Review Article

Methodological considerations in network meta-analysis

Shabbeer Hassan[1], Ravishankar N[1], N Sreekumaran Nair[2]

[1]Department of Statistics, Manipal University, Manipal, Karnataka, India.
[2]Public Health Evidence South Asia (PHESA), Department of Statistics, Manipal University, Manipal, Karnataka, India.
Correspondence to: N Sreekumaran Nair, E-mail: sree.nair@manipal.edu

Received January 21, 2015. Accepted February 1, 2015

Abstract

Network meta-analysis (NMA) is an extension of pairwise meta-analysis that facilitates comparisons of multiple interventions over a single analysis. It is the method in which multiple interventions (that is, three or more) are compared using both direct comparisons of interventions within randomized controlled trials and indirect comparisons across trials based on a common comparator. NMA is methodologically complex compared to simple pairwise meta-analysis as it accounts for a broader evidence base. Results from NMA are more useful to policy makers, service commissioners, and providers when making choices between multiple alternatives than those from multiple, separate pairwise meta-analyses. It can be an ideal choice to be extended to compare complex interventions that are multifaceted. Apart from the numerous benefits the NMA offers, it is prone to methodological complications that need to be understood, implemented, and finally reported correctly. This article is meant to provide a primer to the various methodological issues pertaining to NMA. The NMA can be as valid as a standard pairwise meta-analysis if these methodological issues are taken care of.

KEY WORDS: Network meta-analysis, indirect comparison, multiple treatment comparison, mixed treatment comparison

Introduction

Randomized controlled trials (RCTs) have been at the top of evidence chain[1] for a long time and has been the accepted methodology to be used a gold standard[2] to assess whether or not a health intervention works. They still exist as a key component of evidence-based medicine.[3–7] However, single RCTs are not enough to address the demands of policy makers for conclusive results. In order to address this problem, systematic reviews with meta-analysis (SRMA) are being increasingly used to synthesize results of different clinical trials that evaluate same interventions/treatments for obtaining an overall estimate of the treatment effect with respect to a control (active comparator/placebo).[8]

Access this article online	
Website: http://www.ijmsph.com	Quick Response Code:
DOI: 10.5455/ijmsph.2015.21012015131	

Network meta-analysis (NMA), multiple treatments meta-analysis (MTM), or mixed treatment comparison (MTC) provides a flexible and powerful generalization of pairwise meta-analysis traditionally used with SRMA[9] for comparison of multiple interventions simultaneously. As has been explained in numerous publications,[10,11] NMA allows for indirect comparisons of treatments that have not been studied in a head-to-head manner. For example, treatment effects from clinical trials comparing X with respect to Y (in XY trials) and trials comparing Z with respect to Y (in YZ trials) can be then pooled together in NMA to obtain an indirect estimate for indirect comparison between X and Z[12] by means of the common comparator Y. Even if direct evidences for X and Z (via XZ trials) exist, combining them with indirect estimates will strengthen the evidence base.[13] In the mid-1990s, the technique of "adjusted indirect comparisons" was used to compare multiple interventions. This technique was subjected to extension and periodical revisions, leading to the development of NMA. Higgins and Whitehead,[10] Bucher et al. (1998), Lumley (2002), Lu and Ades (2004) have made enormous contributions toward the development of NMA.

Network plot is the pictorial representation of all comparisons to be made within a NMA. A well-designed network plot will foster the analysis and will be helpful in procuring reliable estimates. A network plot can either be

an open loop or a closed loop. In a closed loop, direct and indirect estimates of the same comparison can be combined to give a mixed estimate, thereby accounting for a broader evidence base.[9]

NMA relies on two basic assumptions: consistency and transitivity. A perfect agreement between direct and indirect estimates of a comparison ensures consistency, and a balanced distribution of effect modifiers across all sets of trials grouped by a comparison guarantees the accomplishment of transitivity.

The fundamental difference between pairwise meta-analysis and NMA is that the former results in only one pooled estimate while the latter results in many pooled estimates. Because of the fact that the NMA includes multiple comparisons (both direct and indirect), it is prone to a high degree of heterogeneity. It is not simple to obtain such pooled estimates, as NMA is more complex than simple pairwise meta-analysis in SRMA.[14] Various methodological issues come up due to this generalization that needs to be *understood, implemented, and finally reported correctly*. Therefore in this paper, we would mainly discuss the various methodological issues surrounding the NMA.

