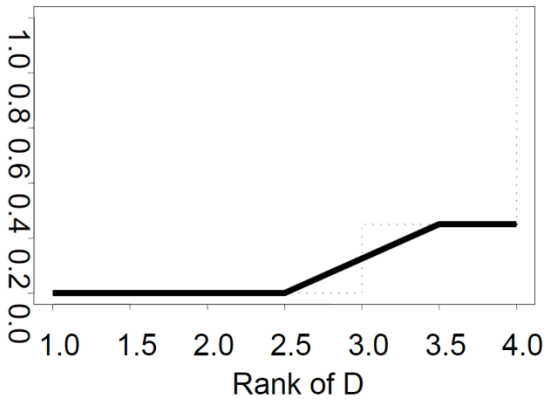
% probability	A	B	C	D
j=1	0.25	0.50	0.25	0.00
j=2	0.50	0.75	0.75	0.00
j=3	0.75	1.00	1.00	0.25
j=4	1.00	1.00	1.00	1.00

i the treatment
j the rank



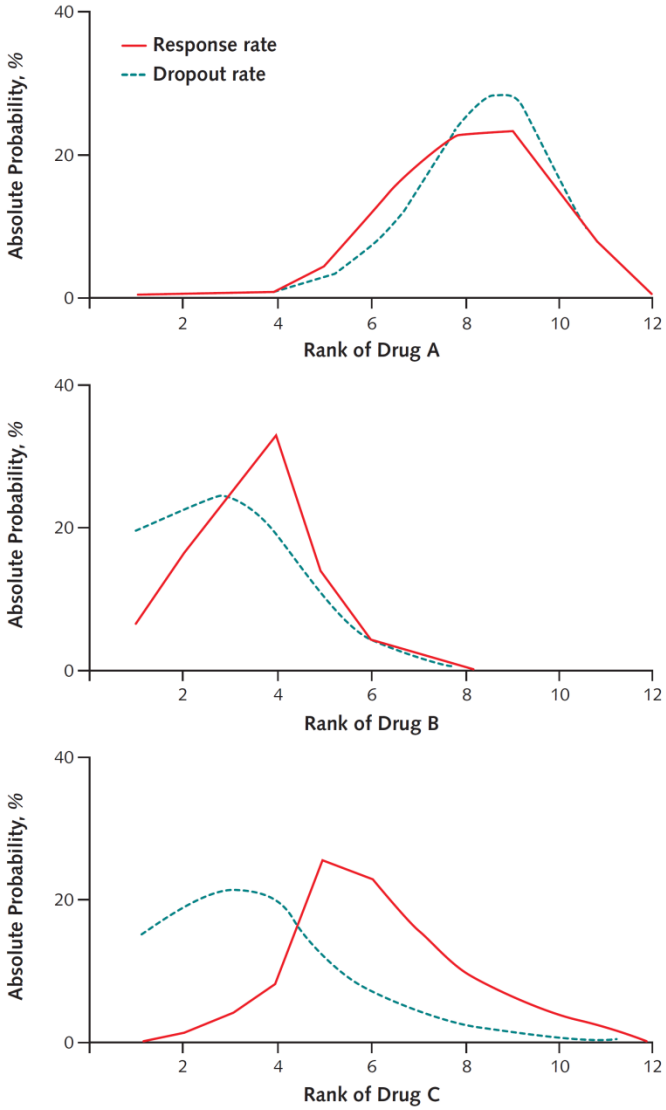




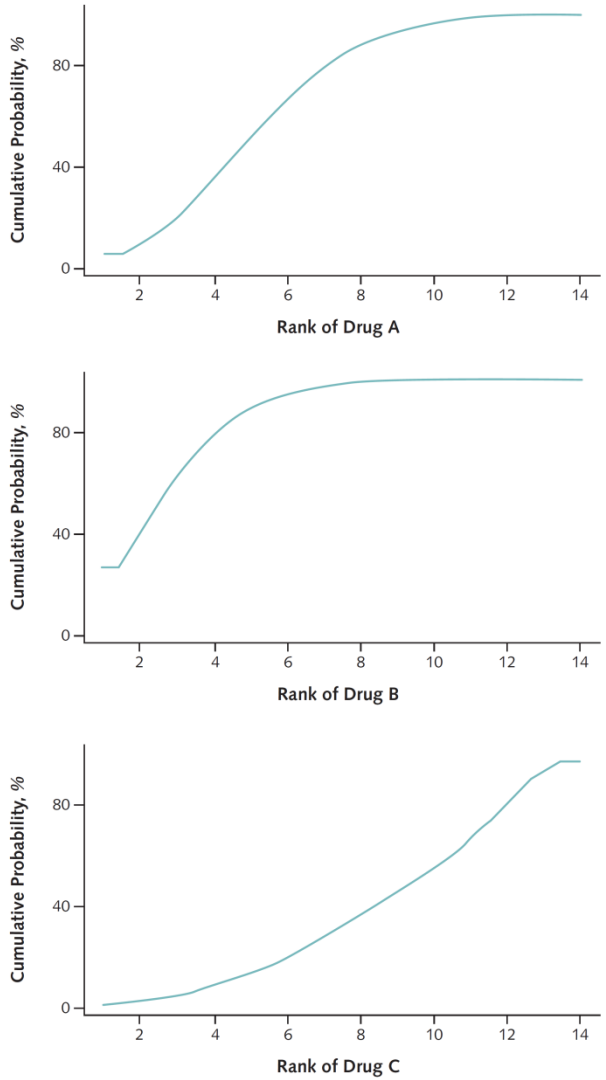


The areas under the cumulative curves for the four treatments of the example above are
A=0.5
B=0.75
C=0.67
D=0.08

Appendix Figure 2. **Absolute rankograms for presenting probabilities and rankings in network meta-analysis.**



Appendix Figure 3. **Cumulative rankograms for presenting probabilities and rankings in network meta-analysis.**



Treatment of depression

CLINICAL GUIDELINES | Annals of Internal Medicine

Comparative Benefits and Harms of Second-Generation Antidepressants: Background Paper for the American College of Physicians

Gerald Gartlehner, MD, MPH; Bradley N. Gaynes, MD, MPH; Richard A. Hansen, PhD, RPh; Patricia Thieda, MA; Angela DeVeaugh-Geiss, MS; Erin E. Krebs, MD, MPH; Charity G. Moore, PhD, MSPH; Laura Morgan, MA; and Kathleen N. Lohr, PhD

Background: Second-generation antidepressants dominate the management of major depressive disorder, dysthymia, and subsyndromal depression. Evidence on the comparative benefits and harms is still accruing.

Purpose: To compare the benefits and harms of second-generation antidepressants (bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine) for the treatment of depressive disorders in adults.

Data Sources: MEDLINE, EMBASE, PsychLit, Cochrane Central Register of Controlled Trials, and International Pharmaceutical Abstracts from 1980 to April 2007, limited to English-language articles. Reference lists of pertinent review articles were manually searched and the Center for Drug Evaluation and Research database was explored to identify unpublished research.

Study Selection: Abstracts and full-text articles were independently reviewed by 2 persons. Six previous good- or fair-quality systematic reviews or meta-analyses were included, as were 155 good- or fair-quality double-blind, placebo-controlled, or head-to-head randomized, controlled trials of at least 6 weeks' duration. For harms, 35 observational studies with at least 100 participants and follow-up of at least 12 weeks were also included.

Data Extraction: Using a standard protocol, investigators abstracted data on study design and quality-related details, funding, settings, patients, and outcomes.

Data Synthesis: If data were sufficient, meta-analyses of head-to-head trials were conducted to determine the relative benefit of response to treatment and the weighted mean differences on specific depression rating scales. If sufficient evidence was not available, adjusted indirect comparisons were conducted by using meta-regressions and network meta-analyses. Second-generation antidepressants did not substantially differ in efficacy or effectiveness for the treatment of major depressive disorder on the basis of 203 studies; however, the incidence of specific adverse events and the onset of action differed. The evidence is insufficient to draw conclusions about the comparative efficacy, effectiveness, or harms of these agents for the treatment of dysthymia and subsyndromal depression.

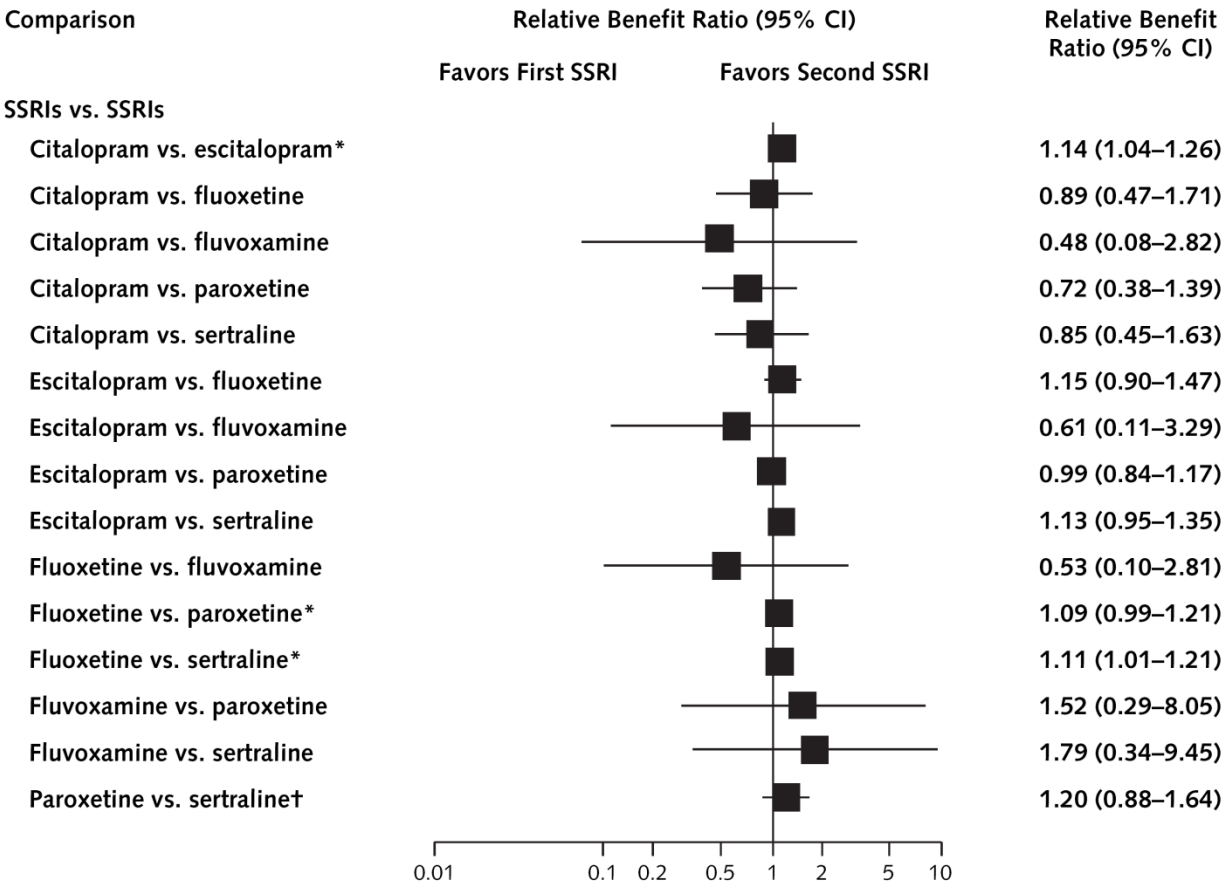
Limitation: Adjusted indirect comparisons have methodological limitations and cannot conclusively rule out differences in efficacy.

Conclusion: Current evidence does not warrant the choice of one second-generation antidepressant over another on the basis of differences in efficacy and effectiveness. Other differences with respect to onset of action and adverse events may be relevant for the choice of a medication.

Ann Intern Med. 2008;149:734-750. www.annals.org

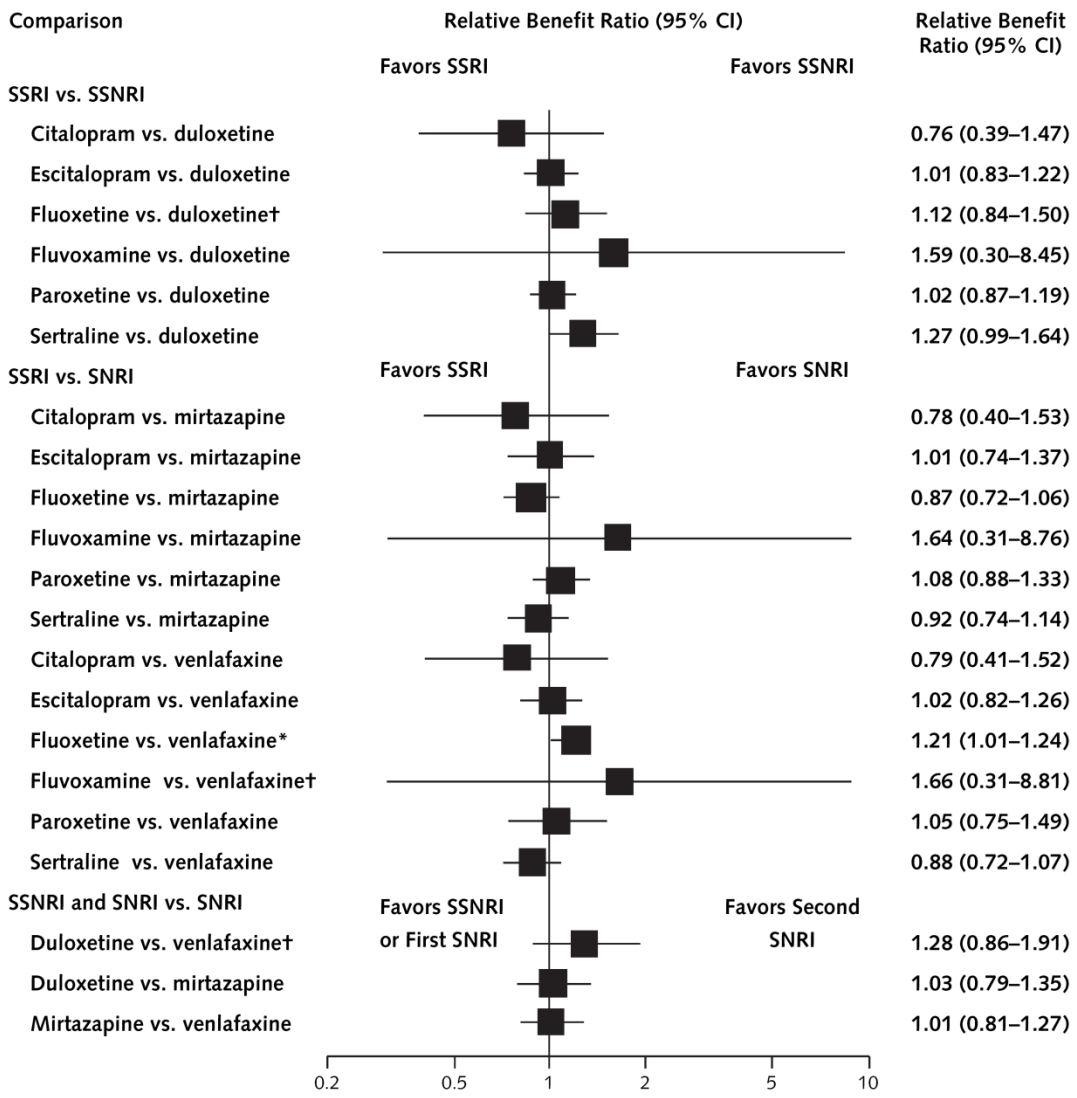
For author affiliations, see end of text.

Figure 2. **Relative benefit of response comparing selective serotonin reuptake inhibitors (*SSRIs*) with other SSRIs.**



All estimates are based on network meta-analyses except for those marked with an asterisk or a dagger.
* Based on meta-analysis of head-to-head trials.
† Based on indirect comparisons with meta-regression.

Figure 3. **Relative benefit of response comparing selective serotonin reuptake inhibitors (*SSRIs*) with selective serotonin and norepinephrine reuptake inhibitors (*SSNRIs*) and SSRIs with serotonin and norepinephrine reuptake inhibitors (SNRIs).**



All estimates are based on network meta-analyses except for those marked with an asterisk or a dagger.
* Based on meta-analysis of head-to-head trials.
† Based on indirect comparisons with meta-regression.

Consultations

- Preventing unintentional road injuries among under 15s: road design - evidence consultation
- Soft tissue sarcoma - trabectedin: appraisal consultation
- Cervical cancer (recurrent) - topotecan (appraisal consultation)
- Autism spectrum disorders in children and young people (draft scope consultation)

> See all consultations

> Read how NICE involves patients and the public

Join a committee

Attend a meeting

Welcome to the National Institute for Health and Clinical Excellence

NICE is an independent organisation responsible for providing national guidance on promoting good health and preventing and treating ill health.

