

# The development of network meta-analysis

**Andrew Lee**

University of Oxford, UK

**Corresponding author:** Andrew Lee. Email: andrewlee1@nhs.net

As a clinician, I have been keen to apply the principles of evidence-based medicine, informed by evidence from systematic reviews and the specific needs of each of my patients. However, in my attempts to practise evidence-based medicine, I have often had to confront two frustrations. First, many clinical decisions involve more than two possible interventions, yet systematic reviews are usually based only on comparisons in pairs of interventions. Second, the comparisons researched do not cover all the relevant interventions.

Many others have experienced these frustrations. For example, after assessing 18 trials testing 24 interventions for children with acute pyelonephritis, John Ioannidis<sup>1</sup> asked: ‘*How do we make sense of this complex network and guess the best choice(s)?*’

It was in 2009 that I first became aware of an approach to addressing quandaries resulting from such complex networks. The solution involves combining direct and indirect treatment comparisons using evidence from randomised trials to produce a synthesis linking data from all the interventions in a ‘network’. In doing this, each trial shares at least one direct comparison with another trial. Analysis of the network – network meta-analysis (NMA) – could then produce a ranking of the likelihood of each of the interventions being the most effective.

## NMA as part of the development of research synthesis

NMA is a method of research synthesis, so any history of this specific method must acknowledge that it is just one part of the overall history of research synthesis. That wider history has been described by others.<sup>2,3</sup> This article focuses on the origins and development of this method of synthesis, which has also been described as a ‘mixed treatment comparison’ or ‘multiple treatments comparison’.<sup>4</sup>

## The concept of NMA

NMA is an extension of traditional, pairwise meta-analysis, but three main advantages are claimed for

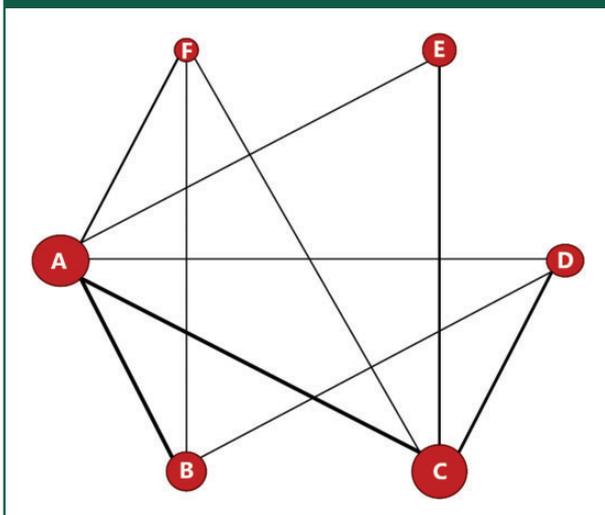
NMA. First, it allows more than two interventions to be compared simultaneously; second, interventions can be compared even if they have not been directly compared in trials; and third, it increases precision of the estimate of effect size by ‘borrowing strength’<sup>5</sup> if the key assumptions described below are valid.

When seeking to compare the effects of interventions A, B and C, if there are trials of A versus B and B versus C, but not of A versus C, an estimate of the relative effects of A compared to C can be made by indirect comparison using the results of the trials of A versus B and B versus C. This indirect comparison uses the results of meta-analysis of all trials in each direct comparison, as explained by Bucher et al.,<sup>6</sup> to combine, in a single synthesis, all the available direct comparisons and all the indirect comparisons that can be estimated between the selected interventions.

A network diagram can be drawn illustrating the direct comparisons between A, B and C. In more complex networks, multiple combinations of direct comparisons are used to estimate the same indirect comparison, which are then incorporated in the synthesis. These networks might have more than three interventions and have varying complexity of geometry. Among the simplest is a star, in which the intervention at the centre is the common comparator to all the other interventions. More complex forms of network geometry, when there are common comparators for only a few of a more diverse range of interventions, involve multiple conjoined loops and sidearms.

Figure 1 shows an example of a network diagram from an NMA comparing six interventions, A to F. The solid lines show where trial results of direct comparisons exist. In this example, there is no single common comparator for all the interventions, but each intervention shares at least one comparator with another intervention in the network. The thickness of each connecting line represents the number of trials between the pair of interventions at either end of the line, and the size of each node (letter within a circle) represents the number of participants receiving the intervention. Some network diagrams also

Figure 1. Network diagram example.



include the actual numbers of trials, participants or both.