Eligibility Criteria: Comparator Preference Bias

Specification of an appropriate eligibility criterion is one of the key aspects of any NMA, because the structure of network plot is solely dependent on it. Consequently, discrepancies in the eligibility criteria are bound to have an effect on the estimates of NMA. In most practical situations, the new drugs or treatments are compared with placebo (inactive comparator) instead of comparing against available standard treatment. Even though such comparisons are not of primary interest, they are compelled to be included in NMA to obtain indirect estimates.[9]

Network Geometry

An important aspect of any NMA is the structure of the treatment network. This geometry details in a graphical manner about all the treatment comparisons, both direct and indirect, which have been made and the strength of those comparisons.[15] For example, in the network given below, we see a network showing both direct and indirect comparisons.

Figure 1 shows network plot of 22 RCTs of the systematic review "incident diabetes in clinical trials of antihypertensive drugs."[16] The objective of the review was to assess the effects of five antihypertensive drugs, namely angiotensin receptor blockers (ARB), angiotensin converting-enzyme (ACE) inhibitors, calcium-channel blocker (CCB), beta-blocker (BB), and diuretics on incident diabetes. Circles (nodes) in the figure represent the individual treatments including placebo; the size of the nodes corresponds to the number of subjects receiving the treatment; all the lines represent direct comparisons; and the thickness of the lines represents the number of clinical trials.

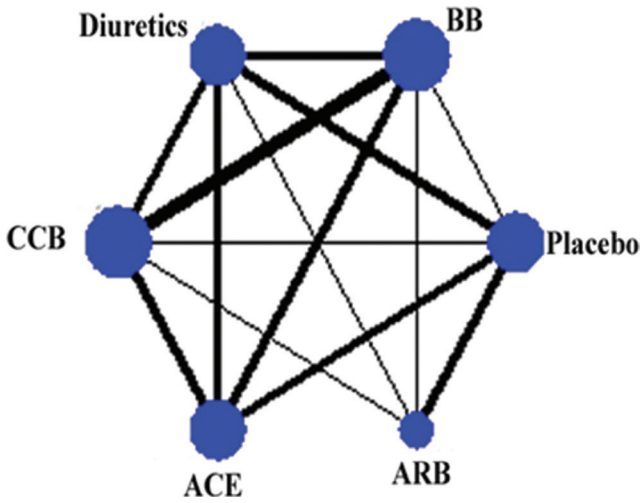

Figure 1: Network plot.

The figure has 15 pairwise comparisons and the most common comparison seems to be between CCB and BB. There are no trials comparing ACE and ARB, which constitutes an indirect comparison.

Hence, the strength and thus diversity of a network stem from the number of varied interventions it includes and the comparisons between those interventions that are available.[15] The number of evidences available for each comparison affects the influence of the study and also affects the reliability of the final pooled estimates so obtained.[17,18] Evidence coming from smaller trials would inherently contain many biases, both studied in this article and published in other studies. Therefore, if such smaller trials are included in the network, then the pooled estimates so obtained in the network would be unreliable.[19]

Network geometry can also give researchers an idea about certain peculiar tendencies in trials being conducted. An important concern is many trials test efficacy of interventions via comparison with a placebo and not other interventions. Indeed, this peculiarity can be easily seen in the treatment of partial epilepsy with second generation antiepileptic drugs[8,20] or biologic drugs for rheumatoid arthritis.[8,21] Hence, visualizing the direct comparisons available now (trials present) and the indirect comparisons (trials not present) would help researchers know which trials to conduct in future.