NHS Evidence

NHS Evidence helps you find, access, and use high quality clinical information

Guidance latest

Guidance implementation

Quality initiatives

- Percutaneous endoscopic laser cervical discectomy
- Corneal endothelial transplantation
- Head and neck cancer (squamous cell carcinoma) - cetuximab

> See all latest guidance

Search NICE guidance

Want to know what NICE recommends?

Devices and diagnostics

NICE has been asked by the Department of Health to coordinate the evaluation of medical devices and diagnostics used in the NHS.

News

- NICE guideline helps end regional variation of access to fertility treatment
  23 Jun 2009

page 1 of 5

DEPRESSION

THE NICE GUIDELINE ON THE TREATMENT AND MANAGEMENT OF DEPRESSION IN ADULTS

UPDATED EDITION

Systematic reviews using meta-analysis suggest that antidepressant drugs, when considered individually or by class, are more effective than placebo in the treatment of major depression, and are generally equally effective (Cochrane Database of Systematic Reviews; Gartlehner *et al.*, 2008; NICE, 2004a). SSRIs are considerably safer in overdose than TCAs, are generally better tolerated than antidepressants from other classes and most are available as generic preparations. An SSRI was recommended as first-line pharmacological treatment of moderate to severe depression in the previous guideline, and SSRIs are now the most commonly prescribed group of antidepressants in the UK (see also Section 9.2).

12 new generation antidepressants

19 **meta-analyses** published in the last two years

"Although **Mirtazapine** is likely to have a faster onset of action than **Sertraline and Paroxetine** no significant differences were observed..."

"...statistically significant differences in terms of efficacy .... between **Fluoxetine and Venlafaxine**, but the clinical meaning of these differences is uncertain..."

"...**meta-analysis** highlighted a trend in favour of **Sertraline** over other **Fluoxetine**"

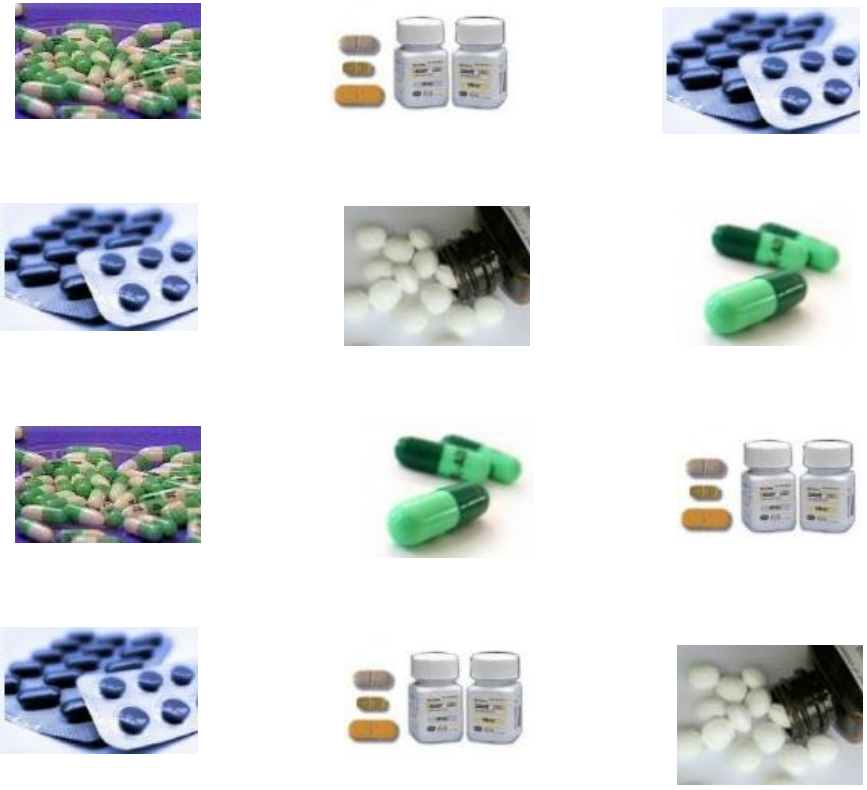
"**Venlafaxine** tends to have a favorable trend in response rates compared with **duloxetine**"

Traditional meta-analysis does not help!

	Number of studies	Number of patients	Efficacy		Acceptability	
			Response rate (responders/ total randomised)	OR (95% CI)	Dropout rate (dropouts/ total randomised)	OR (95% CI)
Bupropion *vs*						
Escitalopram	3	842	163/279 *vs* 172/287	0·93 (0·60–1·45)	105/417 *vs* 109/425	0·98 (0·72–1·34)
Fluoxetine	3	740	187/369 *vs* 206/371	0·82 (0·62–1·10)	134/369 *vs* 134/371	1·01 (0·75–1·36)
Paroxetine	2	240	34/48 *vs* 40/52	0·73 (0·30–1·79)	22/117 *vs* 26/123	0·86 (0·45–1·63)
Sertraline	3	727	237/364 *vs* 231/363	1·07 (0·79–1·45)	63/242 *vs* 82/237	0·66 (0·38–1·16)
Venlafaxine	3	1127	307/563 *vs* 329/564	0·85 (0·63–1·16)	150/563 *vs* 152/564	0·99 (0·76–1·31)
Citalopram *vs*						
Escitalopram	5	1604	319/622 *vs* 426/725	0·68 (0·53–0·87)	127/750 *vs* 141/854	1·17 (0·83–1·64)
Fluoxetine	3	740	216/364 *vs* 219/376	1·05 (0·77–1·43)	75/364 *vs* 68/376	1·17 (0·80–1·70)
Fluvoxamine	1	217	33/108 *vs* 31/109	1·11 (0·62–1·98)	22/108 *vs* 29/109	0·71 (0·37–1·33)
Mirtazapine	1	270	117/133 *vs* 116/137	1·32 (0·66–2·66)	8/133 *vs* 18/137	0·42 (0·18–1·01)
Paroxetine	1	406	77/199 *vs* 102/207	1·54 (1·04–2·28)	41/199 *vs* 43/207	1·01 (0·62–1·63)
Reboxetine	2	451	145/227 *vs* 110/224	1·72 (1·01–2·93)	51/227 *vs* 73/224	0·86 (0·22–3·46)
Sertraline	2	615	139/200 *vs* 136/200	0·93 (0·61–1·42)	60/307 *vs* 82/308	0·67 (0·46–0·98)
Venlafaxine	1	151	50/75 *vs* 49/76	1·10 (0·56–2·16)	..	..
Duloxetine *vs*						
Escitalopram	3	1120	260/562 *vs* 286/558	0·77 (0·52–1·13)	131/411 *vs* 87/414	1·93 (0·99–3·77)

We included only randomised controlled trials (RCTs) that compared any of the following 12 new-generation antidepressants as monotherapy in the acute-phase treatment of adults with unipolar major depression:

- bupropion
- citalopram
- duloxetine
- escitalopram
- fluoxetine
- fluvoxamine
- milnacipran
- mirtazapine
- paroxetine
- reboxetine
- sertraline
- venlafaxine



Search strategy

To identify the relevant studies, we reviewed the Cochrane Collaboration Depression, Anxiety, and Neurosis review group controlled trials registers (CCDANDTR-studies and CCDANCTR-references) up to Nov 30, 2007.

These registers are compiled from systematic and regularly updated searches of Cochrane Collaboration CENTRAL register,* AMED, CINAHL, EMBASE, LiLACS, MEDLINE, UK National Research Register, PSYCINFO, PSYNDEX, supplemented with hand searching of 12 conference proceedings (http://www.thecochranelibrary.com).

•The Cochrane Central Register of Controlled Trials (CENTRAL) serves as the most comprehensive source of reports of controlled trials. CENTRAL is published as part of *The Cochrane Library* and is updated quarterly. As of January 2008 (Issue 1, 2008), CENTRAL contains nearly 530,000 citations to reports of trials and other studies potentially eligible for inclusion in Cochrane reviews, of which 310,000 trial reports are from MEDLINE, 50,000 additional trial reports are from EMBASE and the remaining 170,000 are from other sources such as other databases and hand-searching.
Many of the records in CENTRAL have been identified through systematic searches of MEDLINE and EMBASE. CENTRAL, however, includes citations to reports of controlled trials that are not indexed in MEDLINE, EMBASE or other bibliographic databases; citations published in many languages; and citations that are available only in conference proceedings or other sources that are difficult to access. It also includes records from trials registers and trials results registers (full details available at http://www.cochrane-handbook.org/).

We asked pharmaceutical companies, regulatory agencies, and study investigators to supply all available information.

No language restrictions were applied.

The following drug-approving agencies - (the Food and Drug Administration in the USA, the Medicines and Healthcare products Regulatory Agency in the UK, the European Medicines Agency in the EU, the Pharmaceuticals and Medical Devices Agency in Japan, the Therapeutic Goods Administration in Australia) and ongoing trial registers (clinicaltrials.gov in the USA, ISRCTN and National Research Register in the UK, Netherlands Trial Register in the Netherlands, EUDRACT in the EU, UMIN-CTR in Japan and the Australian Clinical Trials Registry in Australia) were hand-searched for published, unpublished and ongoing controlled trials.

The following phrase was used: [*depress** or *dysthymi** or *adjustment disorder** or *mood disorder** or *affective disorder* or *affective symptoms*] and combined with a list of the 12 specific second-generation antidepressants.

Two persons within the reviewing team independently reviewed references and abstracts retrieved by the search, assessed the completeness of data abstraction, and confirmed quality rating. Any disagreements were resolved via discussion with a third member of the reviewing team. A structured data abstraction form was used to ensure consistency of appraisal for each study (see Appendix 2).

Websites of pharmaceutical companies

- Eli Lilly: www.lilly.com
- Lundbeck: www.lundbeck.com
- Organon: www.organon.com
- Solvay: www.solvay.com
- Pfizer: www.pfizer.com
- GlaxoSmithKline: www.gsk.com
- Bristol Myers Squibb: www.bms.com
- Pierre Fabre : www.pierre-fabre.com
- Wyeth: www.wyeth.com

Medical Control Agencies

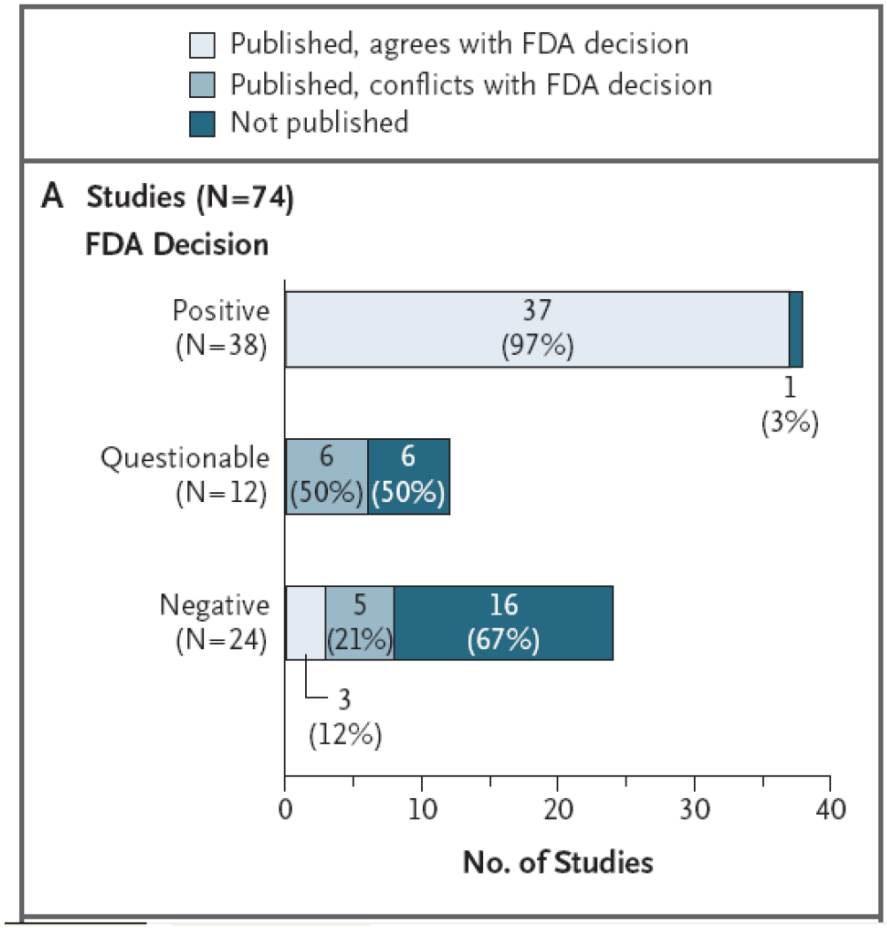
- Food and Drug Administration (USA): www.fda.gov
- European Medicines Agency (EU): www.emea.europa.eu
- Pharmaceuticals and Medical Devices Agency (Japan): www.pmda.go.jp
- Therapeutic Goods Administration (Australia): www.tga.gov.au

The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S., Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.



N Engl J Med 2008;358:252-60.

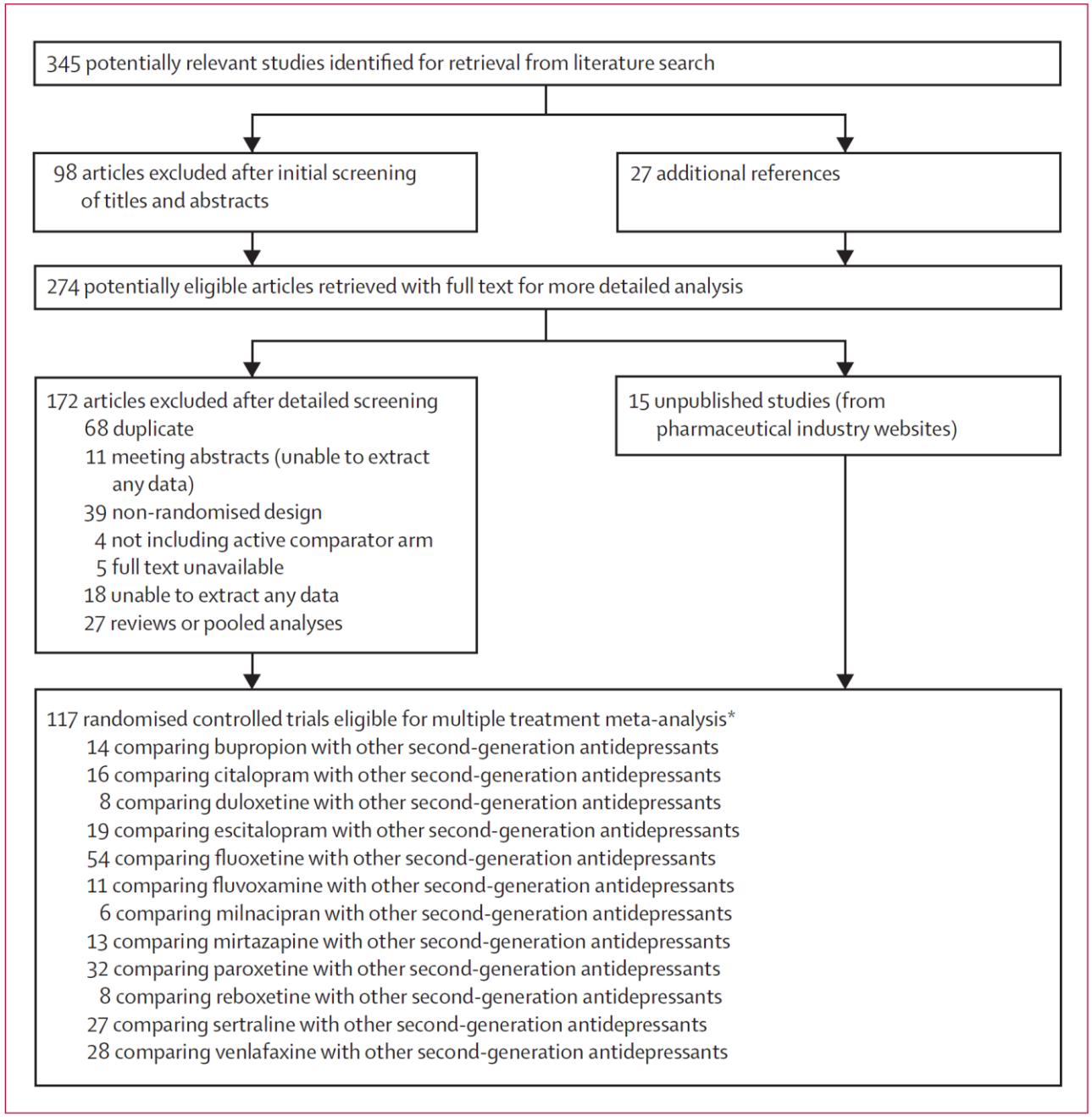


Figure 1: **Study selection process**

*117 randomised controlled trials correspond to 236 arms because two three-arm studies comparing fluoxetine with paroxetine and sertraline were included in this multiple-treatments meta-analysis.

117 RCTs

- 15 unpublished

25 928 individuals

- 24 595 (111 RCTs) EFF
- 24 693 (112) ACC
- 64% women

8·1 weeks (follow up)

109·8 (9–357) (sample)

Comparability of dosages

The dose is an important issue when comparing pharmacological treatments. We used a modified version of a published classification, described by Gartlehner and colleagues.

