

## RESEARCH

# The effects of excluding treatments from network meta-analyses: survey

 OPEN ACCESS

Edward J Mills *associate professor*<sup>1,2</sup>, Steve Kaners *biostatistician*<sup>1</sup>, Kristian Thorlund *associate professor*<sup>2,3</sup>, Anna Chaimani *PhD student*<sup>4</sup>, Areti-Angeliki Veroniki *PhD student*<sup>4</sup>, John P A Ioannidis *professor and director*<sup>2</sup>

<sup>1</sup>Faculty of Health Sciences, University of Ottawa, Ottawa, Canada; <sup>2</sup>Stanford Prevention Research Center, Stanford University, Stanford, USA;

<sup>3</sup>Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Canada; <sup>4</sup>Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece

## Abstract

**Objective** To examine whether the exclusion of individual treatment comparators, including placebo/no treatment, affects the results of network meta-analysis.

**Design** Survey of networks with individual trial data.

**Data sources** PubMed and communication with authors of network meta-analyses.

**Study selection and methods** We included networks that had five or more treatments, contained at least two closed loops, had at least twice as many studies as treatments, and had trial level data available. Investigators abstracted information about study design, participants, outcomes, network geometry, and the exclusion of eligible treatments.

**Results** Among 18 eligible networks involving 757 randomised controlled trials with 750 possible treatment comparisons, 11 had upfront decided not to consider all treatment comparators and only 10 included placebo/no treatment nodes. In 7/18 networks, there was at least one node whose removal caused a more than 1.10-fold average relative change in the estimated treatments effects, and switches in the top three treatments were observed in 9/18 networks. Removal of placebo/no treatment caused large relative changes of the treatment effects (average change 1.16-3.10-fold) for four of the 10 networks that had originally included placebo/no treatment nodes. Exclusion of current uncommonly used drugs resulted in substantial changes of the treatment effects (average 1.21-fold) in one of three networks on systemic treatments for advanced malignancies.

**Conclusion** Excluding treatments in network meta-analyses sometimes can have important effects on their results and can diminish the usefulness of the research to clinicians if important comparisons are missing.

## Introduction

Network meta-analysis (also called multiple or mixed treatment comparison meta-analysis, MTC) permits the evaluation of the comparative effectiveness of multiple interventions.<sup>1,2</sup> This approach has an inherent appeal for clinicians and decision makers as new or existing interventions must be placed within the context of all available evidence.<sup>3-6</sup> Often, those undertaking an MTC will selectively choose interventions to include in the analysis. For example, some MTCs exclude placebo or no treatment from consideration because it is sometimes believed that placebo trials vary over time or are set in favourable conditions to appease regulatory authorities.<sup>7</sup> Other MTCs may include only the treatments available in particular settings (for example, a specific country), only those of perceived dose relevance, or (often in the case of industry submissions to health technology assessment bodies) only specific competing treatments.<sup>8</sup> To obtain empirical evidence on whether these choices make a difference in the results such as treatment effect estimates and treatment rankings, we examined a sample of complex networks and reanalysed their data after excluding specific treatment nodes.

## Methods

### Eligibility criteria and retrieval of data from existing networks

We considered networks that had five or more treatments, contained at least two closed loops, had at least twice as many studies as nodes, and had individual trial level data or estimates available. The eligibility criteria aimed to generate a sample of

Correspondence to: E J Mills Faculty of Health Sciences, University of Ottawa, 43 Templeton Street, Ottawa, Canada K1N6X1  
edward.mills@uottawa.ca

Extra material supplied by the author (see <http://www.bmj.com/content/347/bmj.f5195?tab=related#webextra>)

Figure displaying the concepts of loops and connectivity

networks that had many treatments and studies and sufficient data to explore the impact of exclusions. We used a systematic literature search that has been published previously that identified potentially eligible networks.<sup>9</sup> We also attempted to contact study authors for missing individual data at trial level. We included an additional network from an MTC conducted by our team, where we had direct access to the primary data at trial level. In studies that considered more than one outcome using MTCs, we favoured the efficacy outcome over safety outcomes.

## Data abstraction

For each eligible network with available trial level data, we recorded whether the eligibility criteria excluded specific types of active or inactive/control (placebo, no treatment, best supportive care) treatment comparators, and the rationale for such exclusions. We recorded for each network the number of studies, treatments, and loops; the geometry of the network (the distribution of treatments and comparisons thereof in each network); the condition being treated; the primary outcome measure and the statistical effect measure used; and the range of node connectivity (the number of direct comparisons connected to each node). The supplementary figure displays the concepts of loops and connectivity.