Inconsistency

Consistency means agreement between direct and one or more indirect evidences.[20] The assumption of consistency is linked to transitivity assumption. For example, in a highly simple triangular loop between X, Y, and Z treatments, consistency would hold if X and Y are transitive, then Z would be transitive as well. The Bucher et al. method[11] gives us a statistical method of evaluating inconsistency in a XYZ network as inconsistency factor (IF), which is equal to the

absolute value of the difference between the direct and indirect estimates. If consistency holds, then direct and indirect estimates can be combined to obtain a pooled estimate that is referred to as the mixed estimate. In certain cases, it would be of great interest to compare the mixed estimate and direct estimate as these quantities are correlated.[18]

Hence, consistency represents a very important assumption to be fulfilled before conducting an NMA as the generalizability of NMA depends on it. However, absence of consistency does always mean all the indirect estimates so obtained from the network are invalid.[21] When inconsistency is quite evident within the NMA, researchers can synthesize the data in such a manner which might reflect the extra uncertainty present in the NMA because of this inconsistency. Lu and Ades[22] in 2006 introduced a model that accounts for inconsistency as it adds an extra random effect to each link where inconsistency may occur. In the presence of such inconsistency, the direct and indirect treatment effects are allowed to differ by a small random quantity instead of them being assumed identical. This model is quite analogous to the random effect model. The variances of such inconsistencies are called inconsistency variances, which is quite similar to the heterogeneity variances. This inconsistency variance can be compared to the heterogeneity variance that can be used to assess the assumption that heterogeneity alone can explain all the differences between the evidence sources.

Sample Size and Power in NMA

We have discussed earlier about how pooled estimates can become unreliable due to the insufficiently powered individual trials used to obtain the network. Individual trials are the not the only important consideration but also the sample size required to obtain a precise estimate from NMA, and hence the statistical power is quite important. Only a few methodological papers have been published to address the important issue of power and effective sample size in NMA. A paper published by Thorlund and Mills[18] in 2012 explained three methods of quantifying the power and precision of NMA. The sample size required for a particular treatment comparison in an NMA can be understood as the number of patients in a traditional pairwise meta-analysis that would give the same degree and strength of evidence as that from NMA.

Evaluation of Risk of Bias and Quality of Evidence

Risk of bias assessment is quite difficult in NMA than in traditional pairwise meta-analysis. Such risk of bias occurs when the individual trials used for constructing the network have some inherent problems in their design and execution.[22] In NMA, more than one pooled estimate can be obtained and hence this is quite different from a traditional SRMA where only one such pooled estimate is obtained. So, if individual trials have been poorly designed and executed, then in conventional SRMA, only one pooled estimate is affected whereas in NMA, several pooled estimates can be affected in different degrees. For example, in direct estimates for XY trials the risk of bias may be low but in all the indirect estimates obtained from the XZ or YZ trials the risk can be quite high. Apart from this, various regions of a network can have different risk of bias that makes evaluation of them quite difficult in NMA.[23]

An important method to evaluate the quality of evidence of NMA had been developed by the Grading of Recommendation Assessment, Development and Evaluation (GRADE) Working Group.[24] This GRADE method allows the readers to judge the confidence with which a particular estimate of a treatment effect (indirect or direct) for any given outcome can be believed, using four grades: high, moderate, low, and very low.[24,25] The approach consists of five components: indirectness, inconsistency, imprecision, publication bias, and study limitations.[25]

A modification of GRADE approach was developed by Higgins and Green[23] in 2011 where the evaluation of NMA depends on indirect comparisons, contribution of the direct evidence to the NMA pooled estimates, assumption of transitivity, and the disagreement between the direct and indirect estimates. They summarize by mentioning that transitivity remains a key assumption in NMA and the assessment of this assumption in the indirectness component of pooled estimates is very important both within the GRADE approach and Higgins and Green[23] framework. However, if the NMA has some degree of inconsistency, then methods should be used to detect it; and if important inconsistency remains in the network, then NMA should be avoided by the researchers.

Rankograms

An important method helpful for policy making from NMA is the use of rankograms, that is, graphical methods used to rank treatments considered based on certain probabilities. In Bayesian framework, each treatment k is ranked according to the own estimated effect sizes. So, the proportion of Markov chain Monte Carlo (MCMC) cycles in which a given treatment would rank first among others gives the probability $p(k = 1)$ that treatment k ranks first, and is the best among the other available treatments.[26] In this way, probabilities are calculated for the second best and third best treatments, in such a way that these probabilities sum to one for each treatment and each ranking. This can be represented in a two-dimensional treatment specific plot[26] in which the horizontal axis represents the possible ranks of all the treatments in NMA and the vertical axis has the probability for that particular treatment to assume each of the possible ranks with respect to a given specific outcome.