	Range (mg/day)	Low	Medium	High
Bupropion	150–450	<337·5	337·5–412·5	>412·5
Citalopram	20–60	<30	30–50	>50
Duloxetine	60–100	<70	70–90	>90
Escitalopram	10–30	<15	15–25	>25
Fluoxetine	20–60	<30	30–50	>50
Fluvoxamine	50–300	<75	75–125	>125
Milnacipran	50–300	<75	75–125	>125
Mirtazapine	15–45	<22·5	22·5–37·5	>37·5
Paroxetine	20–60	<30	30–50	>50
Reboxetine	4–12	<5	5–9	>9
Sertraline	50–200	<75	75–125	>125
Venlafaxine	75–250	<156·3	156·25–218·7	>218·75

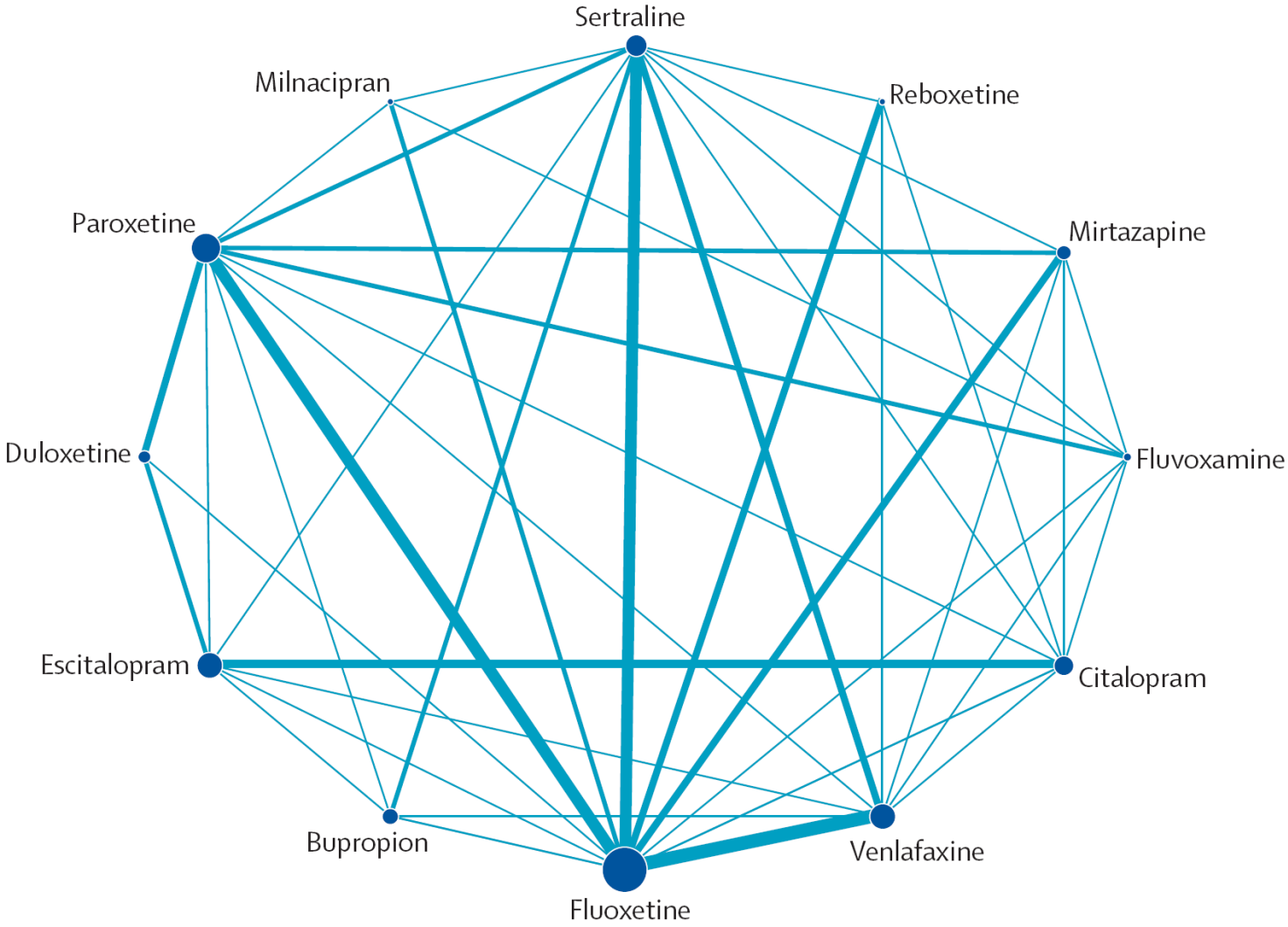


Figure 2: **Network of eligible comparisons for the multiple-treatment meta-analysis for efficacy (response rate)**
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of randomised participants (sample size). The network of eligible comparisons for acceptability (dropout rate) analysis is similar.

Outcome measures

We defined acute treatment as 8-week treatment for both efficacy and acceptability analyses (Bauer et al., 2002).

Response and dropout rates were chosen as primary outcomes, being the most consistently reported estimates of acute-treatment efficacy and acceptability.

Response: the proportion of patients who had a reduction of at least 50% from the baseline score on HDRS or MADRS

Treatment discontinuation (acceptability): the number of patients who terminated the study early for any reason during the first 8 weeks of treatment (dropouts).

BUP	1·00 (0·78–1·28)	0·75 (0·55–1·01)	1·06 (0·86–1·32)	0·89 (0·74–1·08)	0·73 (0·53–1·00)	0·87 (0·58–1·24)	0·87 (0·66–1·14)	0·81 (0·65–1·00)	**0·62 (0·45–0·86)**	1·01 (0·82–1·27)	0·84 (0·68–1·02)
0·98 (0·78–1·23)	CIT	0·75 (0·55–1·02)	1·07 (0·86–1·31)	0·90 (0·73–1·09)	**0·73 (0·54–0·99)**	0·87 (0·60–1·24)	0·87 (0·66–1·15)	0·81 (0·65–1·01)	**0·62 (0·45–0·84)**	1·02 (0·81–1·28)	0·84 (0·67–1·06)
1·09 (0·83–1·43)	1·12 (0·87–1·44)	DUL	**1·43 (1·09–1·85)**	1·19 (0·91–1·57)	0·98 (0·67–1·41)	1·16 (0·77–1·73)	1·16 (0·83–1·61)	1·08 (0·84–1·40)	0·83 (0·57–1·22)	**1·36 (1·01–1·83)**	1·12 (0·84–1·50)
0·82 (0·67–1·01)	0·84 (0·70–1·01)	**0·75 (0·60–0·93)**	ESC	0·84 (0·70–1·01)	**0·69 (0·50–0·94)**	0·81 (0·55–1·15)	0·81 (0·62–1·07)	**0·76 (0·62–0·93)**	**0·58 (0·43–0·81)**	0·95 (0·77–1·19)	**0·78 (0·64–0·97)**
1·08 (0·90–1·29)	1·10 (0·93–1·31)	0·99 (0·79–1·24)	**1·32 (1·12–1·55)**	FLU	0·82 (0·62–1·07)	0·97 (0·69–1·32)	0·97 (0·77–1·21)	0·91 (0·79–1·05)	**0·70 (0·53–0·92)**	1·14 (0·96–1·36)	0·94 (0·81–1·09)
1·10 (0·83–1·47)	1·13 (0·86–1·47)	1·01 (0·74–1·38)	**1·35 (1·02–1·76)**	1·02 (0·81–1·30)	FVX	1·18 (0·76–1·75)	1·18 (0·87–1·61)	1·10 (0·84–1·47)	0·85 (0·57–1·26)	**1·38 (1·03–1·89)**	1·14 (0·86–1·54)
1·07 (0·77–1·48)	1·09 (0·78–1·50)	0·97 (0·69–1·38)	1·30 (0·95–1·78)	0·99 (0·74–1·31)	0·97 (0·68–1·37)	MIL	0·99 (0·69–1·53)	0·94 (0·68–1·31)	0·72 (0·48–1·10)	1·17 (0·84–1·72)	0·97 (0·69–1·40)
0·79 (0·72–1·00)	0·80 (0·63–1·01)	**0·72 (0·54–0·94)**	0·96 (0·76–1·19)	**0·73 (0·60–0·88)**	**0·71 (0·55–0·92)**	0·74 (0·53–1·01)	MIR	0·93 (0·75–1·17)	0·72 (0·51–1·03)	1·17 (0·91–1·51)	0·97 (0·76–1·23)
1·06 (0·87–1·30)	1·08 (0·90–1·30)	0·97 (0·78–1·20)	**1·30 (1·10–1·53)**	0·98 (0·86–1·12)	0·96 (0·76–1·23)	1·00 (0·74–1·33)	**1·35 (1·11–1·64)**	PAR	0·77 (0·56–1·05)	**1·25 (1·04–1·52)**	1·03 (0·86–1·24)
1·60 (1·20–2·16)	**1·63 (1·25–2·14)**	**1·46 (1·05–2·02)**	**1·95 (1·47–2·59)**	**1·48 (1·16–1·90)**	**1·45 (1·03–2·02)**	**1·50 (1·03–2·18)**	**2·03 (1·52–2·78)**	**1·50 (1·16–1·98)**	REB	**1·63 (1·19–2·24)**	1·34 (0·99–1·83)
0·87 (0·72–1·05)	0·88 (0·72–1·07)	0·79 (0·62–1·01)	1·06 (0·88–1·27)	**0·80 (0·69–0·93)**	0·79 (0·61–1·01)	0·81 (0·60–1·11)	1·10 (0·90–1·36)	**0·82 (0·69–0·96)**	**0·54 (0·41–0·71)**	SER	0·82 (0·67–1·00)
0·85 (0·70–1·01)	0·86 (0·71–1·05)	**0·77 (0·60–0·99)**	1·03 (0·86–1·24)	**0·78 (0·68–0·90)**	**0·77 (0·59–0·99)**	0·79 (0·58–1·08)	1·08 (0·87–1·33)	**0·79 (0·67–0·94)**	**0·53 (0·40–0·69)**	0·98 (0·82–1·16)	VEN

***Figure 3*: Efficacy and acceptability of the 12 antidepressants**

Drugs are reported in alphabetical order. Results are the ORs in the column-defining treatment compared with the ORs in the row-defining treatment. For efficacy, ORs higher than 1 favour the column-defining treatment (ie, the first in alphabetical order). For acceptability, ORs lower than 1 favour the first drug in alphabetical order. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken (eg, the OR for FLU compared with CIT is 1/1·10=0·91). Significant results are in bold and underscored. BUP=bupropion. CIT=citalopram. DUL=duloxetine. ESC=escitalopram. FLU=fluoxetine. FVX=fluvoxamine. MIL=milnacipran. MIR=mirtazapine. PAR=paroxetine. REB=reboxetine. SER=sertraline. VEN=venlafaxine. MTM=multiple-treatments meta-analysis. OR=Odds ratio. CI=credibility interval.

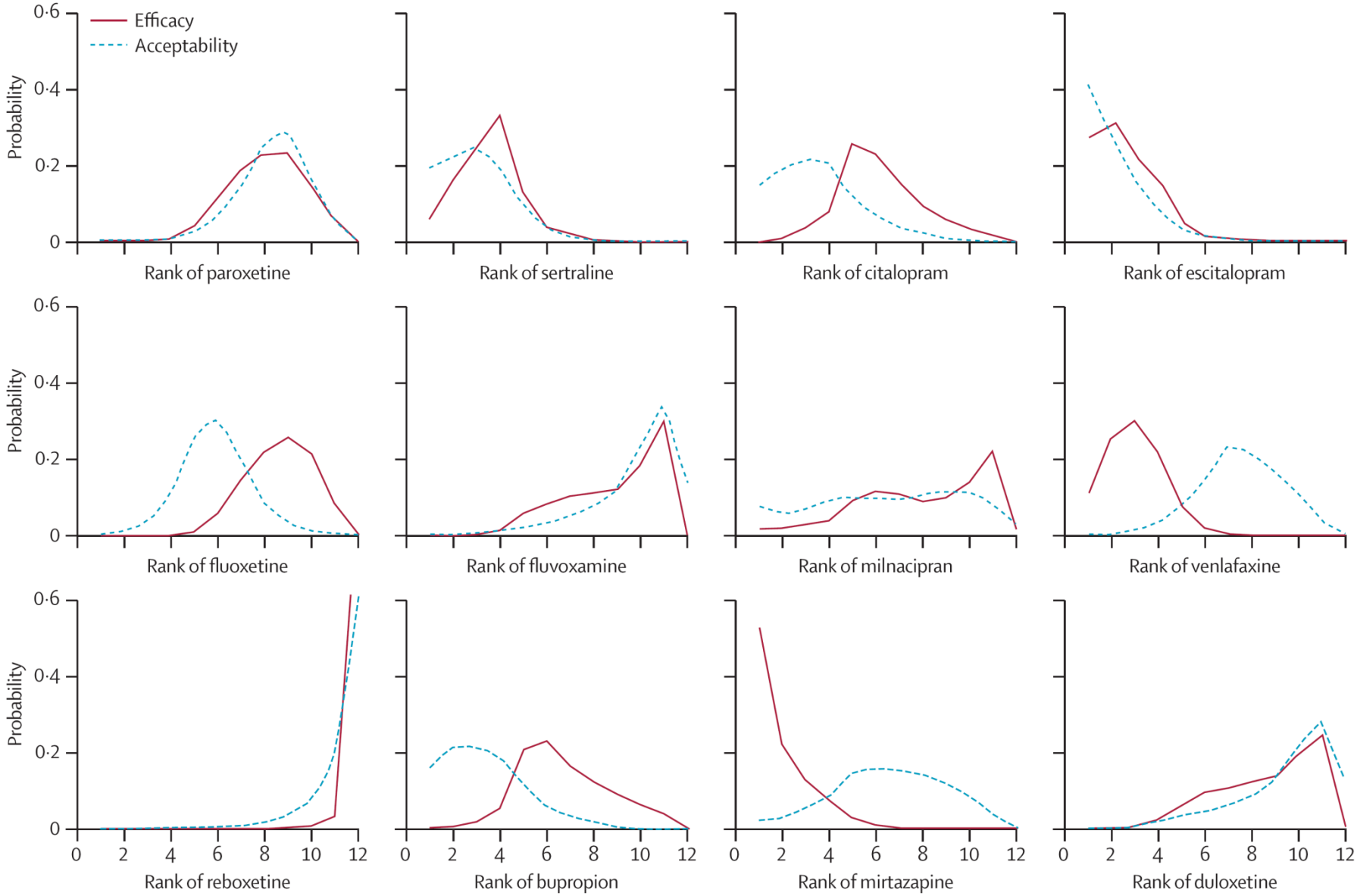


Figure 4: **Ranking for efficacy (solid line) and acceptability (dotted line)**
Ranking indicates the probability to be the best treatment, the second best, the third best, and so on, among the 12 antidepressants.

The cumulative probabilities of being among the four most efficacious treatments

	Efficacy	
Rank	**Drug**	***%***
1.	Mirtazapine	*24·4*
2.	Escitalopram	*23·7*
3.	Venlafaxine	*22·3*
4.	Sertraline	*20·3*
5.	Citalopram	*3·4*
6.	Milnacipran	*2·7*
7.	Bupropion	*2·0*
8.	Duloxetine	*0·9*
9.	Fluvoxamine	*0·7*
10.	Paroxetine	*0·1*
11.	Fluoxetine	*0·0*
12.	Reboxetine	*0·0*

The cumulative probabilities of being among the four most efficacious treatments and among the four best treatments in terms of acceptability

	Efficacy		Acceptability	
Rank	**Drug**	**%**	**Drug**	**%**
1.	Mirtazapine	*24·4*	Escitalopram	*27·6*
2.	Escitalopram	*23·7*	Sertraline	*21·3*
3.	Venlafaxine	*22·3*	Bupropion	*19·3*
4.	Sertraline	*20·3*	Citalopram	*18·7*
5.	Citalopram	*3·4*	Milnacipran	*7·1*
6.	Milnacipran	*2·7*	Mirtazapine	*4·4*
7.	Bupropion	*2·0*	Fluoxetine	*3·4*
8.	Duloxetine	*0·9*	Venlafaxine	*0·9*
9.	Fluvoxamine	*0·7*	Duloxetine	*0·7*
10.	Paroxetine	*0·1*	Fluvoxamine	*0·4*
11.	Fluoxetine	*0·0*	Paroxetine	*0·2*
12.	Reboxetine	*0·0*	Reboxetine	*0·1*

BMJ

RESEARCH

Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials

Dirk Eyding, project manager,[1] Monika Lelgemann, senior researcher,[2] Ulrich Grouven, statistician,[3,4] Martin Härter, head of department of medical psychology,[5] Mandy Kromp, statistician,[3] Thomas Kaiser, head of department of drug assessment,[3] Michaela F Kerekes, data manager,[3] Martin Gerken, researcher,[6] Beate Wieseler, deputy head of department of drug assessment[3]

[1]German Cancer Society, Berlin, Germany (former affiliation: Institute for Quality and Efficiency in Health Care)

[2]Medical Advisory Service of Social Health Insurance, Essen, Germany

[3]Institute for Quality and Efficiency in Health Care, Cologne, Germany

[4]Hanover Medical School, Hanover, Germany

[5]University Clinic Eppendorf, Hamburg, Germany

[6]Health Technology Assessment Centre, University of Bremen, Bremen, Germany

Correspondence to: B Wieseler, Institute for Quality and Efficiency in Health Care, Dillenburger Strasse 27, 51105 Cologne, Germany beate.wieseler@iqwig.de

Cite this as: *BMJ* 2010;341:c4737
doi:10.1136/bmj.c4737

ABSTRACT

Objectives To assess the benefits and harms of reboxetine versus placebo or selective serotonin reuptake inhibitors (SSRIs) in the acute treatment of depression, and to measure the impact of potential publication bias in trials of reboxetine.