## Statistical analysis

Regardless of the analyses chosen in the original publications, we analysed each network using random effects Bayesian MTCs with uninformed priors, the most common analytical approach used for network meta-analysis.<sup>8</sup> Details on code and specific analysis are available from the authors.

For each network we analysed the complete available data (full model) and also performed analyses excluding one or multiple treatment nodes—that is, disregarding in the calculations data from trials where the excluded nodes were comparators. Firstly, we investigated the effect of excluding the treatment node with the largest expected impact from each network. We used the Brier score to identify the treatment node with the largest expected impact on results.<sup>10</sup> The Brier score is the average of the squared differences between the log ratios (odds, relative risk or hazard) estimated with the full treatment network data versus the treatment network data where one or more treatment nodes are excluded. Secondly, we investigated the effect of excluding other single treatment nodes that could be classified as active interventions (that is, not placebo/no treatment). Thirdly, we investigated the impact of excluding placebo/no treatment from the treatment network. Lastly, we focused on selected examples of situation specific exclusions that were chosen owing to perceived relevance for clinical practise or decision making. Specifically, for the network of thrombolytic therapy we excluded data from trials involving anisoylated plasminogen streptokinase activated complex (ASPAC) and urokinase because these treatments were not available in the United Kingdom and one previous UK based network meta-analysis had excluded them<sup>11</sup>; and for the networks of breast, colorectal, and ovarian cancer, we excluded nodes of treatment regimens that are currently uncommonly used (all those deemed miscellaneous older agents alone or in combination for breast cancer; fluorouracil monotherapy and older drug monotherapy for colorectal cancer; platinum monotherapy and non-platinum/non-taxane drug regimens for ovarian cancer).

We used odds ratios for dichotomous outcomes, rate ratios for Poisson outcomes, and hazard ratios for survival outcomes.

From each network we estimated treatment effects for each treatment comparison and the probability of each treatment being the best treatment.<sup>12</sup>

To evaluate the impact of excluding treatment nodes we calculated the relative change in the estimated treatment effects of the remaining treatments; changes in the top three ranked treatments, according to the probability of being best; and changes in the probability of being the best treatment for the top treatment of the full model. Relative changes in the estimated treatment effects of remaining nodes are always expressed as a value greater than 1.00—for example, if a specific odds ratio is 0.80 in the full network and 0.88 in the network with one node excluded, then the relative change is  $0.88/0.80=1.10$ -fold; if these odds ratios are 0.80 and 0.60, then the relative change is  $0.80/0.60=1.25$ -fold. Given that most treatment effects are relatively small or modest,<sup>13</sup> relative changes of 1.10-fold can be considered large and of 1.20-fold can be considered substantial. Choice of reference group does not affect the values of the fold change. As most networks have a natural choice for the reference group we chose to keep this node as a reference group to best mimic changes that would be observed by researchers. For each network with excluded nodes, we noted the maximum and geometric average of the relative change.

Analyses were conducted using WinBUGS version 1.4 (Medical Research Council Biostatistics Unit, Cambridge) and R version 2.15.1 ([www.r-project.org/](http://www.r-project.org/)).

## Results

### Analysed networks

The search identified 890 relevant abstracts, of which 276 were assessed as potentially eligible and their full articles were screened. After adding an eligible network from our own available data, 18 networks with individual trial data met the inclusion criteria,<sup>14-31</sup> involving 757 randomised controlled trials with 750 possible treatment comparisons and 1 036 701 patients. Table 1 shows the characteristics of the individual networks. Networks ranged in size from 11 studies on five treatments to 128 studies on 22 treatments (figure 1).

### Exclusion of treatment nodes in analysed networks

Table 1 shows the eligibility criteria that had been set by the original authors of these meta-analyses and that resulted in exclusion of specific treatment comparators from the network. Eleven of the 18 networks upfront decided not to consider all treatment comparators. One network mentioned that a previous network meta-analysis had excluded data on two treatments (ASPAC and urokinase for thrombolysis) that were not licensed in the United Kingdom, but the full data were available in the current analysis. Several networks deliberately excluded placebos or no treatment options (for example, best supportive care for cancers). Several other networks focused on specific types of treatments that they considered to be of clinical interest and made statements that other treatments were not considered at all, because they were of different class, old (for example, trials published before 1997), or to be considered in a separate review.

