

Is Network Meta-analysis a Revolutionary Statistical Tool for Improving the Reliability of Clinical Trial Results? A Brief Overview and Emerging Issues Arising

GEORGE BEIS¹ and IOANNIS PAPASOTIRIOU²

¹Research Genetic Cancer Centre S.A., Florina, Greece;

²Research Genetic Cancer Centre International GmbH Headquarters, Zug, Switzerland

Abstract. *Network meta-analysis (NMA) as the quantification of pairwise meta-analysis in a network format has been of particular interest to medical researchers in recent years. As a powerful tool with which direct and indirect evidence from multiple interventions can be synthesized simultaneously in the study and design of clinical trials, NMA enables inferences to be drawn about the relative effect of drugs that have never been compared. In this way, NMA provides information on the hierarchy of competing interventions for a given disease concerning clinical effectiveness, thus giving clinicians a comprehensive picture for decision-making and potential avoidance of additional costs. However, estimates of treatment effects derived from the results of network meta-analyses should be interpreted with due consideration of their uncertainty, because simple scores or treatment probabilities may be misleading. This is particularly true where, given the complexity of the evidence, there is a serious risk of misinterpretation of information from aggregated data sets. For these reasons, NMA should be performed and interpreted by both expert clinicians and experienced statisticians, while a*

more comprehensive search of the literature and a more careful evaluation of the body of evidence can maximize the transparency of the NMA and potentially avoid errors in its interpretation. This review provides the key concepts as well as the challenges we face when studying a network meta-analysis of clinical trials.

Meta-analysis is a highly recognized scientific discipline capable of providing high-quality evidence in medical research, particularly in clinical oncology (1-3). The main goal of a meta-analysis is to provide a definitive answer to the fundamental clinical research question about which treatment is most effective when it has been evaluated by multiple studies but with inconsistent results (4). In particular, the purpose of meta-analysis steps in clinical trials, as in any intervention study in general, is to calculate the true effect size of a specific treatment, that is, the same type of intervention compared with similar control groups. Therefore, it is possible to assess whether a particular type of treatment is effective. For the strategy followed in randomized controlled trials (RCTs), pairwise meta-analysis is a well-known statistical tool for synthesizing evidence from multiple trials but refers only to the relative effectiveness of two specific interventions. Therefore, the utility of pairwise meta-analyses is very limited in medical reality. Because there are usually many competing interventions for a given disease, studies related to some of the pairwise comparisons may be missing. Only a small percentage of these have been examined in head-to-head studies. For these reasons, needs have led to the development of network meta-analysis (5-7) (NMA), which is also called mixed treatment comparisons (8-11) (MTC) and may provide more accurate estimates of treatment effects than a pairwise meta-analysis (12). Especially when comparisons between important treatments are missing (13), NMA may be a more useful technique, as it is used to compare multiple treatments

Correspondence to: George Beis, Research Genetic Cancer Centre, 53100 Florina, Greece. Tel: +30 2385041950, e-mail: beis.giorgos@rgcc-genlab.com and Ioannis Papisotiriou, Research Genetic Cancer Centre International GmbH, Baarerstrasse 95, Zug 6300, Switzerland. Tel: +41 (0)417250560, email: office@rgcc-international.com

Key Words: Network meta-analysis, Bayesian network models, Markov Chain Monte Carlo algorithms, indirect evidence, treatment effects, treatment network, review.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC-ND) 4.0 international license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

simultaneously in a statistical study, whereby combining direct and indirect evidence in a network of randomized controlled trials (RCTs) (13-16), by providing a more complete picture to clinicians and thus enabling them to more clearly ‘rank’ treatments using summary results. This is achieved by assessing a composite (mixed) effect size as the weighted average of these direct and indirect components, which then allows competing interventions to be ordered more clearly according to their relative effectiveness, even if they have not been compared in a single trial (17, 18).

In recent years this statistical approach has matured as a technique (19, 20), where models are available for raw data that produce different aggregated outcome measures, using both frequentist and Bayesian models through statistical software packages (16). Especially in the last decade, many applications have been published (21, 22), as there are methodological developments in the subject of NMA. The study of the concept of NMA came to the fore to ‘open wider horizons’ for clinicians, by drawing information from the evaluation of a connected network of studies comparing the results of several interventions simultaneously (23). This approach has gained great popularity among clinicians and decision-makers because the costs involved in the development of new or unnecessary clinical studies may be reduced.