Design Systematic review and meta-analysis including unpublished data.

Data sources Bibliographic databases (Medline, Embase, PsycINFO, BIOSIS, and Cochrane Library), clinical trial registries, trial results databases, and regulatory

difference in response rates between patients receiving reboxetine and those receiving placebo (OR 1.24, 95% CI 0.98 to 1.56; P=0.071; I^2=42.1%). Reboxetine was inferior to SSRIs (fluoxetine, paroxetine, and citalopram) for remission rates (OR 0.80, 95% CI 0.67 to 0.96; P=0.015) and response rates (OR 0.80, 95% CI 0.67 to 0.95; P=0.01). Reboxetine was inferior to placebo for both harm outcomes (P<0.001 for both), and to fluoxetine for withdrawals owing to adverse events (OR 1.79, 95% CI 1.06 to 3.05; P=0.031). Published data overestimated the benefit of reboxetine versus placebo by up to 115% and

IQWiG Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen

Selektive Serotonin- und Noradrenalin-Wiederaufnahmehemmer (SNRI) bei Patienten mit Depressionen

Abschlussbericht

Auftrag A05-20A
Version 1.1
Stand: 18.08.2010

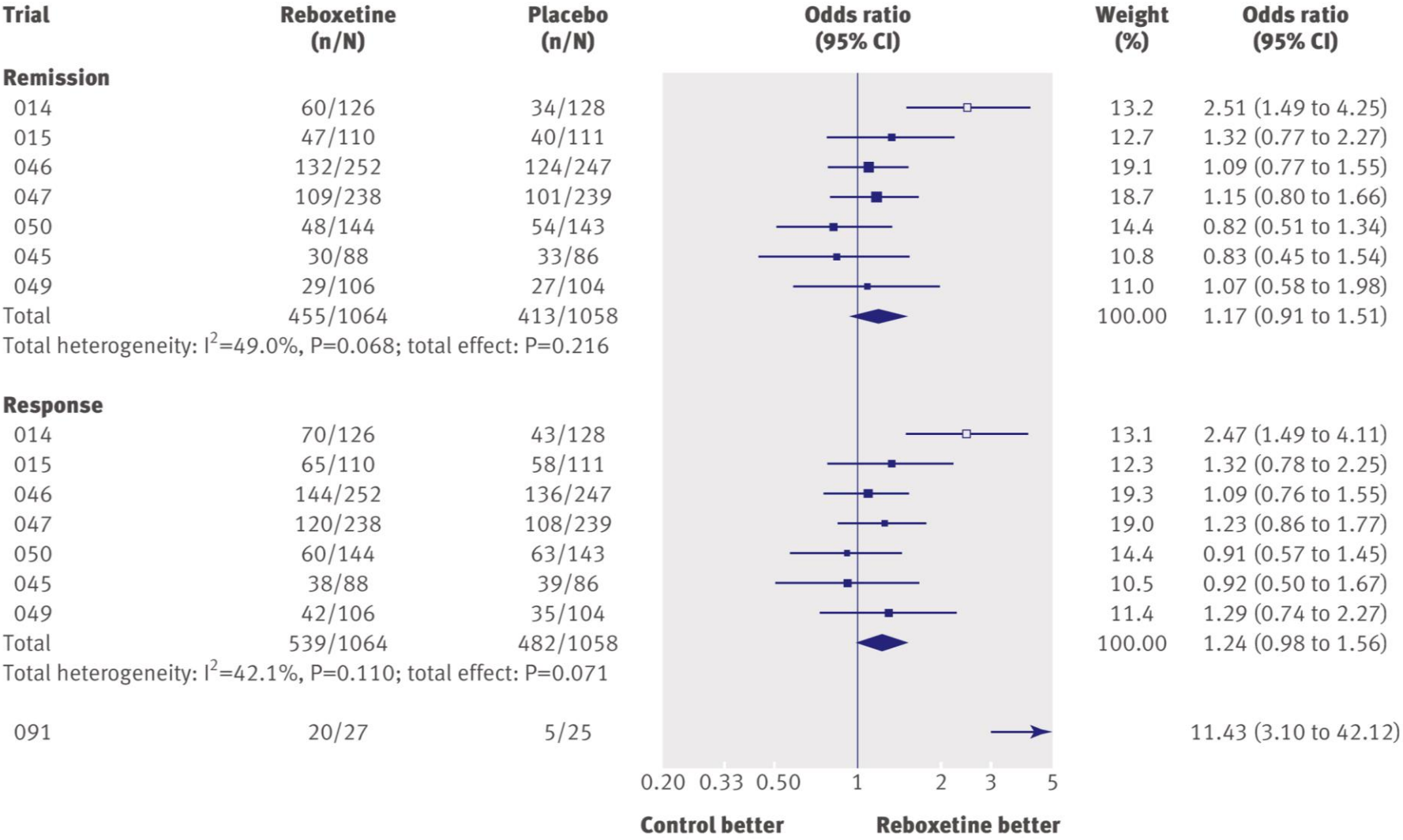


Fig 2 | Forest plot showing meta-analyses of remission and response rates for trials that compared reboxetine with placebo. Empty boxes show published studies and filled boxes show unpublished studies. Study 091 is not included in the pooled analysis of response of reboxetine versus placebo because of high heterogeneity (see text for details). CI, confidence interval; n, number of patients with event; N, number of patients in treatment group

Adjusting for sponsorship bias

y_i^{AB} the outcome of experiment A vs B

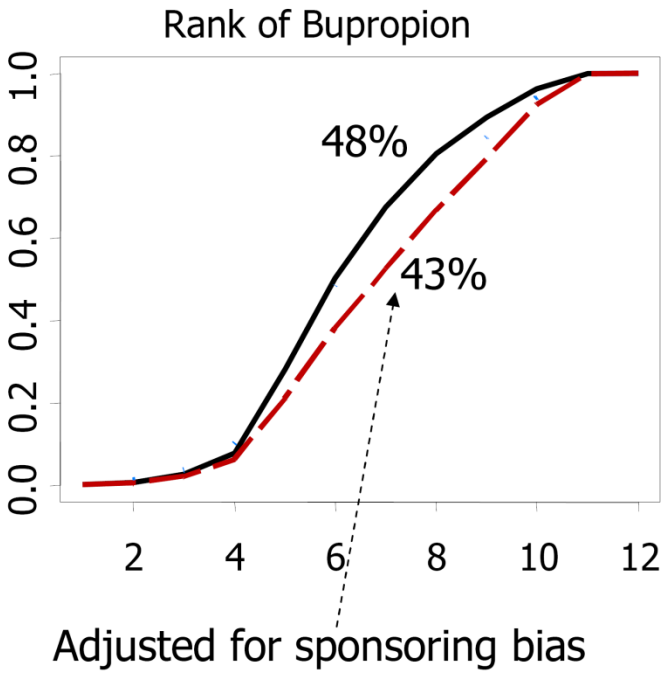
Likelihood: $y_i^{AB} \sim N(\theta_i^{AB}, (var_i^{AB})^2)$

Bias adjusted estimate $\theta_i^{AB} = \delta_i^{AB} + \beta_i I^{AB}$

$\beta_i \sim N(B, \sigma^2)$

$\delta_i^{AB} \sim N(\mu^{AB}, \tau^2)$

Index, (0 or 1) depending on whether A is sponsored (and favored) compared to B



Result: No effect of sponsoring!
B was practically zero and the change in effectiveness was negligible

Lancet 2009 *Cipriani, Fukurawa, Salanti et al*

Meta-regression

In a meta-regression analysis to assess potential sponsorship bias, ORs and final rankings did not substantially change. The cumulative probability of being among the four best treatments became slightly smaller for those drugs in trials which were sponsored by the marketing company, with the comparators moving up the ranking slightly.

SPONSORSHIP					
	Rank	**Drug**	***%***	**N of RCTs sponsored by drug manufacturer (or unclear) [%]**	**Total N of RCTs included in the MTM**
Efficacy	1.	Mirtazapine	*24·4*	13 [100%]	*13*
	2.	Escitalopram	*23·7*	16 [84.2%]	*19*
	3.	Venlafaxine	*22·3*	17 [60.7%]	*28*
	4.	Sertraline	*20·3*	12 [44.4%]	*27*

Meta-regression

In a meta-regression analysis to assess potential sponsorship bias, ORs and final rankings did not substantially change. The cumulative probability of being among the four best treatments became slightly smaller for those drugs in trials which were sponsored by the marketing company, with the comparators moving up the ranking slightly.

SPONSORSHIP					
	Rank	**Drug**	***%***	**N of RCTs sponsored by drug manufacturer (or unclear) [%]**	**Total N of RCTs included in the MTM**
Efficacy	1.	Mirtazapine	*24·4*	13 [100%]	*13*
	2.	Escitalopram	*23·7*	16 [84.2%]	*19*
	3.	Venlafaxine	*22·3*	17 [60.7%]	*28*
	4.	Sertraline	*20·3*	12 [44.4%]	*27*
Acceptability	1.	Escitalopram	*24·4*	16 [84.2%]	*19*
	2.	Sertraline	*23·7*	12 [44.4%]	*27*
	3.	Bupropion	*22·3*	13 [92.8%]	*14*
	4.	Citalopram	*20·3*	11 [68.7%] (5 RCTs are CIT vs ESC)	*16*

Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis

Andrea Cipriani, Toshiaki A Furukawa, Georgia Salanti, John R Geddes, Julian P T Higgins, Rachel Churchill, Norio Watanabe, Atsuo Nakagawa, Ichiro M Omori, Hugh McGuire, Michele Tansella, Corrado Barbui

Summary

Background Conventional meta-analyses have shown inconsistent results for efficacy of second-generation antidepressants. We therefore did a multiple-treatments meta-analysis, which accounts for both direct and indirect comparisons, to assess the effects of 12 new-generation antidepressants on major depression.

Methods We systematically reviewed 117 randomised controlled trials (25 928 participants) from 1991 up to Nov 30, 2007, which compared any of the following antidepressants at therapeutic dose range for the acute treatment of unipolar major depression in adults: bupropion, citalopram, duloxetine, escitalopram, fluoxetine, fluvoxamine, milnacipran, mirtazapine, paroxetine, reboxetine, sertraline, and venlafaxine. The main outcomes were the proportion of patients who responded to or dropped out of the allocated treatment. Analysis was done on an intention-to-treat basis.

Findings Mirtazapine, escitalopram, venlafaxine, and sertraline were significantly more efficacious than duloxetine (odds ratios [OR] 1·39, 1·33, 1·30 and 1·27, respectively), fluoxetine (1·37, 1·32, 1·28, and 1·25, respectively), fluvoxamine (1·41, 1·35, 1·30, and 1·27, respectively), paroxetine (1·35, 1·30, 1·27, and 1·22, respectively), and reboxetine (2·03, 1·95, 1·89, and 1·85, respectively). Reboxetine was significantly less efficacious than all the other antidepressants tested. Escitalopram and sertraline showed the best profile of acceptability, leading to significantly fewer discontinuations than did duloxetine, fluvoxamine, paroxetine, reboxetine, and venlafaxine.

Interpretation Clinically important differences exist between commonly prescribed antidepressants for both efficacy and acceptability in favour of escitalopram and sertraline. Sertraline might be the best choice when starting treatment for moderate to severe major depression in adults because it has the most favourable balance between benefits, acceptability, and acquisition cost.

Funding None.

Correspondence to:
Dr Andrea Cipriani, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, Policlinico "G B Rossi", Piazzale L A Scuro, 10, 37134, Verona, Italy
andrea.cipriani@univr.it

THE LANCET

Volume 373 · Number 9665 · Pages 693–780 · February 28–March 6, 2009 www.thelancet.com

"Sertraline might be the best choice when starting treatment for moderate to severe major depression in adults because it has the most favourable balance between benefits, acceptability, and acquisition cost."

See Articles page 746

£5.00 Registered as a newspaper · ISSN 0140-6736
Founded 1823 · Published weekly

Lancet 2009; 373: 746–58

Treatment of mania

Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis

Andrea Cipriani, Corrado Barbui, Georgia Salanti, Jennifer Rendell, Rachel Brown, Sarah Stockton, Marianna Purgato, Loukia M Spineli, Guy M Goodwin, John R Geddes

Department of Public Health and Community Medicine, Section of Psychiatry and Clinical Psychology, University of Verona, Verona, Italy (A Cipriani PhD, Prof C Barbui MD, M Purgato PsyD); **Department of Hygiene and Epidemiology University of Ioannina School of Medicine, Ioannina, Greece** (Prof G Salanti PhD, L M Spineli MSc); **and Department of Psychiatry, University of Oxford, Oxford, UK** (J Rendell PhD, S Stockton BA (Hons), R Brown BPharm, Prof G M Goodwin FMedSci, Prof J R Geddes MD)

Correspondence to: Dr Andrea Cipriani, WHO Collaborating Centre for Research and Training in Mental Health and Service Evaluation, Policlinico "G.B. Rossi", Piazzale LA Scuro, 10, 37134 Verona, Italy **andrea.cipriani@univr.it**

Lancet 2011; 378: 1306–15

Summary

Background Conventional meta-analyses have shown inconsistent results for efficacy of pharmacological treatments for acute mania. We did a multiple-treatments meta-analysis, which accounted for both direct and indirect comparisons, to assess the effects of all antimanic drugs.

Methods We systematically reviewed 68 randomised controlled trials (16 073 participants) from Jan 1, 1980, to Nov 25, 2010, which compared any of the following pharmacological drugs at therapeutic dose range for the treatment of acute mania in adults: aripiprazole, asenapine, carbamazepine, valproate, gabapentin, haloperidol, lamotrigine, lithium, olanzapine, quetiapine, risperidone, topiramate, and ziprasidone. The main outcomes were the mean change on mania rating scales and the number of patients who dropped out of the allocated treatment at 3 weeks. Analysis was done by intention to treat.

Findings Haloperidol (standardised mean difference [SMD] –0·56 [95% CI –0·69 to –0·43]), risperidone (–0·50 [–0·63 to –0·38]), olanzapine (–0·43 [–0·54 to –0·32], lithium (–0·37 [–0·63 to –0·11]), quetiapine (–0·37 [–0·51 to –0·23]), aripiprazole (–0·37 [–0·51 to –0·23]), carbamazepine (–0·36 [–0·60 to –0·11], asenapine (–0·30 [–0·53 to –0·07]), valproate (–0·20 [–0·37 to –0·04]), and ziprasidone (–0·20 [–0·37 to –0·03]) were significantly more effective than placebo, whereas gabapentin, lamotrigine, and topiramate were not. Haloperidol had the highest number of significant differences and was significantly more effective than lithium (SMD –0·19 [95% CI –0·36 to –0·01]), quetiapine (–0·19 [–0·37 to 0·01]), aripiprazole (–0·19 [–0·36 to –0·02]), carbamazepine (–0·20 [–0·36 to –0·01]), asenapine (–0·26 [–0·52 to 0·01]), valproate (–0·36 [–0·56 to –0·15]), ziprasidone (–0·36 [–0·56 to –0·15]), lamotrigine (–0·48 [–0·77 to –0·19]), topiramate (–0·63 [–0·84 to –0·43]), and gabapentin (–0·88 [–1·40 to –0·36]). Risperidone and olanzapine had a very similar profile of comparative efficacy, being more effective than valproate, ziprasidone, lamotrigine, topiramate, and gabapentin. Olanzapine, risperidone, and quetiapine led to significantly fewer discontinuations than did lithium, lamotrigine, placebo, topiramate, and gabapentin.

Interpretation Overall, antipsychotic drugs were significantly more effective than mood stabilisers. Risperidone, olanzapine, and haloperidol should be considered as among the best of the available options for the treatment of manic episodes. These results should be considered in the development of clinical practice guidelines.