### Changes after removal of treatment node with largest expected impact

When the node with the largest Brier score was removed (table 2) the average relative change of treatments effects exceeded 1.10-fold in seven of the 18 networks.<sup>14 16 18 22 27 29 31</sup> The largest

observed relative change exceeded 1.10-fold (maximum 3.64-fold) in all but four networks. In three networks some non-significant effects became significant, and in 12 networks the opposite change was seen. Switches in the ranking of the top three treatments were observed in nine of the 18 networks.<sup>16-18 21 22 25 27 29 31</sup> Table 3⇓ shows a worked example of the calculations and results for a single node removal in the Cooper 2006 network.<sup>18</sup>

The most influential nodes were typically highly connected. In 10 of the 18 networks they actually represented the most connected node. However, the most influential nodes were never the top ranked treatment of the networks and almost always (15/18) had a 10% or less probability of being the best treatment (0% probability in 11/18 networks). Thus standards of care, placebo, or other treatments that may easily be overlooked represent nodes of which the exclusion could be influential to the network results.

### Changes of removal after other active intervention nodes

Removal of the remaining, less influential nodes caused minor changes (average relative change  $\leq 1.03$ -fold) in 13 of the 18 networks. None the less, seven networks contained a node the exclusion of which caused a relative change greater than 1.10-fold. Of the nine networks with changes in the top ranks by the exclusion of the most influential node, three had an additional node whose removal affected the top ranks.<sup>16 18 29</sup>

### Changes after placebo node removals or context specific node removals

Table 4⇓ presents the results after removal of placebo/no treatment nodes or after removal of context specific sets of treatment nodes. Placebos and no treatment nodes were removed from 10 networks that had originally included such nodes. This resulted in important changes in the results of four of these 10 networks,<sup>16 18 22 27</sup> with average relative changes ranging from 1.16-fold to 3.10-fold, whereas in the other six networks the average relative changes were minor (1.01-1.03-fold). In one network, exclusion of the placebo node resulted in a switch in position between the first and second best treatment. In the other cases, the ranking of the top three treatments did not change, but modest to large changes were seen in the probability of being the best treatment, such as in Mills 2011 in which the probability of high dose nicotine replacement being the best treatment went from 58% to 90% (table 4).<sup>26</sup>

Removal of APSAC and urokinase from the network on thrombolysis and percutaneous intervention did not result in major changes. Removal of uncommonly used treatment regimens from the networks of treatments for malignancies resulted in a large change in estimated effects (on average by 1.21-fold) for ovarian cancer and small changes for colorectal and breast cancer. However, in the case of advanced breast cancer, these removals switched the ranks of novel non-taxane agents+taxanes and taxanes (combination regimen), the first and third ranked treatments, respectively.

## Discussion

In this metaepidemiological study involving 18 large network meta-analyses, we found that it is common for networks to exclude trials and specific comparators based on widely diverse criteria. Furthermore, we documented that exclusions occasionally can have an important impact on the results. The comparators that had the largest potential to change the results

were typically those used in many trials; and they were unlikely to be the best treatments. We also explored the impact of exclusion of each other active treatment and placebo/no treatment as well as the impact of excluding several treatments based on scenarios that may be clinically relevant. Our results show, moreover, that some of these exclusions could affect substantially the estimated treatment effects and occasionally even affect the ranking of the top treatments.

These findings seem particularly relevant to network meta-analysis where analysts choose only to evaluate certain newer treatments or where they have chosen to exclude well established interventions or placebo, or both from a network. Readers of network meta-analysis should examine whether a network represents all available interventions. Excluding treatments from networks may reduce the applicability and usefulness of a multiple or mixed treatment comparison (MTC) if comparisons of interest to doctors are not considered. A substantial literature already exists on traditional pairwise meta-analyses, showing how eligibility criteria may result in differences in the results and conclusions of meta-analyses on the same topic.<sup>32-36</sup> As meta-analyses have become popular, most of the topics of interest are addressed by two or more published meta-analyses and these may differ in their eligibility criteria.<sup>37</sup> Some examples have started to accumulate where independent network analyses on the same topic may reach different conclusions.<sup>38</sup> A key consideration may be which treatment nodes are considered eligible for analysis. For example, on evaluating the relative merits of second generation antidepressants, an MTC that did not consider placebo comparisons yielded different rankings from one that included only placebo controlled comparisons and one that considered both types of comparisons.<sup>17 39-41</sup> We should also note that although treatment rankings and probabilities are arguably easy to interpret, their interpretability is limited by the fact that they are driven predominantly by the estimated effect sizes, and that standard errors play an unduly small role in determining their position.<sup>32</sup> Readers are advised to observe the estimated effects first and use the rankings only as a supplementary measure.