The study of an NMA model during the approval process of a drug can make a decisive contribution to the design of a clinical trial by giving accurate information about both the competitive picture and the corresponding evidence so that the information collected can help to ensure that the clinical trial design is the best possible to receive strong support. Consequently, the NMA is a very useful tool for evaluating the comparative effectiveness of different treatments commonly used in clinical practice, provided, however, that appropriate care is taken in the interpretation of the concepts that characterize it so that the results are not biased or bulging (24). Although this technique is increasingly used by biomedical researchers, it has created several challenges and pitfalls that deserve careful consideration, especially since this technique cultivates all the hypotheses of pairwise meta-analysis but with greater complexity due to the multitude of comparisons involved. Moreover, despite the wider acceptance of NMA, there are concerns about the validity of its findings (25). However, as NMA remains a hot research topic to this day, the purpose of this review is to examine the key concepts underlying it, focusing on its risks and benefits, and outlining relevant emerging issues and concerns.

Network Geometry

In clinical trials it is known that for n treatments in NMA the maximum number of designs (*i.e.* each combination of treatments within a study) is $2^n - 1$, while for each multi-

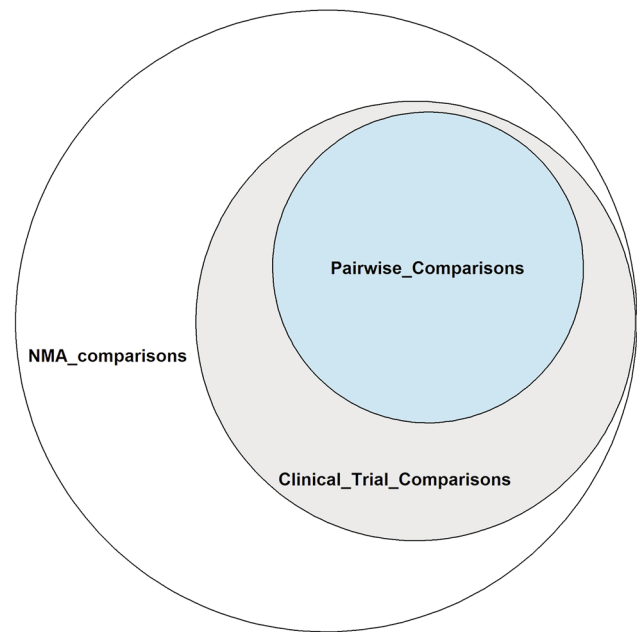


Figure 1. Venn Diagram resulting from overlapping comparisons of network meta-analyses (from the binomial formula), pairwise comparisons (from network diagrams), and trial comparisons (from study-based registry).

arm study, there are $(n!2) = n(n-1)/2$ comparisons including all possible unique comparisons, even if they are not observed in clinical trials or a pairwise meta-analysis (Figure 1), which would lead to a fully connected network. However, some of the comparisons predicted by the combinatorial formula will be ineligible due to protocol compliance or post hoc limitations (26).

The most important parameter in the utility of a treatment network before relevant data analysis is the assessment of its geometry (27-29), showing which interventions have been directly compared in RCTs and which can only be indirectly compared. In particular, the geometry of the network allows one to understand how many choices there are for each treatment, whether or not certain types of comparisons have been avoided, and whether there are particular patterns among the possible choices of the comparators. However, a network can ‘mutate’ over time as more tests are carried out, thus modifying its geometry which must be studied at each evolving step.

NMA Assumptions

NMA requires the same steps as a conventional meta-analysis but is graphically represented with a network, thus providing direct information about treatments that can be compared with each other and identifying all interventions linked to a

common comparator (the linking treatment). For example, two different treatments have been compared with a placebo in different trials. An NMA allows a hypothesis test to be created that compares these active treatments to each other based on their effectiveness against a common comparator (usually a placebo), thus providing ‘indirect’ evidence. These indirect comparisons provide the opportunity to fill the knowledge gaps of efficacy comparisons of existing treatments, thereby providing a more comprehensive understanding of the multitude of treatment options for the clinician. In short, the network estimate is an aggregate result of the direct and indirect evidence for a given comparison or the indirect evidence if no direct evidence is available. Then, once all the treatments in the existing network have been compared, there are different methods for ranking (30-33) the treatments in terms of their net effectiveness.