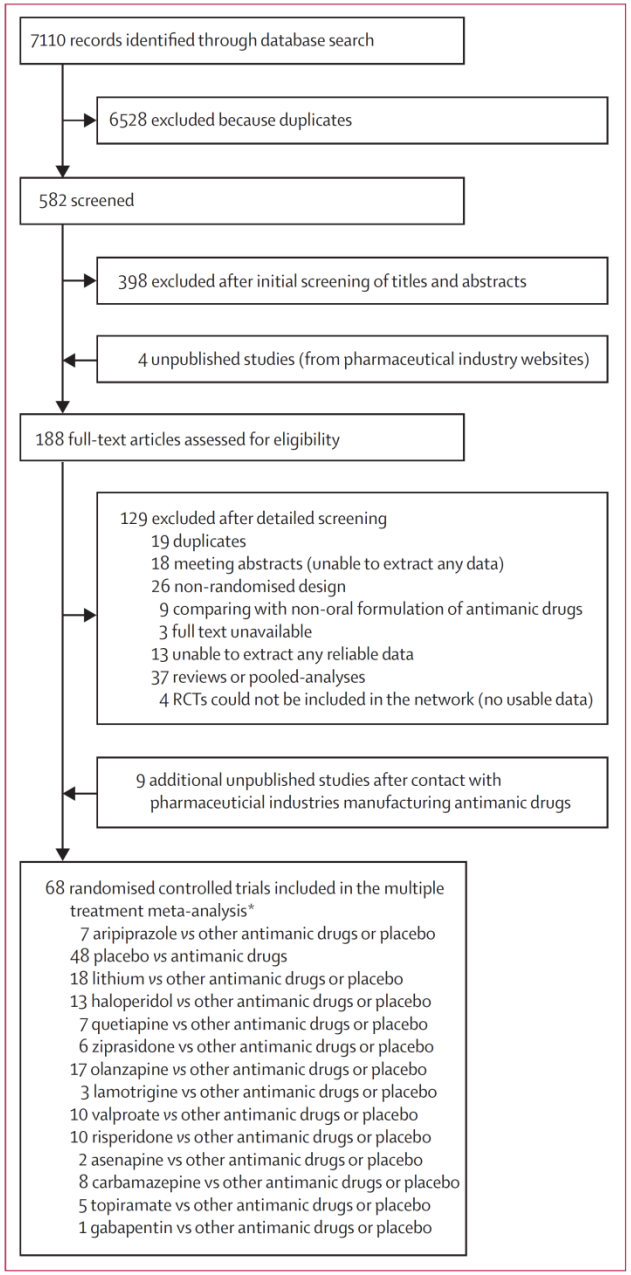


Figure 1: **Included and excluded studies**
*68 randomised trials correspond to 155 groups because three-group or four-group studies were included in this multiple-treatments meta-analysis.

14 treatments were analysed: aripiprazole, asenapine, carbamazepine, valproate, gabapentin, haloperidol, lamotrigine, lithium, olanzapine, paliperidone, quetiapine, risperidone, topiramate, ziprasidone, and placebo.

Most trials (54 [79%] of 68) were two-grouped studies and the rest were three-grouped studies in which one active comparator was usually haloperidol.
17 trials had a combination design, in which the antimanic drugs of interest were added to lithium or valproate. Of these trials, only one was a three-grouped study and the remaining 16 were two-grouped.

Overall, 16 073 patients were randomly assigned to one of the 14 antimanic treatments or to placebo and were included in the multiple-treatments meta-analysis.

	Number of studies	Overall number of patients	Efficacy		Acceptability
			Standardised mean difference (95% CI)	Response rate OR (95% CI)	Dropout rate OR (95% CI)
Aripiprazole *vs*					
Haloperidol	2	679	0·05 (–0·10 to 0·20)	1·16 (0·76 to 1·77)	0·58 (0·25 to 1·35)
Lithium	1	315	–0·06 (–0·28 to 0·16)	1·09 (0·70 to 1·70)	1·07 (0·69 to 1·66)
Placebo	6	1959	–0·31 (–0·42 to –0·20)	1·75 (1·37 to 2·24)	0·86 (0·62 to 1·19)
Asenapine *vs*					
Olanzapine	2	774	0·22 (0·08 to 0·37)	0·68 (0·46 to 1·03)	2·04 (1·49 to 2·86)
Placebo	2	582	–0·42 (–0·59 to –0·24)	2·04 (1·20 to 3·45)	0·80 (0·56 to 1·14)
Carbamazepine *vs*					
Valproate	1	30	0·85 (0·10 to 1·60)	0·41 (0·09 to 1·92)	1·00 (0·16 to 5·88)
Haloperidol	3	70	–0·09 (–0·56 to 0·38)	0·80 (0·12 to 5·56)	0·81 (0·06 to 10·00)
Lithium	2	67	0·23 (–0·30 to 0·76)	..	0·81 (0·08 to 8·33)
Placebo	1	443	–0·50 (–0·69 to –0·30)	3·12 (2·08 to 4·76)	0·71 (0·49 to 1·04)
Gabapentin *vs*					
Placebo	1	118	0·32 (–0·08 to 0·72)	..	1·75 (0·83 to 3·70)
Haloperidol *vs*					
Aripiprazole	2	679	–0·05 (–0·20 to 0·10)	0·86 (0·56 to 1·32)	1·72 (0·74 to 4·00)
Carbamazepine	3	70	0·09 (–0·38 to 0·56)	1·25 (0·18 to 8·44)	1·23 (0·10 to 15·43)
Lithium	2	44	–1·11 (–1·89 to –0·33)	..	0·98 (0·09 to 11·11)
Olanzapine	2	578	–0·15 (–0·32 to 0·03)	1·14 (0·76 to 1·70)	1·86 (0·81 to 4·30)
Placebo	6	1285	–0·58 (–0·77 to –0·39)	2·27(1·54 to 3·33)	0·72 (0·50 to 1·06)
Quetiapine	1	201	–0·42 (–0·71 to –0·14)	1·71 (0·98 to 3·00)	0·52 (0·28 to 0·98)
Risperidone	3	433	0·02 (–0·17 to 0·21)	0·95 (0·60 to 1·51)	1·36 (0·72 to 2·57)
Ziprasidone	1	350	–0·51 (–0·72 to –0·29)	2·05 (1·33 to 3·14)	0·83 (0·55 to 1·28)
Lamotrigine *vs*					
Lithium	3	303	0·21 (–0·02 to 0·50)	0·76 (0·18 to 3·23)	1·01 (0·26 to 3·85)
Placebo	2	331	0·01 (–0·21 to 0·22)	..	1·25 (0·81 to 1·96)

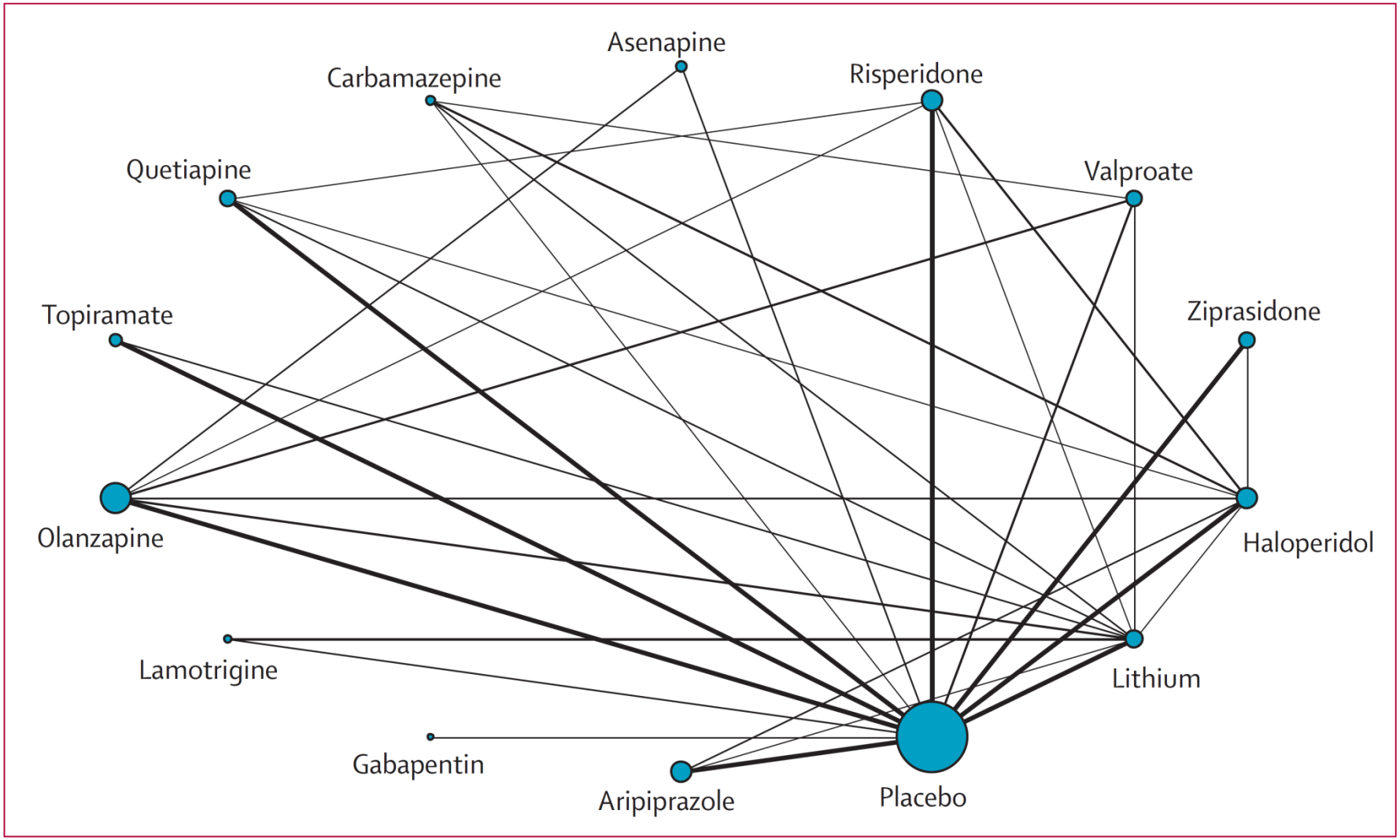


Figure 2: **Network of eligible comparisons for the multiple-treatments meta-analysis for efficacy**
The width of the lines is proportional to the number of trials comparing every pair of treatments, and the size of every node is proportional to the number of randomised participants (sample size). The networks of eligible comparisons for acceptability analysis dropout rate) and for efficacy as binary outcome are similar (webappendix pp 26–27).

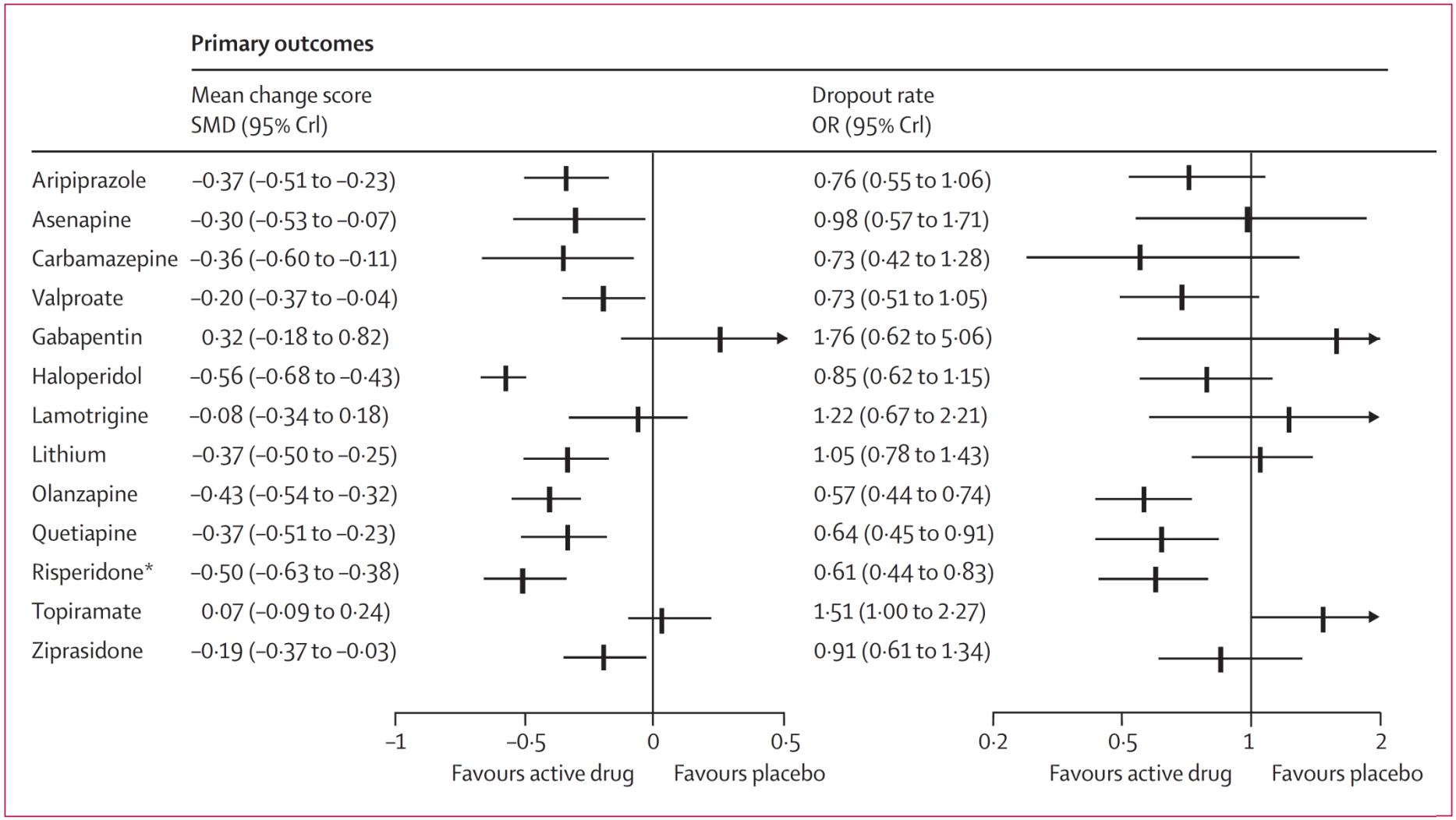


Figure 3: **Forest plots of MTM results for efficacy outcomes and dropout rate with placebo as reference compound**

Standardised mean differences lower than 0 and ORs lower than 1 favour active compound. *As stated in the protocol, data from risperidone and paliperidone were merged. MTM= multiple-treatments meta-analysis. OR=odds ratio. CrI=credibilty interval.

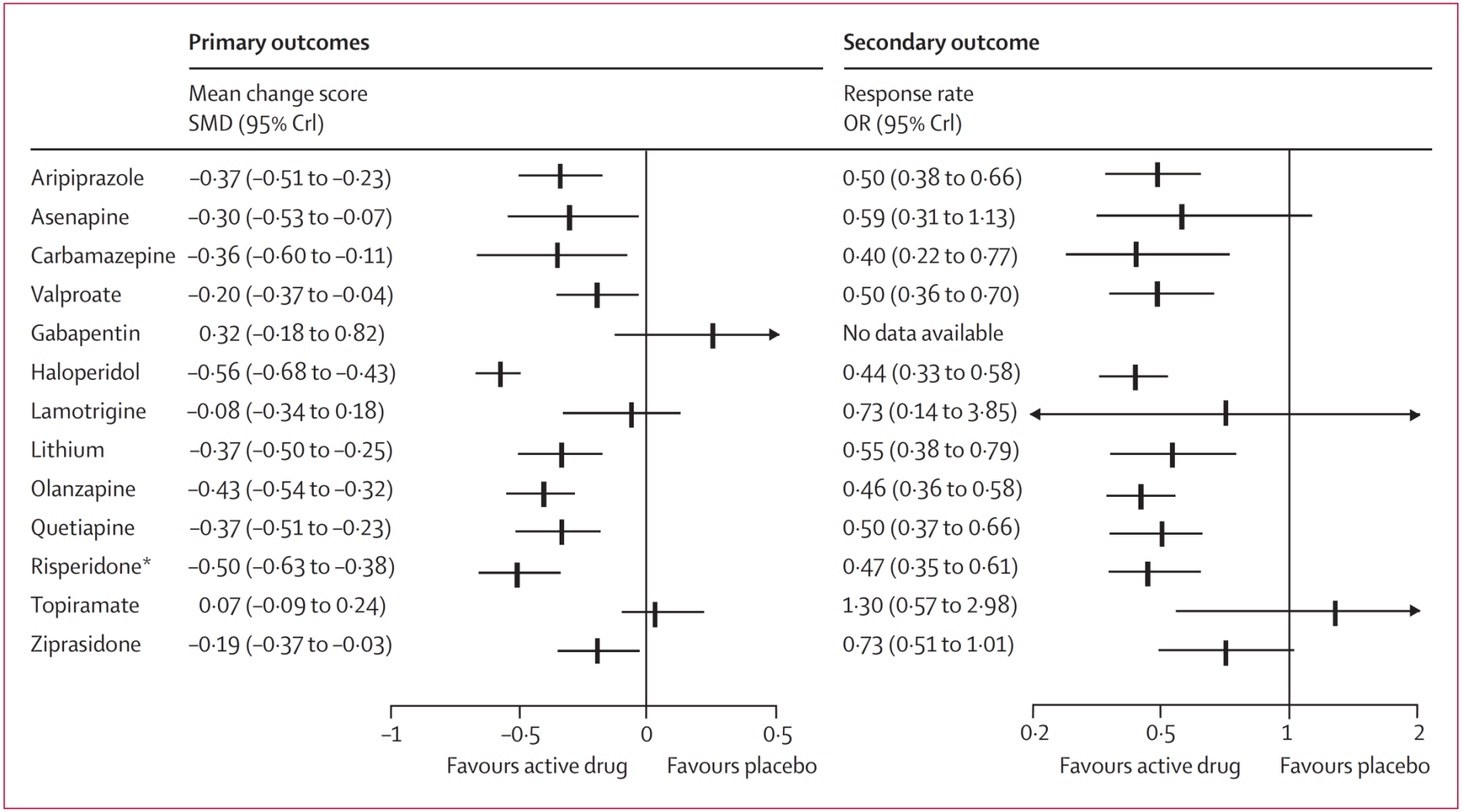


Figure 3: **Forest plots of MTM results for efficacy outcomes and dropout rate with placebo as reference compound**

Standardised mean differences lower than 0 and ORs lower than 1 favour active compound. *As stated in the protocol, data from risperidone and paliperidone were merged. MTM= multiple-treatments meta-analysis. OR=odds ratio. CrI=credibilty interval.