### Limitations of this study

There are some caveats to our analysis. Firstly, we excluded simple networks such as star networks (that cannot be analysed in an MTC)<sup>42</sup> and poorly connected networks. Less well connected networks with less total evidence are likely to be more “fragile” when nodes are removed—that is, the results may change even more from what we documented here for well connected networks. Secondly, we applied a uniform approach—a random effects analysis using non-informative priors—to our reanalysis of each published network meta-analysis. Different analytical choices may introduce some further variability in the results. Thirdly, some of the examined networks had already excluded specific trials and treatments upfront, and it was not practical for us to try to retrieve the excluded trials and evaluate the impact of their inclusion. Conversely, we focused on the potential impact that further exclusions would have had on the results. Fourthly, we did not question the clustering of treatment regimens into specific nodes by the original meta-analysts. However, in some MTCs there may be some ambiguity on whether slightly or modestly different regimens (for example, different doses or schedules for administering drugs or drugs belonging to the same class) should be treated as a single node or separate nodes. This could introduce some additional variability in the results, depending on what choices are made. For example, in the MTC of systemic treatment for colorectal cancer,<sup>21</sup> 242 trials were identified but

only 37 could be used, because the others were deemed to compare regimens that belonged to the same treatment node and thus would not be distinguishable. In all, given these reasons our estimates of the potential impact on the results of choices pertaining to the geometry of the network are probably conservative. To avoid potential bias, network meta-analysts should consider the inclusion of all relevant interventions for the condition of interest.

We should acknowledge that the choice of including all possible interventions that have ever been evaluated in randomised controlled trials on a topic can be daunting. Sometimes specific exclusions are clearly justifiable, and many of these exclusions may not have any substantial impact on the estimated treatment effects of the remaining “relevant” interventions. However, exclusions of potential nodes and decisions about eligibility criteria need to be carefully justified—for example, it may be argued that alternative medicine or non-mainstream comparators should not be included in the same network as standard interventions. It may even be reasonable to perform sensitivity analyses to examine the impact of the removal of specific nodes. Moreover, our findings indicate that the largest impact on the results occurs when well connected nodes are removed. It seems reasonable to advocate that the most evaluated treatments available for a condition should be considered necessary to include for a network to be valid. Typically, this includes older established standards of care, placebo, and well-tested interventions. In particular, some industry sponsored meta-analyses increasingly focus on target competitor agents that compete for market share.<sup>43</sup> However, the exclusion of other interventions in a network can importantly affect the results.

The authors appreciate the contribution of Georgia Salanti for providing data and comments on protocol and analysis.

Contributors: EJM, SK, KT, and JPAI conceived the study design. AC, A-AV, and SK acquired the data. SK, KT, and EJM conducted the analyses. EJM, SK, KT, and JPAI wrote the manuscript and all contributed to the writing. All authors approved the final manuscript. EJM is guarantor for the study.

Funding: This study was supported by the Drug Safety and Effectiveness Network (DSEN) of the Canadian Institutes of Health Research. No funding agency has seen this study.

Competing interests: All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: The technical appendix and statistical code are available from the corresponding author at [joannid@stanford.edu](mailto:joannid@stanford.edu).

- Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
- Mills EJ, Bansback N, Ghement I, Thorlund K, Kelly S, Puhon MA, et al. Multiple treatment comparison meta-analyses: a step forward into complexity. *Clin Epidemiol* 2011;3:193-202.
- Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 2008;26:753-67.
- Wells GA, Sultan SA, Chen L, Khan M, Coyle D. Indirect evidence: indirect treatment comparisons in meta-analysis. Canadian Agency for Drugs and Technologies in Health, 2009.
- Act on the Reform of the Market for Medicinal Products (Gesetz zur Neuordnung des Arzneimittelmarktes, AMNOG). Statute 35a (1), S. 6+7, SGB V.
- Fu R, Gartlehner G, Grant M, Shamliyah T, Sedrakyan A, Wilt TJ, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *J Clin Epidemiol* 2011;64:1187-97.
- Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252-60.
- Coleman CI, Phung OJ, Cappelleri JC, Baker WL, Kluger J, White CM, et al. Use of mixed treatment comparisons in systematic reviews. Methods Research Report. (Prepared by the University of Connecticut/Hartford Hospital Evidence-based Practice Center under