The main objective of NMA is to examine and statistically validate the effects of each treatment by evaluating and analyzing three or more interventions/treatments using both direct and indirect evidence. Therefore, basic assumptions such as transitivity, consistency, and homogeneity of direct evidence should be satisfied for performing NMA to be valid. More specifically, these assumptions should be evaluated with statistical tests (34). However, these methodological aspects, although poorly understood, are nevertheless key concepts for understanding a network meta-analysis (35, 36). For this reason, we will explain the basic principles governing these assumptions.

The Concepts of Transitivity, Consistency, and Heterogeneity in NMA

Transitivity (37) is the composition of studies that makes a direct comparison between 2 meta-analytic estimates A vs. C and B vs. C meaningful when the studies are similar in important clinical characteristics that influence the relevant treatment effects (9) (effect modifiers, *i.e.*, characteristics that influence the relevant outcomes of a clinical intervention) which need not be identical and therefore can be examined by comparing the distribution of potential effect modifiers across the different comparisons (38). Indirect information on the relative effect of 2 interventions will be considered valid if the studies and comparisons in the network do not differ in terms of the distribution of the various effect modifiers (the intervention effects are transitive). A valid indirect comparison (such as AB) requires both AC and BC studies to be similar in terms of the distribution of these characteristics, and only then will the assumption of transitivity apply. Then the indirect comparison (AB) is calculated by subtracting BC from AC as defined by the formula (20, 39):

$$\theta_{\text{indirect}}^{\text{AB}} = \theta_{\text{direct}}^{\text{AC}} - \theta_{\text{direct}}^{\text{BC}}$$

where θ denotes the observed estimates of treatment effects in terms such as odds ratios (OR), mean difference (MD), *etc.* In oncology, time-to-event data (40) are used where the hazard ratio (HR) (41) is taken as the necessary measure to interpret treatment efficacy. The HR is calculated using Cox regression models (42) in the survival analysis and indicates the relative probability of the event (*e.g.*, death). Transitivity, although is an essentially incalculable hypothesis, nevertheless, its validity can be assessed by clinical and epidemiological methods (34), and suitable models have been found through which, with suitable modifications, its valid hypothesis can be ensured (43). However, if the clinical characteristics are different (*e.g.*, different patient populations), then the transitivity assumption is violated, so the estimate of the indirect AB comparison is invalid (44, 45). Furthermore, detecting the absence of transitivity can often be difficult because sufficient details published in clinical trials are not always available to allow a detailed assessment (46).

The transitivity translated into statistical terms (36) is essentially the consistency (or coherence) and occurs when the above abstraction equation is supported by the corresponding data, but it can only be evaluated when there is a loop in the evidence network, that is when there is direct and indirect evidence for a specific comparison of interventions. The basic assumption underlying the validity of indirect and mixed comparisons is that there are no significant differences between trials making different comparisons in addition to the treatments already compared. So, an area that remains open and is one of the biggest challenges in NMA is inconsistency (36, 44, 47) which generally occurs when direct and indirect evidence diverge (37) ($\theta_{\text{indirect}}^{\text{AB}} \neq \theta_{\text{direct}}^{\text{AB}}$).

More specifically, the inconsistency may arise from the characteristics of the studies due to their different design or when the estimates of the size of the direct and indirect effects differ (48).