Transitivity and consistency are two key assumptions for these analyses. Assessing the validity of these requires input of both clinical and methodological expertise. Transitivity is the assumption that effect modifiers (the clinical and methodological characteristics that can affect the outcome) are similar in each direct comparison involving the same intervention.<sup>7</sup> It cannot be assessed statistically but requires critical interpretation of the effect modifiers in the trials whose results are considered for synthesis. Before an NMA is conducted, those trials must be assessed for significant differences in their populations, interventions, outcomes, methodological features and reporting.<sup>8</sup> It is also important to note that even placebo response has been found to vary over time, which might affect the transitivity assumption when placebo is a common comparator.<sup>9</sup>

Consistency is an extension of transitivity. It is the assumption of agreement between the results of direct and indirect comparisons for each pair of interventions. This can be assessed statistically but only when there are both direct and indirect comparisons of one or more pairs of interventions within a network, known as 'closed loops'. The assumption of transitivity needs to be reconsidered if inconsistency is detected. If inconsistency is not detected statistically, however, that does not automatically validate the transitivity assumption.<sup>7</sup>

### Origins and evolution of methods for NMA

In 1989, Eddy described the 'confidence profile method' (CPM),<sup>10</sup> and the publication by Eddy

et al.<sup>11</sup> in the following year, 'A Bayesian method for synthesizing evidence', described: 'a collection of meta-analysis techniques based on Bayesian methods for interpreting, adjusting, and combining evidence to estimate parameters and outcomes important to the assessment of health technologies.'

These techniques were collectively called the 'confidence profile method' and the article explained indirect comparison with the following example:

The approach is to use the available evidence to derive probability distributions for the various pairs that have been directly compared. A distribution for the relative effects of other pairs can then be calculated by a series of convolutions. The concept is illustrated by calculating the difference between the test scores of Tom and Bill from knowledge of the differences in scores between Tom and George, and George and Bill.

The practical application of the methodology was supported by software to conduct the synthesis. Eddy initially used SOFT\*PRO, but the commonest software in published systematic reviews reporting NMA<sup>4</sup> is WinBUGS (initially BUGS).<sup>12</sup> This uses Bayesian methods, and such methods are predominant in both software use and methodology developments in NMA, but frequentist approaches and software are also used.<sup>13,14</sup>

A series of methodology publications through the 1990s built on the CPM approach, including notably those by Smith et al.<sup>15</sup> and by Higgins and Whitehead.<sup>5</sup> Higgins and Whitehead wrote about borrowing strength from external trials in a meta-analysis. They argued that

Many meta-analysis papers include data from three or more treatments, but only consider pairwise comparisons of, say treatment A with control and treatment B with control. There would seem to be little reason not to combine all treatments into one analysis.

They used the BUGS software to combine the results of trials comparing the effects of beta-blockers versus sclerotherapy, beta-blockers versus control and sclerotherapy versus control on preventing cirrhosis-related bleeding. Higgins' and Whitehead's paper<sup>5</sup> was described later by Salanti and Schmid<sup>16</sup> as:

the first to articulate that relative effects of different treatments can be jointly estimated in a single meta-analysis model to improve power. This landmark paper introduced the basis for the methodology

which, now extended and refined, is increasingly known as network meta-analysis.

In 2002, Lumley published 'Network meta-analysis for indirect treatment comparisons' in which he '*presented methods of estimating treatment differences between treatments that have not been directly compared in a randomized trial, and, more importantly, methods of estimating the uncertainty in these differences.*'<sup>17</sup>

Lumley acknowledged the limitation of his methods, which were restricted to each trial only having two intervention groups '*Meta-analyses with large numbers of multi-armed trials present difficulties for network meta-analysis, and extensions to handle multi-armed trials correctly should be investigated.*'

Ades subsequently described methods to encompass multi-arm trials and multiple outcomes. In his 2003 article,<sup>18</sup> he stated

The aim of nearly all meta-analysis has been to summarize evidence comparing one or sometimes more treatments. Usually only a single outcome is examined, and if there is more than one outcome these are explored in separate meta-analyses, rather than simultaneously. This paper concerns the possibility of combining information from different studies on different, but structurally related, outcomes, and using the data to construct a single model which expresses the relationships between the different kinds of data.

The following year, together with Lu, Ades published 'Combination of direct and indirect evidence in mixed treatment comparisons',<sup>19</sup> which was the most frequently cited origin of the current NMA methodology in a survey of published network meta-analyses in 2014.<sup>4</sup> Lu and Ades extended the model proposed by Smith et al.<sup>15</sup> to encompass trials with more than two intervention groups.