- Contract No 290-2007-10067-I.) AHRQ Publication No 12-EHC119-EF. Agency for Healthcare Research and Quality, 2012.
- Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *Int J Epidemiol* 2013;42:332-45.
  - Brier G. Verification of forecasts expressed in terms of probability. *Monthly Weather Review* 1950;78:1-3.
  - Boland A, Dunder Y, Bagust A, Haycox A, Hill R, Mujica Mota R, et al. Early thrombolysis for the treatment of acute myocardial infarction: a systematic review and economic evaluation. *Health Technol Assess* 2003;7:1-136.
  - Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64:163-71.
  - Pereira TV, Horwitz RJ, Ioannidis JP. Empirical evaluation of very large treatment effects of medical interventions. *JAMA* 2012;308:1676-84.
  - Ara R, Pandor A, Stevens J, Rees A, Rafia R. Early high-dose lipid-lowering therapy to avoid cardiac events: a systematic review and economic evaluation. *Health Technol Assess* 2009;13:1-74, 75-118.
  - Bansback N, Sizzo S, Sun H, Feldman S, Willian MK, Anis A. Efficacy of systemic treatments for moderate to severe plaque psoriasis: systematic review and meta-analysis. *Dermatology* 2009;219:209-18.
  - Brown TJ, Hooper L, Elliott RA, Payne K, Webb R, Roberts C, et al. A comparison of the cost-effectiveness of five strategies for the prevention of non-steroidal anti-inflammatory drug-induced gastrointestinal toxicity: a systematic review with economic modelling. *Health Technol Assess* 2006;10:iii-iv, xi-xiii, 1-183.
  - Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet* 2009;373:746-58.
  - Cooper NJ, Sutton AJ, Lu G, Khunti K. Mixed comparison of stroke prevention treatments in individuals with nonrheumatic atrial fibrillation. *Arch Intern Med* 2006;166:1269-75.
  - Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29:932-44.
  - Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. *Lancet* 2007;369:201-7.
  - Golfinopoulos V, Salanti G, Pavlidis N, Ioannidis JP. Survival and disease-progression benefits with treatment regimens for advanced colorectal cancer: a meta-analysis. *Lancet Oncol* 2007;8:898-911.
  - Imamura M, Abrams P, Bain C, Buckley B, Cardozo L, Cody J, et al. Systematic review and economic modelling of the effectiveness and cost-effectiveness of non-surgical treatments for women with stress urinary incontinence. *Health Technol Assess* 2010;14:1-188, iii-iv.
  - Kyrgiou M, Salanti G, Pavlidis N, Paraskevaidis E, Ioannidis JP. Survival benefits with diverse chemotherapy regimens for ovarian cancer: meta-analysis of multiple treatments. *J Natl Cancer Inst* 2006;98:1655-63.
  - Lam SK, Owen A. Combined resynchronisation and implantable defibrillator therapy in left ventricular dysfunction: Bayesian network meta-analysis of randomised controlled trials. *BMJ* 2007;335:925.
  - Mauri D, Polyzos NP, Salanti G, Pavlidis N, Ioannidis JP. Multiple-treatments meta-analysis of chemotherapy and targeted therapies in advanced breast cancer. *J Natl Cancer Inst* 2008;100:1780-91.
  - Mills EJ, Druyts E, Ghement I, Puhon MA. Pharmacotherapies for chronic obstructive pulmonary disease: a multiple treatment comparison meta-analysis. *Clin Epidemiol* 2011;3:107-29.
  - Mills EJ, Wu P, Lockhart I, Thorlund K, Puhon M, Ebbert JO. Comparisons of high-dose and combination nicotine replacement therapy, varenicline, and bupropion for smoking cessation: a systematic review and multiple treatment meta-analysis. *Ann Med* 2012;44:588-97.
  - Psaty BM, Smith NL, Siscovick DS, et al. Health outcomes associated with antihypertensive therapies used as first-line agents. A systematic review and meta-analysis. *JAMA* 1997;277:739-45.
  - Sciarretta S, Palano F, Tocci G, Baldini R, Volpe M. Antihypertensive treatment and development of heart failure in hypertension: a Bayesian network meta-analysis of studies in patients with hypertension and high cardiovascular risk. *Arch Intern Med* 2011;171:384-94.
  - Thijs V, Lemmens R, Fieuws S. Network meta-analysis: simultaneous meta-analysis of common antiplatelet regimens after transient ischaemic attack or stroke. *Eur Heart J* 2008;29:1086-92.
  - Wang H, Huang T, Jing J, Jin J, Wang P, Yang M, et al. Effectiveness of different central venous catheters for catheter-related infections: a network meta-analysis. *J Hosp Infect* 2010;76:1-11.
  - Cook DJ, Reeve BK, Guyatt GH, Heyland DK, Griffith LE, Buckingham L, et al. Stress ulcer prophylaxis in critically ill patients. Resolving discordant meta-analyses. *JAMA* 1996;275:308-14.
  - Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997;156:1411-6.
  - Linde K, Willich SN. How objective are systematic reviews? Differences between reviews on complementary medicine. *J R Soc Med* 2003;96:17-22.
  - Peinemann F, McGauran N, Sauerland S, Lange S. Disagreement in primary study selection between systematic reviews on negative pressure wound therapy. *BMC Med Res Methodol* 2008;8:41.
  - Poolman RW, Abouali JA, Conter HJ, Bhandari M. Overlapping systematic reviews of anterior cruciate ligament reconstruction comparing hamstring autograft with bone-patellar tendon-bone autograft: why are they different? *J Bone Joint Surg* 2007;89:1542-52.
  - Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ* 2013;347:14501.
  - Thorlund K, Druyts E, Avina-Zubieta JA, Wu P, Mills EJ. Why the findings of published multiple treatment comparison meta-analyses of biologic treatments for rheumatoid arthritis are different: an overview of recurrent methodological shortcomings. *Annals Rheum Dis* 2013;72:5124-35.
  - Gartlehner G, Gaynes BN, Hansen RA, Thieda P, DeVeau-Geiss A, Krebs EE, et al. Comparative benefits and harms of second-generation antidepressants: background paper for the American College of Physicians. *Ann Intern Med* 2008;149:734-50.
  - Ioannidis JP. Ranking antidepressants. *Lancet* 2009;373:1759-60; author reply 61-2.
  - Trinquart L, Abbé A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. *PLoS One* 2012;7:e35219.