The magnitude of inconsistency in an NMA can be statistically calculated by comparing direct and indirect summative effects in predetermined loops (15, 49) or a network by fitting models that allow and disallow inconsistency (50, 51). There are several methods for measuring inconsistency when suspected (48), such as the Akaike (52) and Deviance (53) information criteria for assessing the goodness of fit of models in frequentist/Bayesian approaches to NMA or meta-regression models (50). Also, several methods for detecting inconsistency in an RCT network include the inconsistency parameter approach (48) and the net heat graphical approach (54, 55). ‘Node splitting’ model methods (56-58) have been reported too in the literature to assess inconsistency in NMA, with any direct comparison excluded from the network and then calculating the difference between these direct and indirect components from the network, while appropriate decision rules have been defined to select only those comparisons belonging to potentially inconsistent loops in the network (57). As

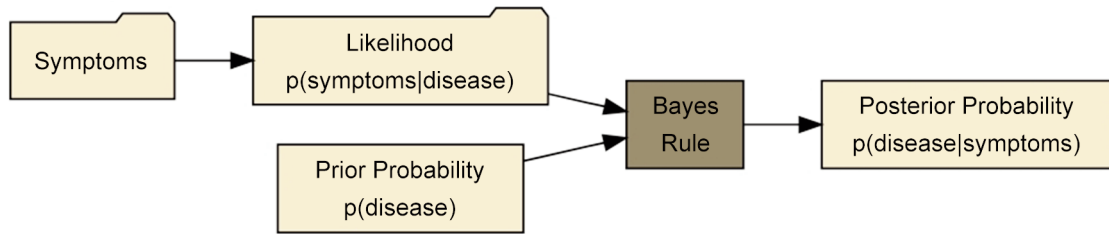


Figure 2. Illustration of Bayes' theorem applied to medical research where data, in the form of symptoms, is used to determine the likelihood of those symptoms if the patient has a particular disease. Bayes' rule combines this probability with prior knowledge to determine the posterior probability that the patient has the disease, given the observed symptoms.

mentioned earlier, inconsistency exists when discrepancies between direct and indirect estimates exist, therefore transitivity is a common cause of inconsistency.

Another very important advantage of NMA is its ability to investigate whether there is homogeneity or heterogeneity between the results from different clinical trials in each of the pairwise comparisons it involves, and therefore, assessment of heterogeneity in the results of different trials in each of the pairwise comparisons is important and should be considered. There are many valuable reviews on assessing and dealing with heterogeneity (59, 60) in a network. Heterogeneity in a meta-analysis is usually assessed with Cochran's Q statistic (61-64) and in particular with Cochran's generalized Q statistic for multivariate meta-analyses, where it can be used in the context of NMA to quantify heterogeneity across the network, both within trials and between trials (the latter is known as inconsistency). Although heterogeneity variance is often the most difficult parameter to estimate, several alternative approaches to estimating this variance have been explored in NMA studies (65) in recent years such as the use of the I^2 statistic (62, 66-68) (proportion description between-study variation) or meta-regression models (69, 70) are mainly used to reduce heterogeneity (and inconsistency) between RCTs in the network. Measures have also been considered to assess its confidence in the results of an NMA, where the impact of its variability on the corresponding clinical decisions is analyzed (71). In the special case when variances in between-study heterogeneity are estimated with considerable imprecision (because the data are sparse), including external evidence usually improves the conclusions (72). However, as the power and precision of indirect comparisons included in the NMA study depend on sample size and extensive statistical information, further improvements in methodology should be made.

Ranking Treatments in NMA

The results of the studies are closely associated with uncertainty, and consequently, we cannot be sure that the treatment is the most effective. But we can determine the

probability of a particular outcome about which treatment is best. With Bayesian thinking for each treatment, the probability of having a particular rank is derived from the posterior distributions of all possible treatments. The treatments are then ranked by the area under the cumulative rank curve SUCRA (30), which is a quantification of the overall rank and presents a unique number associated with each treatment. The higher the SUCRA value and the closer to 100%, the higher the probability that a treatment is close to the first place, while the opposite is true when this value is close to 0. To compare treatments in an NMA, a frequentist analog of SUCRA -by considering the frequentist perspective- called *P*-score is also used. Both concepts allow the ranking of treatments on a continuous scale of 0-1 (73), while rankograms represent these values graphically (74).