A review of the methods for NMA, with particular emphasis on the issue of inconsistency between direct and indirect evidence, was published in 2008 by Salanti et al.<sup>20</sup> They explained that inconsistency in estimates of intervention effects obtained from direct and indirect comparisons may indicate diversity, bias or a combination of both, and they described modelling to test for consistency. Their review considered potential sources of inconsistency, including genuine diversity in the characteristics of included trials, selection bias, study quality and sponsorship bias. It stresses the importance of planning in advance for investigation of inconsistency, because clinical and epidemiological assessment of inconsistency may be difficult because of factors such as reporting

deficiencies or lack of sufficient studies for some comparisons. Salanti et al.<sup>20</sup> also highlighted that attention to the geometry (the overall pattern of comparisons among interventions) and the asymmetry of networks (the extent to which specific comparisons of interventions are represented more heavily than others in the number of included trials or participants) can be used to inform the design of the new trials that would most usefully add to the overall network.

A review of NMA methods, published in 2016 by Efthimiou et al.,<sup>8</sup> summarised newer publications on the use of NMA methods. This included various models for performing NMA, statistical methods for assessing inconsistency, software options, investigating sources of potential bias and reporting results.

The use of individual participant data (IPD) in meta-analyses has many advantages over the use of aggregate data, including improving the quantity and quality of data, which has resulted in it being considered '*the gold standard in evidence synthesis.*'<sup>21</sup> The use of this approach, initially using traditional, pairwise meta-analysis methods, increased between the early 1990s and 2008 to around 50 publications per year.<sup>22</sup> The number of systematic reviews using IPD was found to be 10 to 22 per year with no discernible growth trend in the years leading up to 2015.<sup>23</sup> Gao et al. found that the first IPD using NMA methods was published in 2007 and that 21 IPDs using NMA methods had been published by June 2019.<sup>24</sup> There are limitations as well as advantages to use of IPD. Guidance has been published on the best use of IPD meta-analysis generally<sup>25</sup> and specifically on the use of NMA methods with IPD.<sup>21</sup>

Multiple outcomes multivariate meta-analysis (MOMA) is another approach to meta-analysis that has been increasing in recent years.<sup>26</sup> Relevant studies that might be considered for synthesis may not report the same outcomes, which could result in their exclusion from traditional meta-analyses, but MOMA allows for inclusion where outcomes can be regarded as highly correlated. Guidance on conducting this type of synthesis using NMA methods has been published in recent years,<sup>26-28</sup> including the use of IPD.<sup>29</sup>

Interest has developed recently in creating and maintaining continuously updated meta-analyses using NMA methods<sup>30</sup> and a major project of this kind for COVID-19-related interventions began in 2020.<sup>31</sup>

## Approaches to conducting NMA

A simple meta-regression approach can be used for NMA if there is no multi-arm trial in the network.<sup>32</sup>

However, if the network includes multi-arm trials, other methods are more appropriate. Bayesian methods have been used most frequently,<sup>33</sup> partly because this approach can most naturally produce estimates of ranking probabilities for the interventions being compared (to give the probability that each intervention is most effective through to least effective)<sup>34</sup>; but frequentist methods to approximate ranking have also been described.<sup>35</sup> The hierarchical model approach is detailed by Lu and Ades<sup>19</sup> and by Salanti et al.<sup>20</sup>

An alternative approach is multivariate meta-analysis, which can be conducted using Bayesian<sup>36</sup> or frequentist methods.<sup>37</sup> A further approach, based on graph-theoretical methods, has been described by Rucker.<sup>38</sup>

The frequentist approach assumes that the intervention effect has a true value with a confidence interval, which defines the range within which the true value would fall with a minimum probability, usually 95%. The Bayesian approach assumes that the intervention effect has a fixed value but within a probability distribution based on a 'prior', which might be a value chosen from existing evidence or might be a 'best guess'. The credible interval results of a Bayesian meta-analysis provide the probability of the range of values within which the fixed value lies, given the data, and this range is the 'posterior' that includes 95% of the probability.

### Use of network meta-analyses in published systematic reviews

In 1999, Dominici et al.<sup>39</sup> used Bayesian methods and data from 46 trials of treatments to prevent migraine headache, to produce a ranking of treatments. They stated their aims as follows:

In this article we present a meta-analysis of these 46 trials with the goal of synthesizing existing evidence about which treatments are most effective and of quantifying the remaining uncertainty about treatment effectiveness. We hope that the results and methods will be useful in supporting clinical treatment decisions and will help guide the planning of new trials. The critical statistical aspects of this goal are the estimation of treatment effects on a common scale and the relative ranking of treatments, both within classes and overall. This requires indirect comparisons among treatments that may never have been tested together in the same trial.