HAL	1·40 (0·93 to 2·11)	**1·49 (1·03 to 2·15)**	0·81 (0·53 to 1·22)	1·32 (0·85 to 2·06)	1·11 (0·75 to 1·66)	1·16 (0·63 to 2·14)	0·86 (0·46 to 1·60)	1·16 (0·73 to 1·86)	0·93 (0·59 to 1·49)	0·69 (0·36 to 1·36)	0·85 (0·62 to 1·15)	**0·56 (0·34 to 0·93)**	0·48 (0·16 to 1·44)
−0·06 (−0·22 to 0·11)	RIS	1·06 (0·72 to 1·56)	**0·58 (0·37 to 0·88)**	0·94 (0·60 to 1·47)	0·80 (0·51 to 1·25)	0·83 (0·44 to 1·57)	0·62 (0·33 to 1·16)	0·83 (0·51 to 1·34)	0·67 (0·41 to 1·10)	**0·50 (0·25 to 0·98)**	**0·61 (0·44 to 0·83)**	**0·40 (0·24 to 0·68)**	0·34 (0·11 to 1·03)
−0·12 (−0·28 to 0·02)	−0·07 (−0·22 to 0·08)	OLZ	**0·54 (0·37 to 0·79)**	0·88 (0·58 to 1·36)	0·75 (0·49 to 1·13)	0·78 (0·43 to 1·44)	0·58 (0·33 to 1·00)	0·78 (0·52 to 1·17)	0·63 (0·40 to 1·00)	**0·47 (0·24 to 0·89)**	**0·57 (0·44 to 0·74)**	**0·38 (0·23 to 0·61)**	**0·32 (0·11 to 0·95)**
−0·19 (−0·36 to −0·01)	−0·13 (−0·30 to 0·04)	−0·06 (−0·22 to 0·10)	LIT	**1·63 (1·06 to 2·54)**	1·38 (0·91 to 2·12)	1·44 (0·81 to 2·60)	1·07 (0·57 to 2·00)	1·44 (0·92 to 2·28)	1·15 (0·71 to 1·91)	0·86 (0·47 to 1·59)	1·05 (0·78 to 1·43)	0·70 (0·44 to 1·11)	0·60 (0·20 to 1·77)
−0·19 (−0·37 to −0·01)	−0·13 (−0·31 to 0·04)	−0·07 (−0·24 to 0·11)	−0·01 (−0·18 to 0·17)	QTP	0·85 (0·52 to 1·35)	0·88 (0·46 to 1·70)	0·66 (0·34 to 1·25)	0·88 (0·53 to 1·46)	0·71 (0·42 to 1·20)	0·53 (0·27 to 1·05)	**0·64 (0·45 to 0·91)**	**0·43 (0·25 to 0·73)**	0·36 (0·12 to 1·10)
−0·19 (−0·36 to −0·02)	−0·13 (−0·31 to 0·05)	−0·06 (−0·23 to 0·11)	−0·01 (−0·18 to 0·17)	0·00 (−0·19 to 0·20)	ARI	1·04 (0·55 to 1·98)	0·77 (0·41 to 1·47)	1·05 (0·64 to 1·70)	0·84 (0·51 to 1·39	0·62 (0·32 to 1·24)	0·76 (0·55 to 1·06)	**0·50 (0·30 to 0·85)**	0·43 (0·14 to 1·29)
−0·20 (−0·36 to −0·01)	−0·14 (−0·42 to 0·12)	−0·08 (−0·34 to 0·18)	−0·02 (−0·28 to 0·24)	−0·01 (−0·30 to 0·26)	−0·01 (−0·29 to 0·26)	CBZ	0·74 (0·34 to 1·62)	1·00 (0·52 to 1·91)	0·80 (0·41 to 1·59)	0·60 (0·27 to 1·33)	0·73 (0·42 to 1·28)	**0·48 (0·25 to 0·96)**	0·41 (0·13 to 1·37)
−0·26 (−0·52 to −0·01)	−0·20 (−0·46 to 0·05)	−0·14 (−0·36 to 0·10)	−0·08 (−0·41 to 0·27)	−0·07 (−0·34 to 0·20)	−0·07 (−0·34 to 0·20)	−0·06 (−0·39 to 0·28)	ASE	1·35 (0·71 to 2·58)	1·08 (0·56 to 2·14)	0·81 (0·36 to 1·83)	0·98 (0·57 to 1·72)	0·65 (0·33 to 1·30)	0·56 (0·17 to 1·82)
−0·36 (−0·56 to −0·15)	**−0·30 (−0·50 to −0·10)**	**−0·23 (−0·40 to −0·06)**	−0·10 (−0·41 to 0·23)	−0·17 (−0·38 to 0·05)	−0·17 (−0·38 to 0·05)	−0·15 (−0·44 to 0·13)	−0·10 (−0·37 to 0·18)	VAL	0·80 (0·47 to 1·37)	0·60 (0·30 to 1·20)	0·73 (0·51 to 1·05)	**0·48 (0·28 to 0·83)**	0·41 (0·13 to 1·25)
−0·36 (−0·56 to −0·15)	**−0·31 (−0·51 to −0·10)**	**−0·24 (−0·43 to −0·03)**	−0·15 (−0·44 to 0·16)	−0·17 (−0·39 to 0·05)	−0·18 (−0·39 to 0·04)	−0·16 (−0·45 to 0·14)	−0·10 (−0·39 to 0·18)	−0·01 (−0·24 to 0·23)	ZIP	0·75 (0·37 to 1·51)	0·91 (0·61 to 1·34)	0·61 (0 34 to 1 06)	0·52 (0 17 to 1 58)
−0·48 (−0·77 to −0·19)	**−0·43 (−0·71 to −0·14)**	**−0·36 (−0·64 to −0·08)**	−0·32 (−0·67 to 0·06)	−0·29 (−0·58 to 0·00)	−0·29 (−0·58 to 0·00)	−0·28 (−0·63 to 0·08)	−0·22 (−0·57 to 0·12)	−0·13 (−0·43 to 0·18)	−0·12 (−0·43 to 0·19)	LAM	1·22 (0·67 to 2·21)	0·81 (0·40 to 1·65)	0·69 (0·21 to 2·30)
−0·56 (−0·69 to −0·43)	**−0·50 (−0·63 to −0·38)**	**−0·43 (−0·54 to −0·32)**	**−0·37 (−0·63 to −0·11)**	**−0·37 (−0·51 to −0·23)**	**−0·37 (−0·51 to −0·23)**	**−0·36 (−0·60 to −0·11)**	**−0·30 (−0·53 to −0·07)**	**−0·20 (−0·37 to −0·04)**	**−0·20 (−0·37 to −0·03)**	−0·08 (−0·34 to 0·18)	PBO	0·66 (0·44 to 1·00)	0·57 (0·20 to 1·62)
−0·63 (−0·84 to −0·43)	**−0·58 (−0·78 to −0·37)**	**−0·51 (−0·70 to −0·31)**	**−0·45 (−0·75 to −0·14)**	**−0·44 (−0·66 to −0·23)**	**−0·45 (−0·66 to −0·23)**	**−0·43 (−0·72 to −0·14)**	**−0·38 (−0·66 to −0·09)**	**−0·28 (−0·52 to −0·04)**	**−0·27 (−0·51 to −0·04)**	−0·15 (−0·46 to 0·15)	−0·07 (−0·24 to 0·09)	TOP	0·85 (0·28 to 2·63)
−0·88 (−1·40 to −0·36)	**−0·83 (−1·34 to −0·31)**	**−0·76 (−1·27 to −0·24)**	**−0·70 (−1·21 to −0·18)**	**−0·69 (−1·21 to −0·17)**	**−0·69 (−1·21 to −0·17)**	**−0·68 (−1·23 to −0·12)**	**−0·62 (−1·17 to −0·07)**	−0·53 (−1·05 to 0·01)	−0·52 (−1·05 to 0·01)	−0·40 (−0·96 to 0·16)	−0·32 (−0·82 to 0·18)	−0·25 (−0·77 to 0·28)	GBT

Treatment Efficacy (SMD with 95% CrI) Dropout rate (OR with 95% CrI)

Figure 4: **Efficacy and acceptability of all antimanic drugs according to multiple-treatments meta-analysis (primary outcomes)**

Drugs are reported in order according to efficacy ranking. Comparisons between treatments should be read from left to right and the estimate is in the cell in common between the column-defining treatment and the row-defining treatment. For efficacy, SMD below 0 favour the column-defining treatment· For acceptability, ORs higher than 1 favour the column-defining treatment. To obtain SMDs for comparisons in the opposite direction, negative values should be converted into positive values, and vice versa. To obtain ORs for comparisons in the opposite direction, reciprocals should be taken. Significant results are in bold and underscored. ARI=aripiprazole. ASE=asenapine. CBZ=carbamazepine. VAL=valproate. GBT=gabapentin. HAL=haloperidol. LAM=lamotrigine. LIT=lithium. OLZ=olanzapine. PBO=placebo. QTP=quetiapine. RIS=risperidone and paliperidone. TOP=topiramate. ZIP=ziprasidone. CrI=credibility interval. SMD=standardised mean difference.

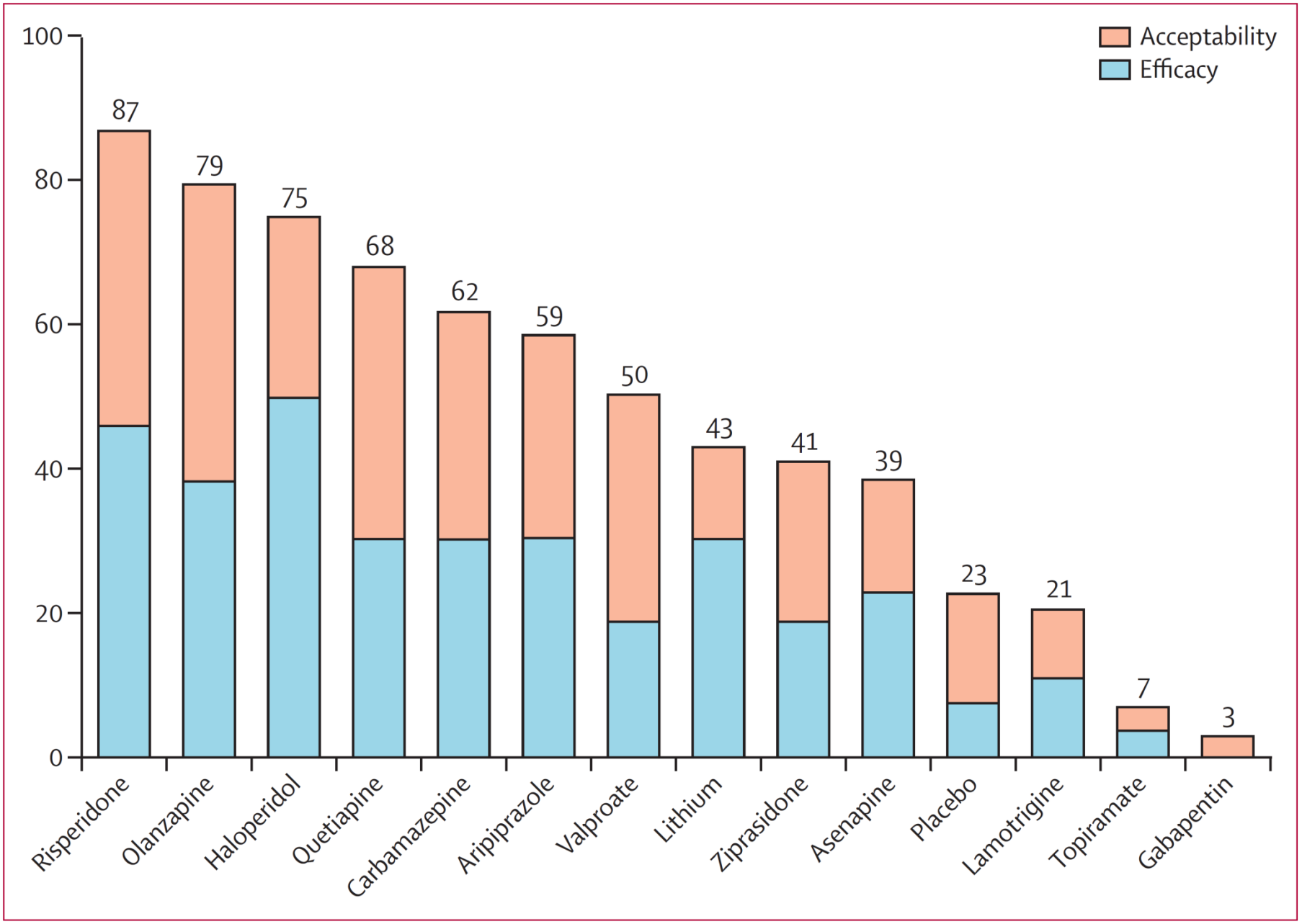


Figure 5: **Drugs ordered by their overall probability to be the best treatment in terms of both efficacy and dropout rate, showing the separate contributions to the overall scores of efficacy and dropout**

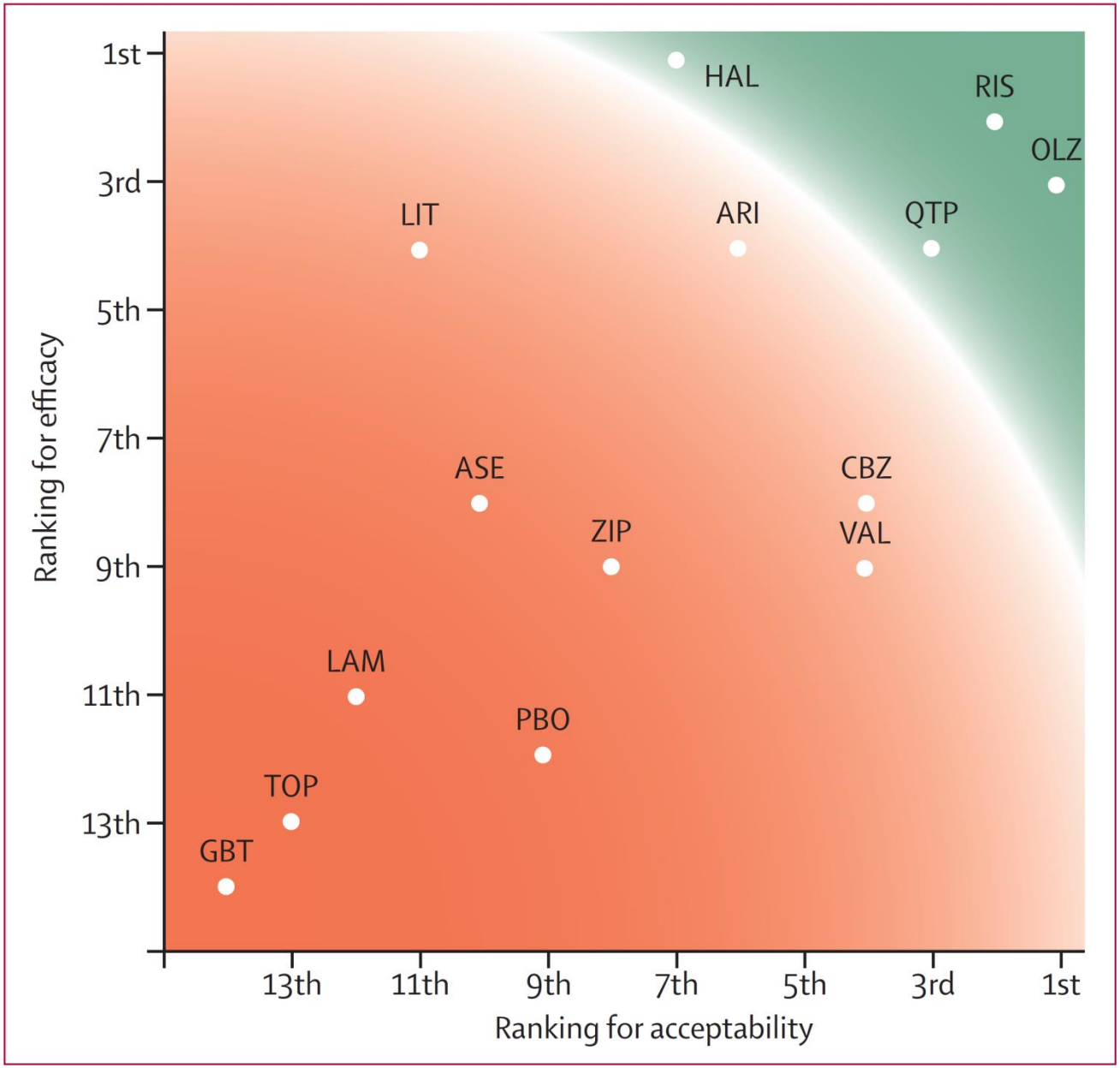


Figure 6: **Ranking of antimanic drugs according to primary outcomes: efficacy (as continuous outcome) and dropout rate**

3rd Course on Network Meta-Analysis

4th – 6th July 2016, Oxford

UNIVERSITY OF OXFORD

How to appraise, interpret and publish a network meta-analysis

A 3-day course for ***clinicians, researchers and policy makers*** wishing to expand their knowledge of indirect and mixed treatment comparisons in the ess evaluation.