A simple numerical summary that can supplement the rankogram plot is to estimate the surface under the cumulative ranking (SUCRA) lines for each of the considered treatment in NMA.[26–28] The larger the SUCRA value for a treatment k, the higher its rank among the available treatments. When a treatment is best, the SUCRA value is 1, and for worst treatment SCURA is 0.

Such SCURA plots can be obtained by using WinBUGS, R, or Stata software. However, as Salanti et al. concluded in their 2013 paper that the usefulness and the applicability of each such graph would depend on the nature of data used to obtain the NMA. For example, in star-shaped networks where all treatments are compared against a placebo, the assumption of consistency cannot be tested.[28] Another important point to note is that graphs alone may not be used for interpretation and only with proper numerical results can over-interpretation be avoided.

Heterogeneity

Heterogeneity remains an important criterion that must be assessed when constructing an NMA.[29] Statistical heterogeneity represents the situation: are the treatment effects seen within the individual trials comparing the same treatments similar or dissimilar? Such statistical heterogeneity as in SRMA can be checked using statistical tests such as Cochran's *Q* and quantified by I^2. An important reminder for researchers is that even though statistical heterogeneity can be insignificant, there may be conceptual heterogeneity that also needs to be looked for. Conceptual heterogeneity refers to any differences in study populations, study setting, follow-up procedures, outcome measures, or any other feature that might make clinical trials different.[8] Even though such heterogeneity exists in traditional SRMA, an important thing to keep in mind in NMA is that there are a lot of multiple comparisons. Hence, conceptual heterogeneity must be checked for both in each comparison and between all comparisons done in NMA.

Effect Modifiers

Effect modifiers represent various study and patient characteristics that are also associated with the patient characteristics. As NMA includes various single trials involving different interventions conducted on different populations, with different underlying characteristics, the distribution of effect modifiers would vary not only across such studies involving same treatment comparison but also between different treatment comparisons. If we do find an imbalance in the distribution of effect modifiers, then the resultant pooled estimates from NMA would be biased.[30–32] If there are no imbalances in the distribution of effect modifiers, then the NMA is as valid as a standard pairwise meta-analysis.

Deciding which covariates would be effect modifiers based on the differences in trial results requires a careful analysis by first generating a potential list of potential effect modifiers for the treatments in question based on prior knowledge from literature.[33] Secondly, potential imbalances between direct comparisons should be checked based on the distribution of the trial characteristics that have been deemed as effect modifiers.

If there are a sufficient number of studies within an NMA, then it is possible to conduct a meta-regression analysis in which the treatment effect obtained from each study included in NMA is not only a function of the treatment comparison for that particular study but also is related to the effect modifier identified previously.[34] NMA is usually conducted on data extracted from study-level data, but the effect modifiers discussed depends on patient level, which in turn could lead to an ecological bias.[32,35–38] Hence, patient-level data should be used as much as possible as it can improve parameter estimation of NMA models.[37,38]

Statistical Models for NMA

Many statistical models have been proposed for comparing mixed treatments in NMA.[39,40] Fixed effect model assumes a common effect to be present behind the observed effects, whereas random effect model assumes that the true effect follows a distribution. When a substantial amount of heterogeneity whether statistical (as measured by I^2) or conceptual (while taking into account the wide differences between studies) is present, then a random effect model is better suited.[41] The rationale behind random effect model is there could be many true values for the treatment effect as there might be a whole range of populations consisting of different patients. So each trial would then estimate its own true value for the considered treatment effect. All such true values from the different studies considered are not exactly the same, but they would all be closely related.

Hence, it is observed that there are two sources of variation: variation within studies that can be attributed to between patients (as in fixed effects) and variation between studies due to some heterogeneity (as there can be many possible true values for the treatment effect depending on dosages provided, duration such dosages are given, etc.), so there can be a whole range of populations of different patients.

Here is introduced the random effect model:

$$f\left(x_{ijk}\right) = \alpha + \beta_i + \tau_j + \left(\beta\tau\right)_{ij} \quad (1)$$

where x_{ijk} is the outcome for subject *k*, which has been subjected to treatment group *j* in the *i*th study. Function *f* in Equation (1) defines the relation between the outcome *x* and the various effects *i*, *j* and *k*.