Dominici et al. used data collected by another team working on a systematic review. The first published

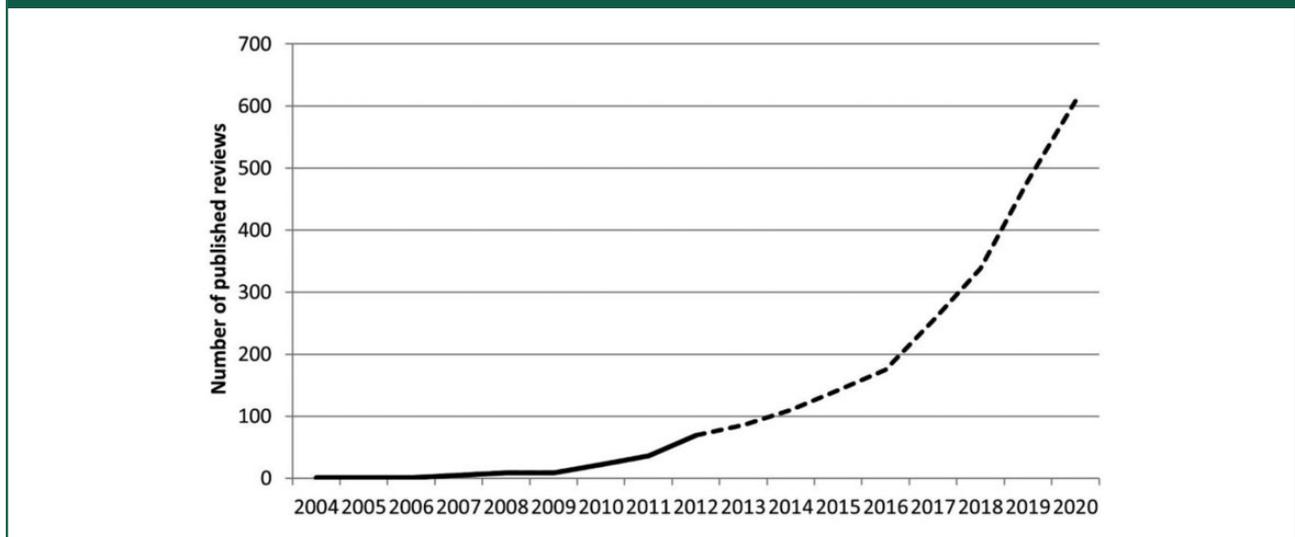
report of an NMA as part of a systematic review conducted by the authors, appears to have been that reported by Psaty et al.<sup>40</sup> Their explanation for selecting this method to study treatments for high blood pressure summarises well the commonest problem in using only pairwise meta-analysis to inform clinical decision making:

The clinical trials in hypertension have provided a patchwork of evidence about the health benefits of antihypertensive agents. Some trials used placebo or untreated controls, and others used active-treatment comparison groups. Among the latter, the choice of treatment and comparison therapies has varied from one trial to the next. Several approaches to the synthesis of these complex data are possible. The Blood Pressure Trialists, for instance, conducted a prospective series of mini-meta-analyses, but this method left many 'unresolved issues' due to multiple comparisons and low power. In this study, we used a new technique, called network meta-analysis, to synthesize the available evidence from placebo-controlled and comparative trials in a single meta-analysis.

Subsequent uses of NMA were initially slow to appear in the healthcare literature. Edwards et al.<sup>41</sup> identified only seven published systematic reviews reporting NMA up to July 2007. Publications rose to an estimated 90 to 100 published in 2012<sup>4</sup> and an estimated 180 to 200 in 2018,<sup>42</sup> with a total of more than 1000 now available in the literature. The increasing trend of publication of network meta-analyses is illustrated in Figure 2. The solid line indicates the number of publications for each year up to 2012 identified by a review<sup>4</sup>; the dashed line extends this to 2020 using the results of Medline searches for each year. This trajectory of increasing publications is similar to that for publications reporting use of traditional, pairwise meta-analysis between 1980 and 2000.<sup>43</sup>

The most significant impact of evidence synthesis on healthcare is likely to come from the use of the evidence generated by these research projects in national clinical guidelines. A review of NICE clinical guidelines published or updated in 2015 and 2016 found that they made extensive use of meta-analysis to identify evidence to support their recommendations.<sup>42</sup> NMA methods were used far less often than traditional, pairwise meta-analysis but were used or considered for nearly one-quarter of the guidelines reviewed, showing that evidence produced using network meta-analysis methods is influencing recommendations for the UK National Health Service.