Course highlights:

- Key features of network meta-analysis (NMA) & underlying methodological assumptions
- Critical appraisal of NMA
- Clinical questions in the context of NMA
- Write a NMA protocol
- Writing the manuscript
- Presenting NMA results
- How to reply to peer reviewers
- STATA workshop

Speakers: *Doug Altman, Tianjing Li, Deborah Caldwell, Stefan Leucht, Georgia Salanti, Huseyin Naci, Toshi Furukawa, Anna Chaimani, John Geddes, Andrea Cipriani*

The course takes place at Wolfson College, with formal lectures and small group activities

Course fee (all inclusive): £1300 (students £1000). Fee includes slides & scientific material, 2-night accommodation/breakfast/lunch at Wolfson College (free wifi)*, Gala Dinner at Merton College.

Course fee (without accommodation): £1200 (students £900).

*Extra-night(s) accommodation at Wolfson College available on a first come, first served basis.

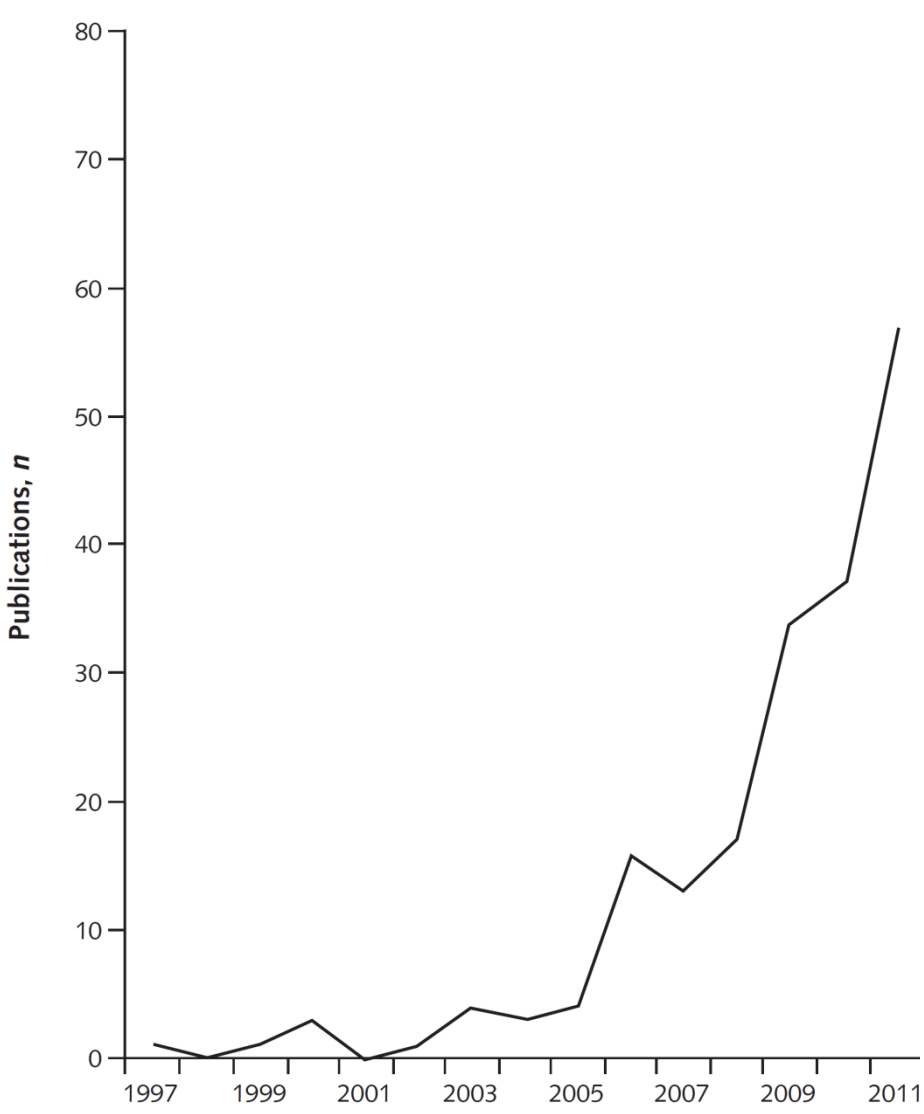
For more information, please contact the Course Coordinator

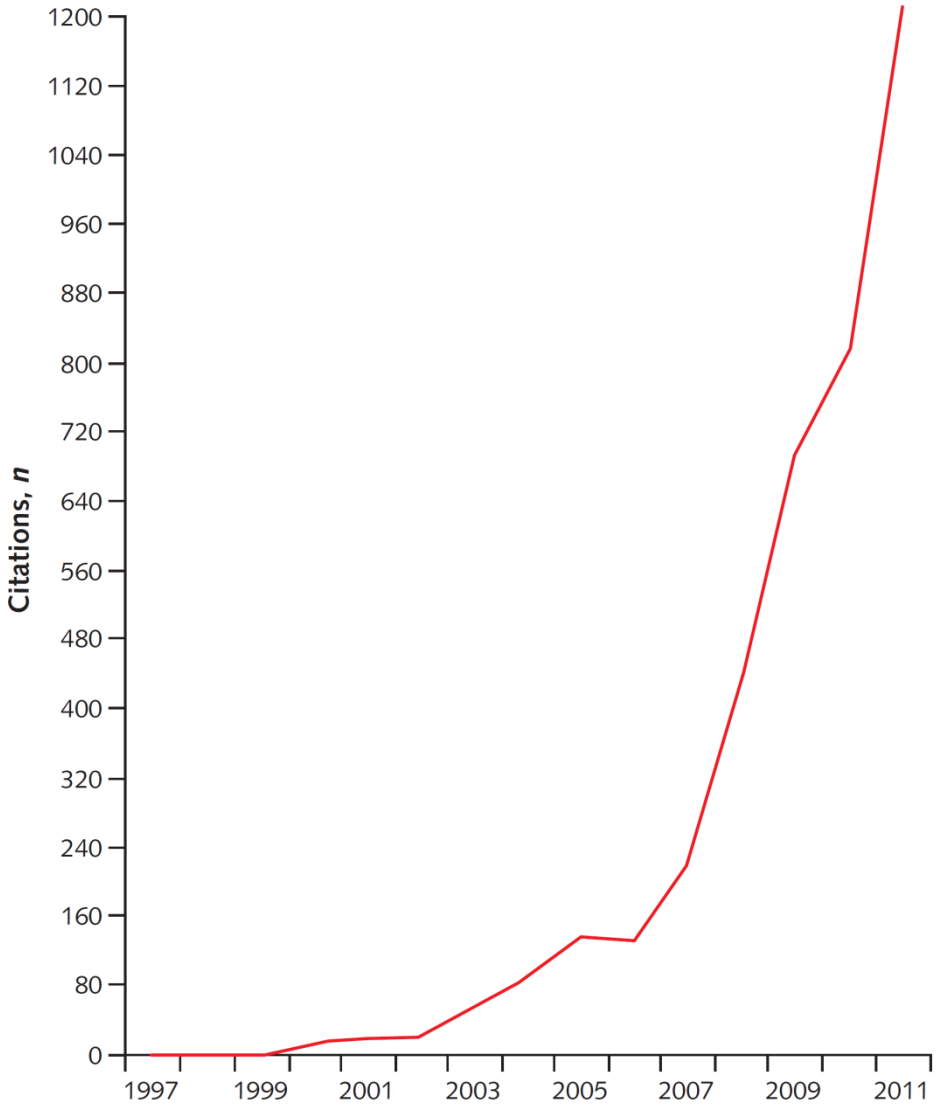
Lucy Curtin

University of Oxford
Department of Psychiatry
Oxford OX3 7JX
Tel: +44 (0) 1865 226451
lucy.curtin@psych.ox.ac.uk

Appendix Figure 1. **Number of network meta-analyses published in the scientific literature and their citations since 1997.**

We defined a network meta-analysis as any meta-analysis that used a form of valid indirect relative treatment estimates.





BMJ

RESEARCH

Efficacy of drug treatments for generalised anxiety disorder: systematic review and meta-analysis

David Baldwin, professor of psychiatry and honorary consultant psychiatrist,[1] Robert Woods, senior research executive,[2] Richard Lawson, statistician,[2] David Taylor, professor of psychopharmacology[3]

[1]University of Southampton Faculty of Medicine, University Department of Psychiatry, Academic Centre, Southampton SO14 3DT, UK

[2]Complete Medical Group, Macclesfield SK10 1AQ

[3]King's College London, London SE1 9NH

Correspondence to: D Baldwin
D.S.Baldwin@soton.ac.uk

Cite this as: *BMJ* 2011;342:d1199
doi:10.1136/bmj.d1199

ABSTRACT

Objective To appraise the evidence for comparative efficacy and tolerability of drug treatments in patients with generalised anxiety disorder.

Design Systematic review of randomised controlled trials. Primary Bayesian probabilistic mixed treatment meta-analyses allowed pharmacological treatments to be ranked for effectiveness for each outcome measure, given as percentage probability of being the most effective treatment. Secondary frequentist mixed treatment meta-analyses conducted with random effects model; effect size reported as odds ratio and 95% confidence interval.

Data sources Medline, Embase, BIOSIS, PsycINFO, Health Economic Evaluations Database, National Health Service Economic Evaluation Database, and Database of Abstracts of Reviews of Effects via DataStar, and Cochrane

meta-analyses, fluoxetine was ranked first for response and remission (probability of 62.9% and 60.6%, respectively) and sertraline was ranked first for tolerability (49.3%). In a subanalysis ranking treatments for generalised anxiety disorder currently licensed in the United Kingdom, duloxetine was ranked first for response (third across all treatments; 2.7%), escitalopram was ranked first for remission (second across all treatments; 26.7%), and pregabalin was ranked first for tolerability (second across all treatments; 7.7%).

Conclusions Though the frequentist analysis was inconclusive because of a high level of uncertainty in effect sizes (based on the relatively small number of comparative trials), the probabilistic analysis, which did not rely on significant outcomes, showed that fluoxetine (in terms of response and remission) and sertraline (in

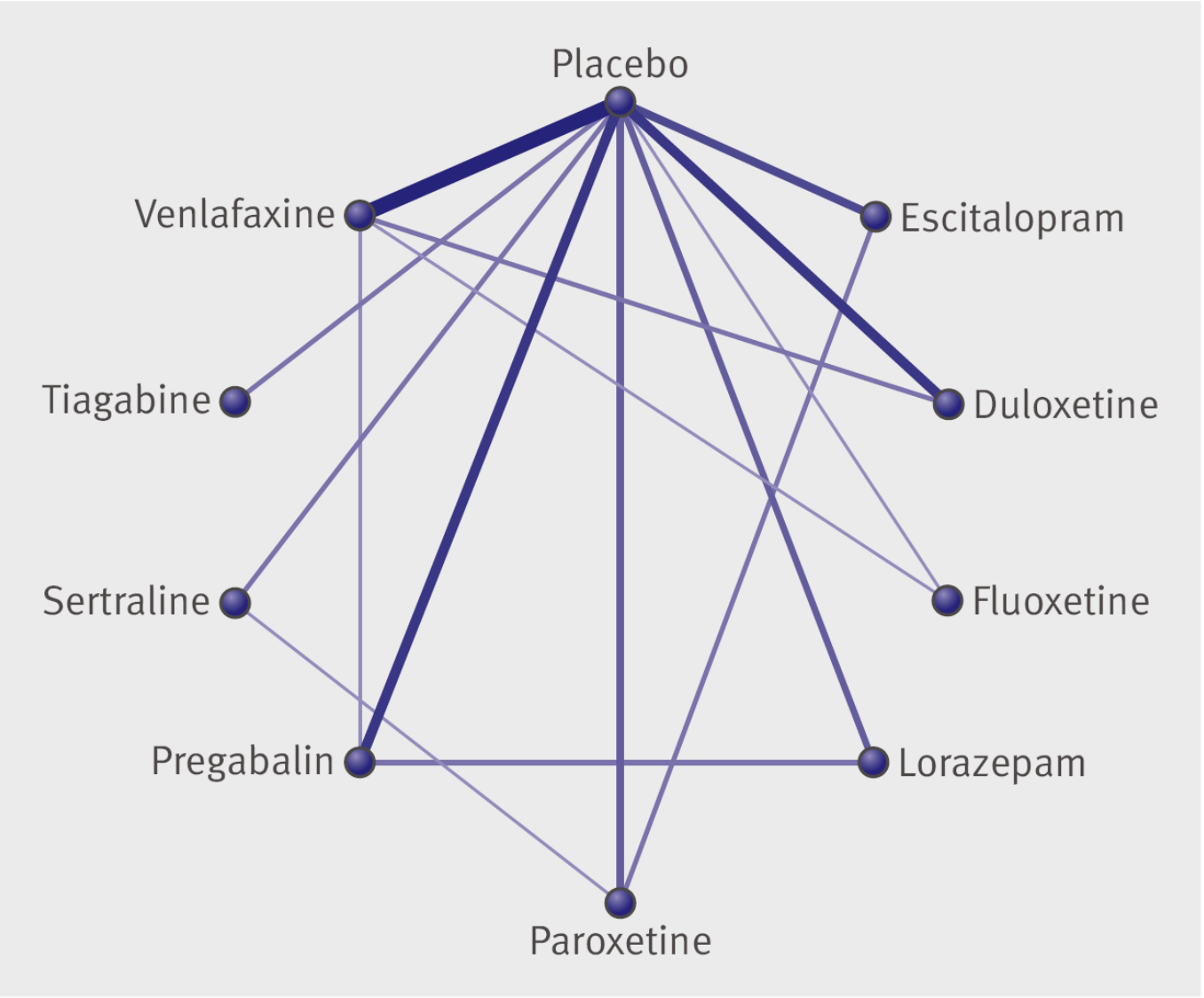


Fig 1 | **Eligible network comparisons between all treatments, with increasing thickness of lines indicating increasing number of direct comparisons**