In meta-analysis and even in NMA, individual patient data (IDP) is seldom available and rates or odds are used as summary estimates. The equation can be rewritten as:

$$\mu_{ij} = \alpha + \beta_i + \tau_j + \left(\beta\tau\right)_{ij} \quad (2)$$

where μ_{ij} is the average treatment effect for the *j*th treatment in the *i*th study. The interpretation of Equation (2) is slightly different from the first one. As here α denotes the average treatment effect for a specific treatment *j*, its treatment effect depends on the study characteristics, various treatment applied (treatments have different effects), and interaction between study in question and treatment in consideration (with a random effect model, the effect of treatment *j* is considered to vary across the different studies).

The above model can then be implemented via any statistical software package such as SAS, Stata, or R.[42,43] Statistical models for NMA tend to become more complex

when there are more than two treatment arms in some of the included studies. This is due to the fact that observed differences in the treatment effects within a study with multiple (>2) arms are not independent of each other. Bayesian approach to NMA has of late become the most popular approach for analyzing NMA models[42] and we look into this in the next section.

Bayesian Networks

Bayesian NMA as said earlier is increasingly being used for evidence consolidation and this has been due to the ease of access to computational tools such as WinBUGS, Stata, SAS, or R. Bayesian networks use the Bayesian statistical framework for synthesizing the direct and indirect treatment comparisons to obtain mixed effects.[43,44] Within this framework, prior beliefs regarding the parameters are assumed initially and are factored into analysis later on. However, in practice non-informative priors are used, which means that the mean treatment effect is taken as zero with a very large standard error component. Posterior distributions are then obtained after taking into account the prior distributions and the obtained data.[44] The main reason for taking non-informative priors is that prior to the conduct of the study, evidence for the treatment effect is scant and not reliable; hence, a large standard error is used to give credence to the wide variability in the main treatment effect. If for example, one wants to study the treatment differences of two interventions, then a flat prior distribution like uniform distribution or normal distribution with a large variance can be used. Posterior distribution can then be later obtained from data.

However, obtaining such posterior distributions is quite a complex task but is made easier with the advent of many statistical packages. Conventionally simulation methods like Markov chain Monte Carlo (MCMC) are used to derive such posterior distributions.[45] As it is a simulation method, which involves taking some initial values and then iterates over and over to get an approximate posterior distribution, it has several advantages. For example, it can estimate predicted values for treatment effects that then can be used to derive rankograms. Another advantage is that the Bayesian softwares used are more adaptable to various situations like multiple treatment arms, parallel group, or split-mouth design.[46] A major disadvantage of such Bayesian methods is that they present the practitioners with a steep learning curve as the techniques used are mathematically quite complex.

Discussion

As we have shown here that a lot of methodological issues exist that should be carefully considered before conducting an NMA. However, NMA has a definite promise to displace SRMA as the gold standard of evidence-based medicine. An encouraging case of NMA being proved correct comes from the work of PROTECT (Patient-Related Outcomes with Endeavor versus Cypher Stenting Trial)[47,48] published in Lancet in 2012. PROTECT was a study that compared sirolimus-eluting stents versus first generation zotarolimus-eluting stents. Before its publication Palmerini et al.[48] published an NMA in March 2012 providing several quantitative estimates and predictive OR that compared the several drug-eluting stents in the network involving 49 trials, 50,844 patients. These results came out months before PROTECT trial, which was published in August 2012.

A comparison of results was performed by Biondi-Zoccai et al. in their article published in 2013 in which they compared both the effect size and precision of OR estimates for stent thrombosis. Similar results were obtained from both studies[48] and hence the PROTECT trial results were validated prior to their publication by Palmerini et al.[48] This brings to a conclusion that NMA results can be very important and valid when properly conducted on a sizeable number of studies.

Caruba et al.[49] used NMA to compare the efficacy and mean cost per patient after 1 and 3 years of follow-up of five treatment modalities: medical therapy (MT), percutaneous coronary intervention (PCI) without stent (PTCA), percutaneous coronary intervention with bare-metal stent (BMS), percutaneous coronary intervention with drug-eluting stent (DES), and elective coronary artery bypass graft (CABG) of stable angina.[49] They observed that "NMA was able to document considerable differences in treatment costs at 3-year follow-up, when comparing five treatment modalities that provided similar clinical results, in terms of death and risk of myocardial infarction."