**Figure 2.** Estimate for the number of published network meta-analyses 2004 to 2020.



## Guidance relating to conduct and reporting of NMA

Efforts to improve the quality of reporting of systematic reviews have been based on defining and promoting internationally recognised standards. There is evidence that publication of such reporting standards results in improved quality of reporting, based on comparisons before and after the publication of both the QUOROM statement<sup>44</sup> and the PRISMA Statement,<sup>45</sup> and that reporting to PRISMA statement standards is strongly associated with higher study quality, as assessed by a widely used critical appraisal tool (CAT).<sup>46</sup> Panic et al.<sup>47</sup> found that endorsement of the PRISMA statement by journals in their instructions for authors was associated with improved quality of reporting, regardless of whether the authors declared that they had followed the statement and was associated with higher study quality.

When the first standards for reporting systematic reviews and meta-analyses were published in 1999 in the QUORUM Statement, the concept of NMA was still in its infancy. However, by the time of the PRISMA Statement in 2009, NMA warranted a mention as a form of meta-analysis that combined direct and indirect comparisons. However, the PRISMA 2009 statement did not make recommendations relating to reporting that addressed the specifics of NMA methodology. For standards for conducting systematic reviews, it directed readers to the guidance published by the Cochrane Collaboration<sup>48</sup> and the Centre for Reviews and Dissemination,<sup>49</sup> both of which contained very limited guidance on the use of

NMA methodology in systematic reviews. Reporting standards for use of NMA methods were subsequently published as a PRISMA extension statement in 2015.<sup>50</sup>

The Methods Guide 2008 Update of the National Institute for Health and Care Excellence (NICE)<sup>51</sup> included a section on NMA for the first time and the NICE Decision Support Unit's Evidence Synthesis TSD series,<sup>52</sup> published initially in 2011, expanded on that guidance. Also in 2011, the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) published reports on the interpretation and conduct of NMA.<sup>13,14</sup> Ades<sup>53</sup> observed that this '*seems to represent the first position statement from an academic body on these methods*'.

The Cochrane Comparing Multiple Interventions Methods Group was established in 2010 and has since produced guidance on the use of NMA methods in Cochrane Reviews and promoted training in using these methods. Version 6 of the Cochrane Handbook for Systematic Reviews of Interventions, published in 2019, contained '*a major new core chapter*' addressing for the first time NMA within the handbook. This guidance emphasises that NMA is '*more statistically complex than a standard meta-analysis*', consequently, '*close collaboration*' between a statistician with expertise in NMA methods and those with expertise in the clinical content area is essential in the design and conduct of a review to ensure that studies selected for inclusion in NMA fulfil the assumptions of transitivity and consistency.<sup>54</sup>

## Critical appraisal of NMA

Critical appraisal involves assessing the report of a study for methodological quality and any likelihood of bias and considering whether these might affect the validity of the reported results. A 2018 review of the published CATs for systematic reviews<sup>42</sup> found that none of the most widely used CATs for systematic reviews contained content specifically relevant to appraising research synthesis using the NMA methodology. Three tools had been published, which include content that is relevant to appraising the use of NMA methods.<sup>13,55,56</sup> These three tools, however, might not be suitable for end-users without specialist statistical knowledge of NMA methodology, so there is still potentially a role for a new CAT to support generalist end-users. A tool has been constructed using the CASP format,<sup>57</sup> which might form the basis for further development.<sup>42</sup>

Other approaches have been developed to assess confidence in the evidence produced using NMA methods. For example, in 2014, the GRADE Working Group reported guidance for establishing the quality of treatment effect estimates obtained from NMA.<sup>58</sup> Their approach involved rating the quality of each direct and indirect effect estimate for each pairwise comparison within the NMA and then rating the NMA effect estimate for each pairwise comparison. In 2018, the GRADE Working Group recommended modifications to their 2014 guidance with a view to making the process more efficient, acknowledging that the original approach, 'may appear onerous in networks with many interventions'.<sup>59</sup>

In 2014, Salanti et al. published a modification of the GRADE guidance, in particular, drawing a distinction between rating the effect estimates for each pairwise comparison and rating the ranking of all the interventions within a network.<sup>60</sup> More recently, in 2020, Salanti and others published a new approach: Confidence in Network Meta-Analysis (CINeMA).<sup>61</sup> They stated that uptake of the earlier 2014 GRADE system and that reported by Salanti and others in the same year had been limited by '*the complexity of the methods and the lack of suitable software*'.