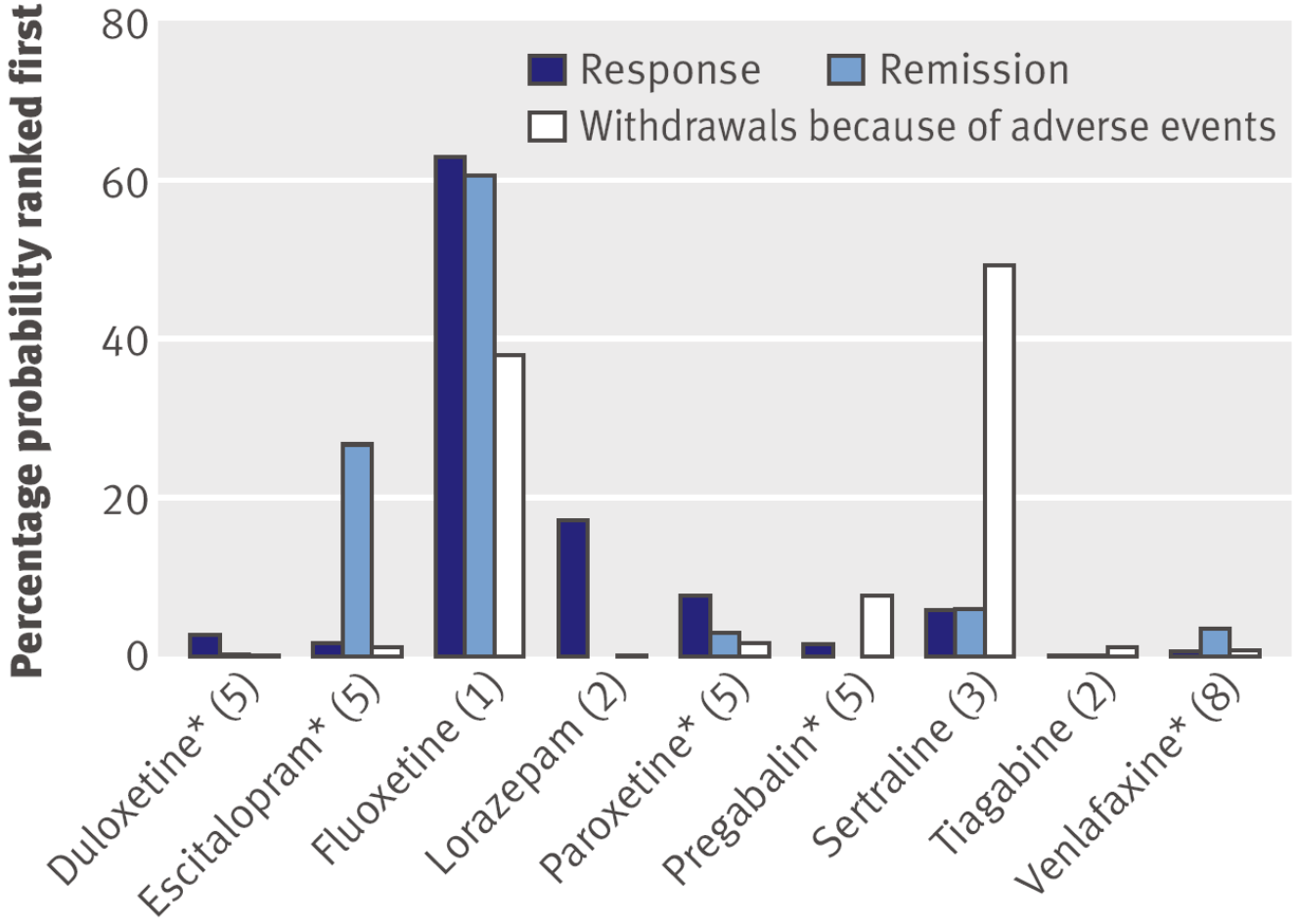
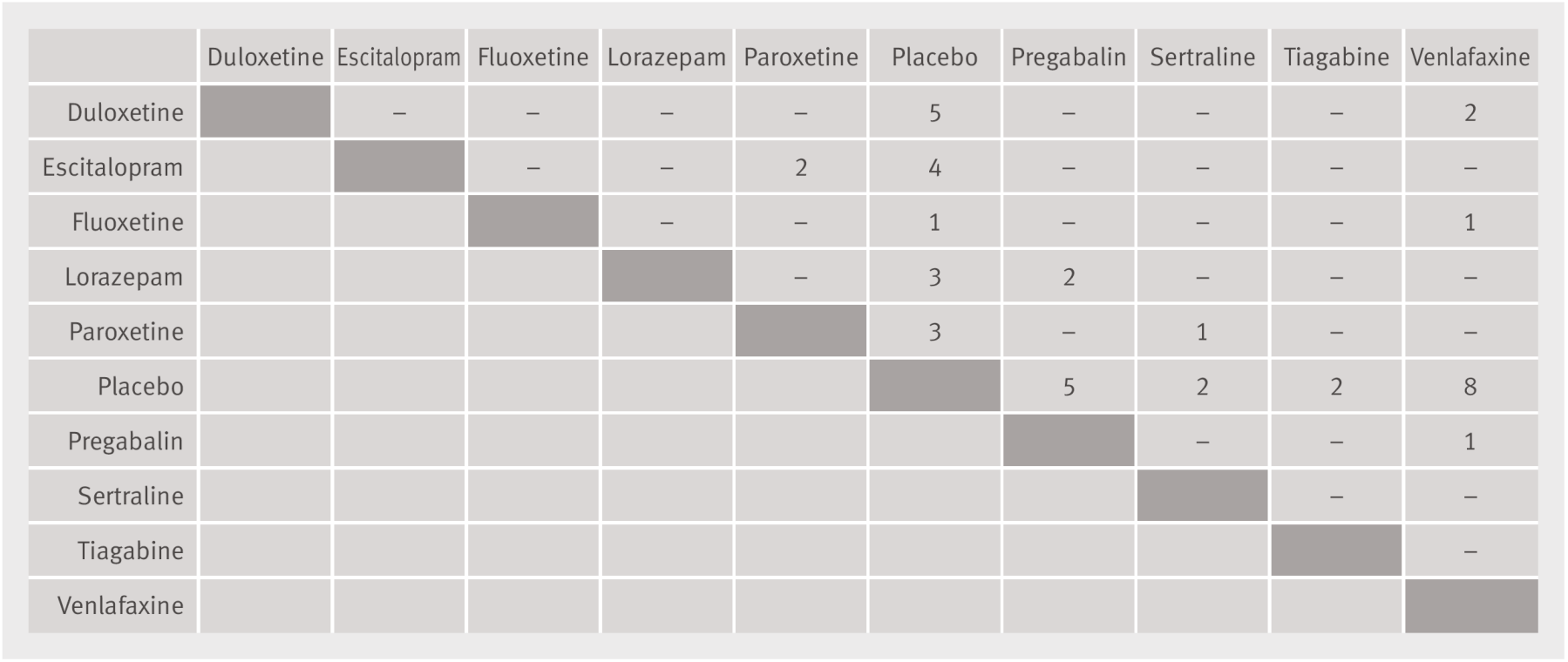


Fig 4 | Probabilistic analysis showing percentage probability of each treatment being ranked first by outcome measure. Numbers in parentheses indicate number of trials analysed for each treatment. Remission data were not available for lorazepam and pregabalin. *Drug licensed in UK



	Duloxetine	Escitalopram	Fluoxetine	Lorazepam	Paroxetine	Placebo	Pregabalin	Sertraline	Tiagabine	Venlafaxine
Duloxetine		–	–	–	–	5	–	–	–	2
Escitalopram			–	–	2	4	–	–	–	–
Fluoxetine				–	–	1	–	–	–	1
Lorazepam					–	3	2	–	–	–
Paroxetine						3	–	1	–	–
Placebo							5	2	2	8
Pregabalin								–	–	1
Sertraline									–	–
Tiagabine										–
Venlafaxine										

Fig 3 | **Number of direct comparisons between treatments (or placebo) for generalised anxiety**

Efficacy of Venlafaxine Extended Release in Patients With Major Depressive Disorder and Comorbid Generalized Anxiety Disorder

Peter H. Silverstone, M.D., F.R.C.P.C., and Eliseo Salinas, M.D.

Background: A subset of patients with comorbid major depressive disorder and generalized anxiety disorder (GAD) was examined from a double-blind, placebo-controlled study comparing the efficacy and safety of venlafaxine extended release (XR) and fluoxetine.

Method: From a total of 368 patients, 92 patients meeting DSM-IV criteria for major depressive disorder who also had comorbid GAD were identified. The comparison group comprised 276 evaluable noncomorbid patients. Patients received venlafaxine XR (75–225 mg/day), fluoxetine (20–60 mg/day), or placebo for 12 weeks. Efficacy evaluations included Hamilton Rating Scale for Depression (HAM-D), Hamilton Rating Scale for Anxiety (HAM-A), and Clinical Global Impressions (CGI) scale.

Results: By the final assessment at week 12, comorbid patients in the venlafaxine XR group, but not in the fluoxetine group, showed a significantly greater decrease than those in the placebo group in the primary efficacy variables of mean HAM-D and in comorbid patients.

(J Clin Psychiatry 2001;62:523–529)

Received Aug. 5, 2000; accepted March 29, 2001. *From the Department of Psychiatry, University of Alberta, Edmonton, Alberta, Canada (Dr. Silverstone) and the CNS Department, Wyeth-Ayerst Research, Paris, France (Dr. Salinas).*

Sponsored by Wyeth-Ayerst Research.

Reprint requests to: Peter H. Silverstone, M.D., Department of Psychiatry, Psychopharmacology Research Unit, 1E7.13 Mackenzie Centre, University of Alberta, Edmonton, Alberta, Canada, T6G 2B7 (e-mail: peter.silverstone@ualberta.ca).

The majority of patients presenting with major depressive disorder in general practice also have some degree of associated or concomitant anxiety symptoms.[1] Indeed, major depression with acute or subsyndromal anxiety is more common than either condition alone.[2] Generally, the occurrence of depression and anxiety symptoms together is associated with greater severity of symptoms, greater impairment, more chronic course of illness, poorer outcome, and higher incidence of suicide.[3,4] In a significant proportion of these patients, the severity, quality, and chronicity of symptoms of depression and anxiety are sufficient to fulfill diagnostic criteria for major depressive disorder and generalized anxiety disorder (GAD) simultaneously.[5]

Table 2. Adjusted Mean Scores and Between-Group Comparisons Versus Placebo in Patients With Major Depressive Disorder and Comorbid GAD[a]

	Fluoxetine (N = 33)			Venlafaxine XR (N = 32)	
Scale	Placebo (N = 25)	Adjusted Mean Score	Adjusted Mean Different From Placebo (95% CL)	Adjusted Mean Score	Adjusted Mean Different From Placebo (95% CL)
HAM-D total					
Baseline	27.9	27.9		27.9	
Week 1	22.8	22.8	–0.1 (–2.8, 2.7)	24.2	–1.4 (–4.2, 1.4)
Week 2	21.2	20.4	0.8 (–2.2, 3.8)	21.0	0.3 (–2.8, 3.3)
Week 3	20.5	18.9	1.6 (–1.4, 4.7)	19.0	1.5 (–1.6, 4.6)
Week 4	18.5	16.9	1.7 (–1.6, 5.0)	17.3	1.2 (–2.2, 4.6)
Week 6	16.5	16.8	–0.3 (–3.9, 3.3)	15.9	0.6 (–3.0, 4.3)
Week 8	16.0	15.5	0.5 (–3.4, 4.5)	14.5	1.5 (–2.5, 5.6)
Week 12	16.5	14.0	2.5 (–1.7, 6.6)	11.7*	4.8 (0.6, 9.0)
HAM-A total					
Baseline	25.7	25.7		25.7	
Week 1	22.4	21.8	0.6 (–1.9, 3.1)	23.6	–1.2 (–3.8, 1.3)
Week 2	20.6	20.0	0.5 (–2.5, 3.6)	20.4	0.2 (–2.9, 3.3)
Week 3	20.2	18.6	1.6 (–1.6, 4.8)	19.2	1.0 (–2.2, 4.3)
Week 4	19.4	17.2	2.3 (–1.3, 5.8)	17.0	2.4 (–1.2, 6.1)
Week 6	17.5	17.6	–0.1 (–3.8, 3.6)	15.6	1.9 (–1.8, 5.6)
Week 8	16.1	15.9	0.2 (–4.1, 4.5)	14.4	1.7 (–2.6, 6.0)
Week 12	16.9	14.4	2.5 (–1.7, 6.7)	12.5*	4.5 (0.2, 8.7)

[a]Based on data from Silverstone and Ravindran.[14] Analysis based on last observation carried forward. Abbreviations: CL = confidence limits, GAD = generalized anxiety disorder, HAM-A = Hamilton Rating Scale for Anxiety, HAM-D = Hamilton Rating Scale for Depression, XR = extended release.
*Significant difference vs. placebo ($p < .05$).

METHOD

Study Design

The present analysis used data from a previously reported prospective, multicenter, double-blind, randomized, placebo-controlled comparative study of the efficacy and tolerability of once-daily venlafaxine XR and fluoxetine in 368 patients with major depressive disorder and concomitant anxiety.[14] In this study, major depressive

RESULTS

From a total of 368 patients, a subset of 92 patients was identified with comorbid major depressive disorder and GAD at baseline. Ninety patients who had at least 1 baseline evaluation for at least 1 of the primary efficacy parameters and had at least 1 on-therapy evaluation for at least 1 of the primary efficacy parameters formed the ITT population of "comorbid patients." For comparative purposes, 269 of the remaining 276 patients were evaluable for the ITT efficacy population of "noncomorbid patients."

14. Silverstone PH, Ravindran A, for the Venlafaxine XR 360 Study Group. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. J Clin Psychiatry 1999;60: 22–28

How to Use an Article Reporting a Multiple Treatment Comparison Meta-analysis

Edward J. Mills, PhD, MSc
John P. A. Ioannidis, MD, DSc
Kristian Thorlund, PhD, MSc
Holger J. Schünemann, MD, PhD, MSc
Milo A. Puhan, MD, PhD
Gordon H. Guyatt, MD, MSc

Multiple treatment comparison (MTC) meta-analysis uses both direct (head-to-head) randomized clinical trial (RCT) evidence as well as indirect evidence from RCTs to compare the relative effectiveness of all included interventions. The methodological quality of MTCs may be difficult for clinicians to interpret because the number of interventions evaluated may be large and the methodological approaches may be complex. Clinicians and others evaluating an MTC should be aware of the potential biases that can affect the interpretation of these analyses. Readers should consider whether the primary studies are sufficiently homogeneous to combine; whether the different interventions are sufficiently similar in their populations, study designs, and outcomes; and whether the direct evidence is sufficiently similar to the indirect evidence to consider combining. This article uses the existing Users' Guides format to address study validity, interpretation of results, and application to a patient scenario.

JAMA. 2012;308(12):1246-1253 www.jama.com

BMJ

BMJ 2013;346:f2914 doi: 10.1136/bmj.f2914 (Published 14 May 2013)

RESEARCH METHODS & REPORTING

Demystifying trial networks and network meta-analysis

Networks of randomized clinical trials can be evaluated in the context of a network meta-analysis, a procedure that permits inferences into the comparative effectiveness of interventions that may or may not have been evaluated directly against each other. This approach is quickly gaining popularity among clinicians and guideline decision makers. However, certain methodological aspects are poorly understood. Here, we explain the geometry of a network, statistical and conceptual heterogeneity and incoherence, and challenges in the application and interpretation of data synthesis. These concepts are essential to make sense of a network meta-analysis.

Edward J Mills *associate professor*[1 2], Kristian Thorlund *associate professor*[2 3], John P A Ioannidis *professor*[2 4]

[1]Faculty of Health Sciences, University of Ottawa, 35 University Drive, Ottawa, ON, Canada, K1N 6N5; [2]Stanford Prevention Research Center, Department of Medicine, Stanford University School of Humanities and Sciences, Stanford, CA, USA; [3]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada; [4]Department of Health Research and Policy, Stanford University School of Medicine, and Department of Statistics, Stanford University School of Humanities and Sciences, Stanford, CA, USA

Annals of Internal Medicine | Research and Reporting Methods

Conceptual and Technical Challenges in Network Meta-analysis

Andrea Cipriani, PhD; Julian P.T. Higgins, PhD; John R. Geddes, MD; and Georgia Salanti, PhD

The increase in treatment options creates an urgent need for comparative effectiveness research. Randomized, controlled trials comparing several treatments are usually not feasible, so other methodological approaches are needed. Meta-analyses provide summary estimates of treatment effects by combining data from many studies. However, an important drawback is that standard meta-analyses can compare only 2 interventions at a time. A new meta-analytic technique, called network meta-analysis (or multiple treatments meta-analysis or mixed-treatment comparison), allows assessment of the relative effectiveness of several interventions, synthesizing evidence across a network of randomized trials. Despite the growing prevalence and influence of network meta-analysis in many fields of medicine, several issues need to be addressed when constructing one to avoid conclusions that are inaccurate, invalid, or not clearly justified. This article explores the scope and limitations of network meta-analysis and offers advice on dealing with heterogeneity, inconsistency, and potential sources of bias in the available evidence to increase awareness among physicians about some of the challenges in interpretation.

Ann Intern Med. 2013;159:* * * FILL THIS IN * * *. www.annals.org

For author affiliations, see end of text.

The increase in alternative medical treatment options has led to the need for comparative effectiveness research (1, 2). Randomized, controlled trials comparing many treatment options are usually not feasible, so other methodological approaches are needed. A meta-analysis embedded in a systematic review is a useful statistical tool that provides a summary estimate of treatment effect by combining data from many studies. However, a key limitation of standard (or pairwise) meta-analyses is that they can compare only 2 interventions at a time. When several treatment options are available, a series of individual meta-analyses provides only partial information because it can only answer questions about pairs of treatments. This does not support optimal clinical decision making because each meta-analysis is only one part of the whole picture.

What Is a Network Meta-analysis?

Figure 1 presents a network of pharmacologic treatments for acute mania. The primary outcome for short-term treatment efficacy is the change score on a rating scale for manic symptoms (11). The lines between treatment nodes indicate which comparisons have been made in randomized trials. If there is no line between 2 nodes, then no studies (that is, no direct evidence) compare the 2 drugs. A network meta-analysis is a simultaneous analysis of the data from all of these randomized trials. With a network meta-analysis, the relative effectiveness of 2 treatments can be estimated even if no studies directly compare them. For example, no single study has compared aripiprazole and risperidone but using a common comparator (placebo) allows for an indirect comparison between them. Denoting

Thanks!

andrea.cipriani@psych.ox.ac.uk