Cooper et al. conducted a study on "evaluating the effectiveness of interventions to increase the uptake of smoke alarms."[50] This was the first application of NMA in the field of public health. The authors used NMA random effect model and also ranked intervention effectiveness based on absolute intervention effects and calculated the probability that each intervention is best for a particular outcome. They elucidate that "results from NMAs are more useful to policy makers, service commissioners, and providers when making choices between multiple alternatives than those from multiple, separate pairwise meta-analyses." They add on saying that NMA is the ideal choice for comparing and testing the relative efficiency of multiple injury prevention interventions that are often complex and multifaceted with a meager number of studies evaluating same interventions. Furthermore, there are attempts to bring about a methodology to incorporate individual participant data into NMA which, would greatly increase the influence of the study and will be helpful to explore subject-level covariates.

NMA has experienced a tremendous transition over the last decade with several international organizations conducting extensive research and numerous scientific conferences and workshops held across the world in several occasions. Methodological innovations, better reporting standards, and availability of state-of-the-art software packages have contributed to its wide-spread acceptance.

Conclusion

The NMA has a tremendous potential to be the gold standard method for health care evidence synthesis. It is a springboard that has the ability to take the evidence-based health care to a higher level. However, the methodology of NMA is prone to several complications that have to be seriously considered. The NMA can be as valid as a standard pairwise meta-analysis if these methodological issues are taken care of.

Acknowledgment

We cordially acknowledge Public Health Evidence South Asia (PHESA) and Dr. TMA Pai, Endowment Chair in Systematic Review and Evidence-Based Public Health, Manipal University, Manipal for supporting this work.

References

1. Akobeng AK .Principles of evidence-based medicine. Arch Dis Child 2005;90:840–44.
2. National Institutes of Health. Interactive Autism Network Gold Standard evidence: the randomized controlled trial. 2005 [cited 2015 Dec 25]. Available at http://www.iancommunity.org.
3. Prasad K. Evidence-based medicine in India. J Clin Epidemiol 2013;66(1):6–9.
4. Lockwood S. “Evidence of me” in evidence-based medicine. BMJ 2004;329:1033.
5. Sackett DL, Rosenberg WMC, Gray JAM, Harnes RB, Richardson WS. Evidence-based medicine: what it is and what it isn’t. BMJ 1996;312:71–2.
6. Rosenberg W, Donald A. Evidence-based medicine: an approach to clinical problem-solving. BMJ 1995;312:1122–6.
7. Kerridge I, Lowe M, Henry D. Ethics and evidence-based medicine. BMJ 1998;316:1151.
8. Mills EJ, Thorlund K, Ioannidis JPA. Demystifying trial networks and network meta-analysis. BMJ 2013;346:f2914.
9. Caldwell DM, Ades AE, Higgins JPT. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005;331:897–900.
10. Higgins J, Whitehead A. Borrowing strength from external trials in a meta-analysis. Stat. Med 1996;15:2733–49.
11. Bucher H, Guyatt G, Griffith L, Walter S. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997;50:683.
12. Sutton A, Ades AE, Cooper N, Abrams K. Use of indirect and mixed treatment comparisons for technology assessment. Pharmaco economics 2008;26:753–67.
13. Madan J, Stevenson MD, Cooper KL, Ades AE, Whyte S, Akehurst R. Consistency between direct and indirect trial evidence: is direct evidence always more reliable? Value Health 2011;14:953–60.
14. Tu YK, Faggion CM. A primer on network meta-analysis for dental research. ISRN Dentistry 2012;2012:276520.
15. Salanti G, Kavvoura FK, Ioannidis JPA. Exploring the geometry of treatment networks. AnnIntern Med 2008;148(7):544–53.
16. Elliott WJ, Meyer PM. Incident diabetes in clinical trials of antihypertensive drugs: a network meta-analysis. Lancet 2007;369:201–07.
17. Mills EJ, Ghement I, O’Regan C, Thorlund K. Estimating the power of indirect comparisons: a simulation study. PLoS One 2011;6:e16237.
18. Thorlund K, Mills EJ. Sample size and power considerations in network meta-analysis. Syst Rev 2012;1:41.
19. Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. JAMA 1998.279:1089–93.
20. 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 Syn Meth 2012;3:80–97.
21. 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. Ann Rheum Dis 2013;72:1524–35.
22. Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. JASA 2006;101(474):447–59.
23. Higgins JPT, Green S, (Eds.). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0. The Cochrane Collaboration. 2011 [updated March 2011]. Available at http://www.cochrane-handbook.org.
24. Li T, Puhan MA, Vedula SS, Singh S, Dickersin K. Network meta-analysis-highly attractive but more methodological research is needed. BMC Med 2011 June 27;9:79.
25. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 8. Rating the quality of evidence–indirectness. J Clin Epidemiol 2011;64(12):1303–10.
26. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JPT. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014;9(7):e99682.
27. Salanti G, Ades A, Ioannidis J. Graphical methods and numeric summaries for presenting results from multiple treatments meta-analysis: an overview and tutorial. J Clin Epidemiol 2011;64:163–71.
28. Comparing Multiple Interventions Methods Group (CMIMG) [internet]. 2013 [updated 2013 Sept 6, cited 2015 Jan 5]. Available at http://cmimg.cochrane.org/about-us.
29. Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS One 2013;8(10):1–12.
30. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity–subgroups, meta-regression, bias, and bias-adjustment. Med Decis Making 2013;33(5):618–40.
31. Pereira TV, Patsopoulos NA, Salanti G, Ioannidis JP. Critical interpretation of Cochran’s Q test depends on power and prior assumptions about heterogeneity. Res Syn Meth 2010;1:172.
32. Hoaglin D, Hawkins N, Jansen J, 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 Health 2011;14:429–37.
33. Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. BMC Med 2013;11:159.
34. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002;21:1559–73.
35. Cooper NJ, Sutton AJ, Morris D, Ades AE, Welton NJ. Addressing between-study heterogeneity and inconsistency in