**What is already known on this topic**

Network meta-analysis is an increasingly popular method that allows the comparative effectiveness of multiple treatments to be evaluated. Examples have started to accumulate where network analyses on the same topic have reached different conclusions based on the exclusion of treatment nodes.

**What this study adds**

It is common for networks to exclude trials and specific comparators based on widely diverse criteria. Excluding treatments from a network meta-analysis can importantly change effect estimates and the probability rankings of being the best treatment. Well connected treatments that are unlikely to be the best treatment are the most likely to be influential.

- 42 Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008;17:279-301.
- 43 Cure S, Diels J, Gavart S, Bianic F, Jones E. Efficacy of telaprevir and boceprevir in treatment-naïve and treatment-experienced genotype 1 chronic hepatitis C patients: an indirect comparison using Bayesian network meta-analysis. *Curr Med Res Opin* 2012;28:1841-56.

**Accepted:** 4 August 2013

Cite this as: [BMJ 2013;347:f5195](#)

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

## Tables

Table 1 | Study characteristics

Network	Type of treatments	Condition/outcome	Consideration/exclusion of specific treatment comparators	No of studies (nodes)	No of loops	Effect measure	Connectivity (min/max)†
Ara 2009 <sup>14</sup>	Statins	Hypercholesterolaemia/discontinuation due to adverse events	Only trials involving high dose statins or low dose statins	12 (5)	5	Odds ratio	2/4
Bansback 2009 <sup>15</sup>	Biologicals	Moderate to severe plaque psoriasis	All options considered	22 (8)	2	Relative risk*	1/3
Brown 2006 <sup>16</sup>	Pharmacological interventions	NSAID induced gastrointestinal toxicity/serious gastrointestinal complications	Only trials involving specific strategies were considered	40 (6)	6	Relative risk*	2/4
Cipriani 2009 <sup>17</sup>	Antidepressants	Unipolar major depression in adults/response to treatment	Only trials involving second generation antidepressants were considered (not placebo)	111 (12)	70	Odds ratio	3/11
Cooper 2006 <sup>18</sup>	Anticoagulants and antiplatelet therapy	Non-rheumatic atrial fibrillation/ischaemic stroke prevention	All options considered	19 (9)	11	Rate ratio	1/7
Dias 2010 <sup>19</sup>	Thrombolytic drugs and angioplasty	Acute myocardial infarction/death	All options considered (previous analysis excluded two treatments not licensed in the UK)	50 (9)	11	Odds ratio	1/7
Elliot 2007 <sup>20</sup>	Antihypertensive drugs	Hypertension, high risk patients/patients who developed diabetes	All options considered	22 (6)	16	Odds ratio	4/5
Golfinopoulos 2007 <sup>21</sup>	Chemotherapy regimens	Colorectal cancer	Excluded best supportive care, radiotherapy, hormonal therapy, and gene therapy	40 (11)	7	Hazard ratio	1/7
Imamura 2010 <sup>22</sup>	Non-surgical	Stress urinary incontinence/cure	Only non-surgical options	38 (13)	26	Odds ratio	1/7
Kyrgiou 2006 <sup>23</sup>	Chemotherapy regimens	Ovarian cancer	Excluded best supportive care, radiotherapy, hormonal therapy, and gene therapy	61 (9)	10	Hazard ratio	1/6
Lam 2008 <sup>24</sup>	Implantable defibrillator	Left ventricular dysfunction	All options considered	11 (5)	3	Odds ratio	2/4
Mauri 2008 <sup>25</sup>	Chemotherapy regimens	Breast cancer	Excluded best supportive care, radiotherapy, hormonal therapy, and gene therapy	128 (22)	27	Hazard ratio	1/10
Mills 2011 <sup>26</sup>	COPD drugs	COPD/exacerbations	All options considered	25 (10)	7	Rate ratio	1/6
Mills 2012 <sup>27</sup>	Nicotine replacement therapy	Smoking addiction/cessation	All options considered	100 (6)	4	Odds ratio	1/4
Psaty 1997 <sup>28</sup>	Antihypertensive therapy	Coronary heart disease/fatal and non-fatal events	Only the six most commonly used antihypertensives	28 (7)	10	Relative risk*	1/5
Sciarretta 2011 <sup>29</sup>	Antihypertensive therapy	Heart failure/prevention of heart failure	Excluded trials before 1997	26 (8)	13	Odds ratio	1/6
Thijs 2008 <sup>30</sup>	Antiplatelets	Transient ischaemic attack or stroke	Excluded oral anticoagulants, trifusal, cilostazol, dipyridamole alone, and glycoprotein IIb/IIIa inhibitors	24 (5)	3	Odds ratio	2/4
Wang 2010 <sup>31</sup>	Central venous catheters	Catheter related infections	Excluded self preparing coated catheters	43 (9)	4	Odds ratio	1/4