Bayesian Statistical Inference

In addition to frequentist inference which is arguably more commonly used in most research fields, Bayesian statistics (75, 76) is another very important statistical inference tool, having as its main advantage a framework that properly accounts for uncertainty in variance heterogeneity (77), and at the same time is more flexible as it can handle problems that frequency techniques cannot, such as handling missing data. In addition, it is considered more robust because it gives more precise effect estimates with smaller credible intervals, thus implying that it should not be considered as a competitive method of frequentist statistical analysis, but as an additional tool that contributes to the success of a more significant result.

Bayesian statistics treat the unknown quantities as random variables and assign a prior probability distribution to each of them, whereby specifying a joint probability distribution for the data (*i.e.*, a likelihood) we get a full probability model for the set of observable and unobservable quantities. In a few words, in Bayesian inference, prior beliefs (represented by prior distributions) are combined with existing data to arrive at a posterior distribution (Figure 2). So let us assume that the observed data are represented by y and the unknown parameters by θ . Then to have relevant inference for we use

Bayes' theorem (78, 79) to get a posterior distribution for making predictions about future events, *i.e.*, the joint distribution of all the parametric models that depend on the observed data: $p(\theta|y) \propto p(\theta)p(y|\theta)$. Here $p(y|\theta)$ is the conditional probability of the data given the model parameters which is known as the likelihood function, while the term $p(\theta)$ is the probability of certain model parameter values in the population which is the prior distribution. Therefore, the posterior distribution $p(\theta|y)$ is proportional to the likelihood function times the prior distribution (80, 81).

Methodology

Bayesian meta-analysis is mainly based on the hierarchical Bayes model, with the basic principles of this model being very similar to the ordinary random-effects model. When fitting Bayesian meta-analysis models, it is critical to test the model for whether it included sufficient iterations for convergence, as well as to perform sensitivity analyses with different prior standards to assess the effect on the overall simulation results. The Markov Chain Monte Carlo (MCMC) algorithm (82) that is used in Bayesian probabilistic models must have found the optimal solution (due to convergence); otherwise, more iterations will have to occur. MCMC simulation plays a very important role here because it allows

the estimation of the posterior distributions of the parameters for the results of the NMA (83).

Software Options for Fitting NMA Models and Assessing Inconsistency

The most popular software R (84) packages accessible and currently available for Bayesian and frequentist inference in NMA are included in Table I. Details on how data is analyzed, its input options, and the corresponding statistical models can be found in each package's respective manuals, which are also mentioned in the references. However, because most of these packages require strong contact with statistical programming for their use (existence of routines for performing NMA), there are also toolkits based on simple and standard instructions, intended to present the results using only the graphs of the analyses (85).

An Example of a Network Meta-analysis in Diabetes

The objective of the NMA that was applied in Diabetes disease was considered as an example for the estimation of the relative effects on *HbA1c* (glycated hemoglobin) change to a baseline sulfonylurea therapy in patients with type 2 diabetes, where the mean *HbA1c* change from baseline was used in the study and

Table I. Summary of the most important R software packages for reporting and interpretation of results regarding the Bayesian and Frequentist inferences in network meta-analyses.

RJags	R package provides an interface from R to the Just Another Gibbs Sampler (JAGS) (57, 86, 87) library for Bayesian data analysis (88, 89).
BUGSnet	R package (Bayesian inference using Gibbs sampling to conduct network meta-analysis) (90) where Bayesian analyses are conducted with JAGS using generalized linear models and accounting for heterogeneity and inconsistency, combined with the surface under the cumulative ranking curve (SUCRA).
gemtc	R package which incorporates evidence of multiple treatments by fitting generalized linear models under Bayesian statistics assumptions using JAGS. Also, it includes methods to assess heterogeneity and inconsistency, and several standard visualizations (8, 57, 83, 87, 91). In this way, NMA models are built using MCMC simulations based on our choice and using in parallel: JAGS called by the rjags R package, OpenBUGS (92) using the BRugs R package (93) or the WinBUGS (80) software called by the R2WinBUGS R package (94).
multinma	R package (95) that includes functions for setting up NMA and multilevel meta-regression (ML-NMR) models (96, 97). More specifically a suite of tools is used for performing ML-NMR and NMA with individual patient data (IPD) (98, 99), aggregate data (AD), or mixtures of both (100, 101). A range of results is thus obtained by performing model fitting estimated in a Bayesian framework using the Stan sampler (102, 103).
Stan	R package (103), a probabilistic programming language for specifying statistical models, defines a log-likelihood function on parameters that depend on specified data and constants that provide Bayesian inference for continuous variable models. These models operate through MCMC methods, such as the No-U-Turn sampler (104), which is an adaptive form of Hamiltonian Monte Carlo sampling. Maximum likelihood (penalized) estimates are calculated using optimization methods, such as the Broyden-Fletcher-Goldfarb-Shanno limited memory algorithm (105).
MCMCpack	R package (106) as a computational tool that uses MCMC methods to perform Bayesian inference so that we can fit innovative models of our choosing.
netmeta	R package (107-109) which provides frequentist methods for network meta-analysis.
pcnetmeta	R package which provides functions for arm-based NMA (110, 111).
Runjags	R package for MCMC models <i>via</i> JAGS evaluating their performance with drop-k validation and providing additional distributions including the Pareto family (112).