mixed treatment comparisons: application to stroke prevention treatments in individuals with non-rheumatic atrial fibrillation. Stat Med 2009;28:1861–81.
36. Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. J Clin Epidemiol 2009;62:857–64.
37. Jansen JP. Network meta-analysis of individual and aggregate level data. Res Syn Meth 2012;3:177–90.
38. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. Stat Med 2012;31:3516–36.
39. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med 2002;21(16):2313–24.
40. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. Health Technol Asses 2005;9(26):130–34.
41. Whitehead A. *Meta-Analysis of Controlled Clinical Trials*. Chichester, UK: Wiley, 2002.
42. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004;23(20): 3105–24.
43. 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.
44. Ades AE, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. Pharmaco Economics 2006; 24(1):1–19.
45. Berg BA. Markov chain Monte Carlo simulations and their statistical analysis. Master Rev 2006;9:1–54.
46. Jansen JP,Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health 2008;11(5): 956–64.
47. Camenzind E, Wijns W, Mauri L, Kurowski V, Parikh K, Gao R, et al. Stent thrombosis and major clinical events at 3 years after zotarolimus-eluting or sirolimus-eluting coronary stent implantation: a randomised, multicentre, open-label, controlled trial. Lancet 2012;380:1396–405.
48. Palmerini T, Biondi-Zoccai G, Della RD, Stettler C, Sangiorgi D, D'Ascenzo F, et al. Stent thrombosis with drug-eluting and bare-metal stents: evidence from a comprehensive network meta-analysis. Lancet 2012;379:1393–402.
49. Caruba T, Katsahian S, Schramm C, Charles Nelson A, Durieux P, Bégué D, et al. Treatment for stable coronary artery disease: a network meta-analysis of cost-effectiveness studies. PloS One 2014;9(6):e98371.
50. Cooper NJ, Kendrick D, Achana F, Dhiman P, Saramago P, He Z, et al. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms. Epidemiol Rev 2012;34:32–45.

How to cite this article: Hassan S, Ravishankar N, Nair NS. Methodological considerations in network meta-analysis. Int J Med Sci Public Health 2015;4:588-594

Source of Support: Dr. TMA Pai, Endowment Chair in Systematic Review and Evidence-Based Public Health, Manipal University, **Conflict of Interest:** None declared.