NSAID=non-steroidal anti-inflammatory drug; COPD=chronic obstructive pulmonary disease.

\*Our analyses use odds ratios to avoid issues of non-transitivity with risk ratios.

†Connectivity reports the smallest and largest number of connections in each network.

**Table 2 | Nodes leading to greatest change in estimated effects and average fold changes in estimated treatment effects when single treatment nodes are removed**

Network	Node leading to largest Brier score*	Connectivity of removed node	Removal of largest expected impact node		Removal of other active intervention nodes	
			Average fold change	Largest fold change	Average fold change	Largest fold change
Change in top three ranked treatments:						
Brown 2006 <sup>16</sup>	Cox 2 preferentials (20%)	3 (max 4)	1.83	3.13	1.38	4.40
Cipriani 2009 <sup>17</sup>	Fluoxetine (0%)	11 (max)	1.06	1.18	1.03	1.16
Cooper 2006 <sup>18</sup>	Fixed low dose warfarin+aspirin (0%)	4 (max 7)	1.25	2.65	1.06	1.60
Golfinopoulos 2007 <sup>21</sup>	Fluorouracil (0%)	5 (max 7)†	1.03	1.05	1.01	1.09
Imamura 2010 <sup>22</sup>	Pelvic floor muscle training+biofeedback (5%)	6 (max 7)	1.22	1.40	1.07	2.18
Mauri 2008 <sup>25</sup>	Standard dose anthracycline (0%)	10 (max)	1.05	1.12	1.01	1.09
Mills 2012 (6 mth) <sup>27</sup>	Low dose nicotine replacement therapy (0%)	4 (max)	1.12	1.39	1.01	1.03
Sciarretta 2011 <sup>29</sup>	Calcium channel blockers (1%)	6 (max)	1.27	1.53	1.04	1.31
Wang 2010 <sup>31</sup>	Chlorhexidine and silver sulfadiazine (0%)	4 (max)	1.27	4.64	1.01	1.08
No changes in top three ranked treatments:						
Ara 2009 <sup>14</sup>	Simvastatin 40 mg (21%)	4 (max)	1.13	1.19	1.04	1.11
Bansback 2009 <sup>15</sup>	Adalimumab (0%)	2 (max 3)	1.08	1.32	1.01	1.02
Dias 2010 <sup>19</sup>	Accelerated alteplase (0%)	7 (max)	1.08	1.16	1.03	1.07
Elliot 2007 <sup>20</sup>	Angiotensin converting enzyme inhibitors (22%)	4 (max 5)	1.08	1.12	1.02	1.06
Kyrgiou 2006 <sup>23</sup>	Platinum monotherapy (0%)	6 (max)	1.08	1.25	1.02	1.13
Lam 2008 <sup>24</sup>	Implantable cardioverter defibrillator (9%)	3 (max 4)†	1.03	1.07	1.01	1.04
Mills 2011 <sup>26</sup>	Inhaled corticosteroids (0%)	3 (max 6)	1.01	1.02	1.00	1.01
Psaty 1997 <sup>28</sup>	Diuretics (10%)	5 (max)	1.06	1.14	1.03	1.08
Thijs 2008 <sup>30</sup>	Aspirin (0%)	4 (max)	1.03	1.04	1.01	1.03

\*Probability of being best treatment.  
†Most connected node that was removed.