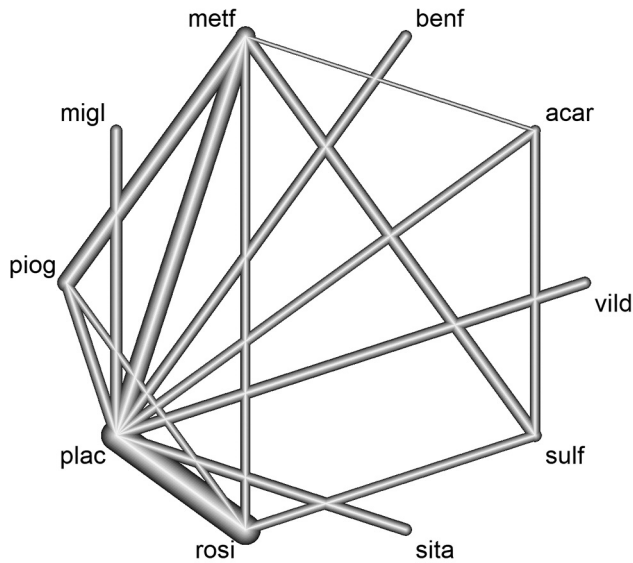


Figure 3. The treatment network for meta-analysis of multiple interventions for investigating efficacy in Diabetes. Nodes represent the treatments in the network and lines represent relative direct comparisons between the treatments. Nodes are weighted by the number of patients receiving each treatment relative to the number of participants in all studies, while the thickness of the line grows proportionally to the number of studies referred to in each comparison.

measurements of *HbA1c* were after a follow-up ranging 3-12 months (113). The studies contained in this data set compared different treatments for blood glucose control in patients with diabetes. The researchers selected 26 studies which consisted of a total of 6,646 patients and 10 drug groups that were acarbose (*acar*), benfluorex (*benf*), metformin (*metf*), miglitol (*migl*), pioglitazone (*piog*), placebo (*plac*), rosiglitazone (*rosi*), sitagliptin (*sita*), sulfonylurea alone (*sulf*), and vildagliptin (*vild*). In the corresponding network, there were 15 different designs (*i.e.*, the set of treatments compared in one study). The data recorded the treatment effect (TE), where the effect was introduced here as MD, the standard error of the effect (*seTE*), the names of the treatments, and finally the name of each study. The effects measure was the MD in blood plasma glucose concentration, while the fixed effects model was used. The visualization of the network (Figure 3) was done *via* the netmeta R statistical package (108, 114). Based on the ranking of treatments (rankogram) in the network with the *P*-score (73) measurement, the top 3 interventions are rosiglitazone treatment which seems to be the most effective (¹*P*-score=0.978), metformin (²*P*-score=0.851), and pioglitazone (³*P*-score=0.768), while the corresponding SUCRA values have very little deviations (rosiglitazone=0.983, metformin=0.852, and pioglitazone=0.766) (108). However, clinicians and decision-makers should not consider an intervention to be best just because it comes first unless the

quality of the evidence used and the confidence in the NMA results are considered (30).