CINeMA is also based on the GRADE framework but instead of considering direct and indirect evidence for each pairwise comparison separately, it considers the impact of every study in the network. A web-based application makes it easy to apply to even large networks.<sup>62</sup>

A further approach, 'threshold analysis', has been developed specifically to assess confidence in NMA results used in guideline development.<sup>63</sup> The authors argue that their approach is needed because GRADE

approaches do not assess the influence of the NMA evidence on a resulting recommendation:

Threshold analysis quantifies precisely how much the evidence could change (for any reason, such as potential biases, or simply sampling variation) before the recommendation changes, and what the revised recommendation would be. If it is judged that the evidence could not plausibly change by more than this amount, then the recommendation is considered robust; otherwise, it is sensitive to plausible changes in the evidence.

These alternative approaches are primarily intended to be used by authors of systematic reviews and guideline developers (to rate confidence in an estimate of effect or confidence that such an estimate adequately supports a specific recommendation), but they are relevant to understanding by those undertaking critical appraisal and can be used to complement use of a CAT.

## Conclusions

NMA is a relatively new form of evidence synthesis and evolution of the methodology is continuing. As outlined by Salanti,<sup>64</sup> the introduction of NMA faced similar scepticism to that raised originally about traditional, pairwise meta-analysis. However, there has been gradually wider uptake of the method by researchers and use of the resulting evidence by decision makers in health and social care. The Cochrane review with the largest number of included studies is a NMA of 585 randomised trials of drugs to prevent postoperative nausea and vomiting.<sup>65</sup> Consensus guidance relating to the conduct and reporting of NMA is now widely available.

The use of NMA methods can overcome some limitations of traditional, pairwise meta-analysis. The design and conduct of NMAs require multidisciplinary input from expert methodologists and clinical topic experts. Further research is needed to clarify whether end-users who do not have specialist statistical knowledge can assess the quality and validity of evidence produced in systematic reviews using NMA methods, even with a critical appraisal tool optimised for such studies.