Table 3| Example of fold change calculation and change in rank probabilities for Cooper 2006<sup>18</sup> network

Treatment	Full model		Reduced model		Fold change*
	Relative rate v placebo (95% CI)	Probability (%)	Relative rate v placebo (95% CI)	Probability (%)	
Warfarin:					
Adjusted standard dose	0.37 (0.26 to 0.53)	3	0.41 (0.27 to 0.7)	2	1.11
Adjusted low dose	0.34 (0.19 to 0.58)	13	0.34 (0.19 to 0.60)	41	1.00
Fixed low dose	0.76 (0.3 to 1.76)	1	0.93 (0.36 to 2.66)	1	1.22
Aspirin	0.62 (0.43 to 0.86)	0	0.66 (0.47 to 1.01)	0	1.06
Fixed low dose warfarin+aspirin	0.98 (0.6 to 1.67)	0	Removed	Removed	NA
Ximelagatran	0.34 (0.18 to 0.62)	14	0.37 (0.21 to 0.83)	13	1.09
Alternate day aspirin	0.17 (0.01 to 1.15)	66	0.45 (0.03 to 18.48)	38	2.65
Indobufen	0.46 (0.19 to 1.14)	5	0.54 (0.23 to 1.57)	6	1.17

NA=not applicable.

\*Average fold change reported in table 2 is averaged on log scale. The average of these fold changes is 1.33, slightly higher than 1.25.

**Table 4 | Average fold changes in treatment effects and impact on treatment ranks and probability of best treatment with clinically relevant removals**

Network	Nodes deleted	Average fold change	Largest fold change	Change in top three treatment ranks	Change in probability of best treatment	
					Full model	Reduced model
Changes after removal of placebo/no treatment nodes:						
Ara 2009 <sup>14</sup>	Placebo	1.03	1.05	None	Rosuvastatin (53%)	Rosuvastatin (56%)
Brown 2006 <sup>16</sup>	Placebo	3.10	4.44	None	Cox 2 inhibitors (41%)	Cox 2 inhibitors (43%)
Cooper 2006 <sup>18</sup>	Placebo	1.16	1.50	None	Alternate day aspirin (66%)	Alternate day aspirin (79%)
Elliot 2007 <sup>20</sup>	Placebo	1.02	1.06	None	Angiotensin receptor blocker (77%)	Angiotensin receptor blocker (83%)
Imamura 2010 <sup>22</sup>	No treatment	1.21	1.77	None	Pelvic floor muscle training+biofeedback+bladder training (61%)	Pelvic floor muscle training+biofeedback+BT (56%)
Mills 2011 <sup>26</sup>	Placebo	1.01	1.02	None	Roflumilast+long acting muscarinic antagonists (49%)	Roflumilast+long acting muscarinic antagonists (56%)
Mills 2012 <sup>27</sup>	Placebo	1.22	1.40	None	High dose nicotine replacement therapy (58%)	High dose nicotine replacement therapy (90%)
Psaty 1997 <sup>28</sup>	Placebo	1.01	1.03	None	Angiotensin converting enzyme inhibitors (55%)	Angiotensin converting enzyme inhibitors (43%)
Sciarretta 2011 <sup>29</sup>	Placebo	1.03	1.06	Top two switch	β blockers (29%)	Angiotensin converting enzyme inhibitors (32%)
Thijs 2008 <sup>30</sup>	Placebo	1.01	1.03	None	Aspirin+dipyridamole (80%)	Aspirin+dipyridamol (70%)
Changes after removal of situation specific relevant treatment nodes						
Dias 2010 <sup>19</sup>	ASPAC and urokinase	1.04	1.06	None	Percutaneous transluminal coronary angioplasty (83%)	Percutaneous transluminal coronary angioplasty (96%)
Golfinopoulos 2007 <sup>21</sup>	Monotherapies: fluorouracil, oxaliplatin, irinotecan	1.06	1.18	None	Fluorouracil and leucovorin+irinotecan+bevacizumab (63%)	Fluorouracil and leucovorin+irinotecan+bevacizumab (48%)
Kyrgiou 2006 <sup>23</sup>	Platinum monotherapy and non-platinum/non-taxane drugs	1.21	1.68	None	Platinum+taxane based combination (ip) (92%)	Platinum+taxane based combination (ip) (89%)
Mauri 2008 <sup>25</sup>	All drugs deemed miscellaneous	1.05	1.17	3rd becomes 1st	Novel non-taxane agents+taxanes (24%)	Taxanes (combination regimen) (21%)

APSAC=anisoylated plasminogen streptokinase activated complex; ip=including at least one agent given intraperitoneally.

BMJ: first published as 10.1136/bmj.f5195 on 5 September 2013. Downloaded from https://www.bmj.com/ on 24 November 2025 by guest. Protected by copyright, including for uses related to text and data mining, AI training, and similar technologies.