## Declarations

**Competing Interests:** None declared.

**Funding:** None declared.

**Ethics approval:** Not applicable.

**Guarantor:** AL.

**Contributorship:** Single author.

**Acknowledgements:** None.

**Provenance:** Not commissioned; invited article from the James Lind Library.

## References

- Ioannidis JP. Indirect comparisons: the mesh and mess of clinical trials. *Lancet* 2006; 368: 1470–1472.
- Chalmers I, Hedges LV and Cooper H. A brief history of research synthesis. *Eval Health Professions* 2002; 25: 12–37.
- Clarke M. History of evidence synthesis to assess treatment effects: personal reflections on something that is very much alive. JLL Bulletin: Commentaries on the history of treatment evaluation. See [www.jameslindlibrary.org/articles/history-of-evidence-synthesis-to-assess-treatment-effects-personal-reflections-on-something-that-is-very-much-alive/](http://www.jameslindlibrary.org/articles/history-of-evidence-synthesis-to-assess-treatment-effects-personal-reflections-on-something-that-is-very-much-alive/)
- Lee AW. Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. *J Clin Epidemiol* 2014; 67: 138–143.
- Higgins JP and Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996; 15: 2733–2749.
- Bucher HC, Guyatt GH, Griffith LE and Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; 50: 683–691.
- Cipriani A, Higgins JP, Geddes JR and Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Int Med* 2013; 159: 130–137.
- Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, et al.; GetReal Methods Review Group. GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* 2016; 7: 236–263.
- Julious SA and Wang S-J. How biased are indirect comparisons, particularly when comparisons are made over time in controlled trials? *Drug Inf J* 2008; 42: 625–633.
- Eddy DM. The confidence profile method: a Bayesian method for assessing health technologies. *Oper Res* 1989; 37: 210–228.
- Eddy DM, Hasselblad V and Shachter R. A Bayesian method for synthesizing evidence. The confidence profile method. *Int J Technol Assess Health Care* 1990; 6: 31–55.
- Lunn DJ, Thomas A, Best N and Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput* 2000; 10: 325–337.
- Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR task force on indirect treatment comparisons good research practices: part 1. *Value in Health* 2011; 14: 417–428.
- Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment comparison and network-meta-analysis studies: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 2. *Value in Health* 2011; 14: 429–437.
- Smith TC, Spiegelhalter DJ and Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Stat Med* 1995; 14: 2685–2699.
- Salanti G and Schmid CH. Research Synthesis Methods special issue on network meta-analysis: introduction from the editors. *Res Synth Methods* 2012; 3: 69–70.
- Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002; 21:2313–2324.
- Ades AE. A chain of evidence with mixed comparisons: models for multi-parameter synthesis and consistency of evidence. *Stat Med* 2003; 22: 2995–3016.
- Lu G and Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; 23: 3105–3124.
- Salanti G, Higgins JP, Ades AE and Ioannidis JP. Evaluation of networks of randomized trials. *Stat Meth Med Res* 2008; 17: 279–301.
- Debray TP, Schuit E, Efthimiou O, Reitsma JB, Ioannidis JP, Salanti G, et al.; GetReal Workpackage. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res* 2018; 27: 1351–1364.
- Riley RD, Lambert PC and Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010; 340: c221.
- Simmonds M, Stewart G and Stewart L. A decade of individual participant data meta-analyses: a review of current practice. *Contemp Clin Trials* 2015; 45: 76–83.
- Gao Y, Shi S, Li M, Luo X, Liu M, Yang K, et al. Statistical analyses and quality of individual participant data network meta-analyses were suboptimal: a cross-sectional study. *BMC Med* 2020; 18:120.
- Tierney JF, Vale C, Riley R, Smith CT, Stewart L, Clarke M, et al. Individual participant data (IPD) meta-analyses of randomised controlled trials: guidance on their use. *PLoS Med* 2015; 12: e1001855.
- Riley RD, Jackson D, Salanti G, Burke DL, Price M, Kirkham J, et al. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017; 358: j3932.
- Achana FA, Cooper NJ, Bujkiewicz S, Hubbard SJ, Kendrick D, Jones DR, et al. Network meta-analysis of multiple outcome measures accounting for borrowing of information across outcomes. *BMC Med Res Meth* 2014; 14: 92.
- Efthimiou O, Mavridis D, Cipriani A, Leucht S, Bagos P and Salanti G. An approach for modelling multiple correlated outcomes in a network of interventions using odds ratios. *Stat Med* 2014; 33: 2275–2287.

29. Riley RD, Price MJ, Jackson D, Wardle M, Gueyffier F, Wang J, et al. Multivariate meta-analysis using individual participant data. *Res Synth Methods* 2015; 6: 157–174.
30. Vandvik PO, Brignardello-Petersen R and Guyatt GH. Living cumulative network meta-analysis to reduce waste in research: a paradigmatic shift for systematic reviews? *BMC Med* 2016; 14: 59.
31. Boutron I, Chaimani A, Meerpohl JJ, Hróbjartsson A, Devane D, Rada G, et al.; COVID-NMA Consortium. The COVID-NMA project: building an evidence ecosystem for the COVID-19 pandemic. *Ann Int Med* 2020; 173: 1015–1017.
32. Cochrane Collaboration (2019). In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. (eds). *Cochrane Handbook for Systematic Reviews of Interventions Version 6.0* (updated July 2019). See [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook) (last checked 27 May 2022).
33. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH and Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. *PLoS One* 2014; 9: e86754.
34. Neupane B, Richer D, Bonner AJ, Kibret T and Beyene J. Network meta-analysis using R: a review of currently available automated packages. *PLoS One* 2014; 9: e115065.
35. White IR. Multivariate random-effects meta-regression: updates to Mvmeta. *Stata J* 2011; 11: 255–270.
36. Mavridis D and Salanti G. A practical introduction to multivariate meta-analysis. *Stat Methods Med Res* 2013; 22: 133–158.
37. White IR, Barrett JK, Jackson D and Higgins JP. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods* 2012; 3: 111–125.
38. Rucker G. Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* 2012; 3: 312–324.
39. Dominici F, Parmigiani G, Wolpert RL and Hasselblad V. Meta-analysis of migraine headache treatments: combining information from heterogeneous designs. *J Am Stat Assoc* 1999; 94: 16–28.
40. Psaty BN, Lumley T, Furberg CD, Schellenbaum G, Pahor M, Alderman MH, et al. Health outcomes associated with various antihypertensive therapies used as first-line agents – a network meta-analysis. *JAMA* 2003; 289: 2534–2544.
41. Edwards SJ, Clarke MJ, Wordsworth S and Borrill J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Practice* 2009; 63: 841–854.
42. Lee A. Developing critical appraisal of systematic reviews reporting network meta-analysis. DPhil, University of Oxford. See <https://ora.ox.ac.uk/objects/uuid:b9b878bd-fc04-4317-9df5-781930224f6b> (last checked 27 May 2022).
43. Lee WL, Bausell RB and Berman BM. The growth of health-related meta-analyses published from 1980 to 2000. *Eval Health Professions* 2001; 24: 327–335.
44. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D and Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* 1999; 354: 896–1900.
45. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; 339: b2700.
46. Tunis AS, McInnes MD, Hanna R and Esmail K. Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? *Radiology* 2013; 269: 413–426.
47. Panic N, Leoncini E, de Belvis G, Ricciardi W and Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One* 2013; 8: e83138.
48. The Cochrane Collaboration. In: Higgins JPT and Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0* (updated March 2011). See <https://training.cochrane.org/handbook/archive/v5.1/> (last checked 27 May 2022).
49. Centre for Reviews and Dissemination (2009). *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. CRD, University of York. See [www.york.ac.uk/inst/crd/index\\_guidance.htm](http://www.york.ac.uk/inst/crd/index_guidance.htm) (last checked 27 May 2022).
50. Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Int Med* 2015; 162: 777–784.
51. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. See [www.nice.org.uk/media/B52/A7/TAMethodsGuide\\_UpdatedJune2008.pdf](http://www.nice.org.uk/media/B52/A7/TAMethodsGuide_UpdatedJune2008.pdf) (last checked 6 September 2012, but no longer available).
52. NICE Decision Support Unit (DSU). Evidence Synthesis TSD series. See [www.nicesu.org.uk/Evidence-Synthesis-TSD-series\(2391675\).htm](http://www.nicesu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm) (last checked 27 May 2022).
53. Ades AE. ISPOR states its position on network meta-analysis. *Value in Health* 2011; 14: 414–416.
54. Chaimani A, Caldwell DM, Li T, Higgins JPT and Salanti G. Undertaking network meta-analyses (Chapter 11). In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. (eds) *Cochrane Handbook for Systematic Reviews of Interventions Version 6.0*. See [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook) (last checked 27 May 2022).
55. Ades AE, Caldwell DM, Reken S, Welton NJ, Sutton AJ and Dias S (2012). NICE DSU technical support

- document 7: evidence synthesis of treatment efficacy in decision making: a reviewer's checklist. See [www.ncbi.nlm.nih.gov/books/NBK395872/](http://www.ncbi.nlm.nih.gov/books/NBK395872/) (last checked 27 May 2022).
56. Ortega A, Fraga MD, Alegre-del-Rey EJ, Puigventós-Latorre F, Porta A, Ventayol P, et al. A checklist for critical appraisal of indirect comparisons. *Int J Clin Pract* 2014; 68: 1181–1189.
  57. Critical Appraisal Skills Programme. CASP Systematic Review Checklist. See [https://casp-uk.net/wp-content/uploads/2018/01/CASP-Systematic-Review-Checklist\\_2018.pdf](https://casp-uk.net/wp-content/uploads/2018/01/CASP-Systematic-Review-Checklist_2018.pdf) (last checked 27 May 2022).
  58. Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al.; GRADE Working Group. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014; 349: g5630.
  59. Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al.; GRADE Working Group. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *J Clin Epidemiol* 2018; 93: 36–44.
  60. Salanti G, Del Giovane C, Chaimani A, Caldwell DM and Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014; 9: e99682.
  61. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, et al. CINeMA: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Medicine* 2020; 17: e1003082.
  62. CINeMA Confidence in Network Meta-Analysis. See <https://cinema.ispm.unibe.ch/> (last checked 27 May 2022).
  63. Phillippo DM, Dias S, Welton NJ, Caldwell DM, Taske N and Ades AE. Threshold analysis as an alternative to GRADE for assessing confidence in guideline recommendations based on network meta-analyses. *Ann Int Med* 2019; 170: 538–546.
  64. Salanti G. Indirect and mixed treatment comparison, network or multiple treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synthesis Methods* 2012; 3: 80–97.
  65. Weibel S, Pace NL, Schaefer MS, Raj D, Schlesinger T, Meybohm P, et al. Drugs for preventing postoperative nausea and vomiting in adults after general anesthesia: an abridged Cochrane network meta analysis. *J Evidence Based Med* 2021; 14: 188–197.