Discussion

In general, NMA can provide increased statistical power when normal network connectivity is possible and sample sizes are sufficient. Mathematical approaches exist to 'measure' network connectivity, but raw data are required to calculate these indicators (39, 115). However, inappropriate use of NMA can lead to erroneous results, such as when there is low network connectivity and therefore low statistical power (1, 44, 116) or when the results are derived from indirect data which, although they remain observations, are nevertheless not interpreted with due care (7, 14). Regarding indirect treatment comparisons, there is disagreement among researchers about the validity of their use in decision-making and especially when direct treatment comparisons are also available (117-119). More specifically, it is argued that decisions should not be made based on rank probabilities alone (especially when treatments are not directly compared) as they may be incompletely informed (120), but also because estimates of rank probabilities are extremely sensitive as they are influenced by factors such as an unequal number of studies per comparison in the network, sample size of individual studies, overall network configuration, and effect sizes between treatments. For example, an unequal number of studies per comparison may lead to biased (121) estimates of treatment rank probabilities for each network considered and thus to an incorrect NMA analysis, as a result of increased variability in the precision of treatment effect estimates (122). For these reasons, it is necessary to provide detailed reports on the strategy researchers intend to follow to assess transitivity and consistency and clarify their calculation methods. Clinicians should also always be cautious about effect sizes and treatment rankings because a good ranking does not necessarily mean a clinically important effect size, and on the other hand, treatment rankings derived from NMAs can often show some degree of inaccuracy (123). This is because their uncertainty can be ignored and so the rankings give the illusion that some interventions are better than others when the relative effects are not different from zero beyond chance (28). However, even so, NMA has serious advantages over pairwise meta-analyses. Especially when there are cycles of evidence (loops), the Bayesian NMA approach has been shown to significantly improve effect estimates compared to separate pairwise meta-analyses (124).

Another equally important aspect that should be considered before constructing an NMA and could help researchers to further improve the results is, as mentioned above, the exploring of the geometry of the network, and by extension the number of nodes (treatments) that will be included in the network, because a decision maker may not

be interested in all pairwise comparisons of the network. Thus, because the therapeutic effect between two active treatments can often be more influential in decision-making than the relative effects of all active treatments versus placebo, researchers could modify the network by using only a subset of available treatments, namely those that are considered clinically more relevant (36) (the most effective treatments). Otherwise, inclusion in the main studies of data comparing treatments without clinical interest provides additional indirect evidence of clinical interest, which may increase the precision of the estimates (9, 125), but may also lead to additional inconsistencies (126). In network studies, it is common to exclude trials and specific comparators based on a variety of different criteria, because choosing to include all possible interventions ever evaluated in an RCT gives unclear and discouraging conclusions. After all, some trials deviate significantly from others and it is not advisable to combine them in NMA (trial-level outliers), where studies suggest corresponding Bayesian outlier detection measures (127). However, deriving treatments from an NMA can significantly change the effect estimates and thus the probability ranking of the most appropriate treatment. Well-connected treatments appear to have the most influence (128). Consequently, the greatest impact on the results occurs when well-connected nodes are removed and so the most evaluated treatments available for a condition must be considered necessary for a network to be valid. Special care is required when it comes to exclusions of potential nodes, and decisions on eligibility criteria must be carefully justified, because small “mutations” in the geometry of the NMA have a direct impact on the analysis and in turn affect the decision-making process. That is why the ‘node-making’ process has been identified as one of the most important problems in NMA, where different ways of generating treatment nodes could significantly affect the results (129-131). But in addition to network size, it is proposed to incorporate the description of specific graph theory statistical measures to complement graph information (132, 133). Particularly for distinguishing similar NMAs, sensitivity analysis is critical to perform when ‘confounding’ is identified in the initial review to infer the absence of heterogeneity, especially when the studies are few (134).

An also very strong reason that the definition of the nodes is critical, is that the interventions are combined with more than one treatment. It is common for researchers to tend to combine treatment arms, where treatments with different characteristics or patients with different subtypes that cannot belong to the same group are merged as one treatment at a node. This has the goal of increasing the statistical power of the network or connecting the network (1), thus introducing bias into the network (135, 136). The simplest approach would be to analyze each combination as a different node in the NMA. Furthermore, evidence has shown that meta-

analysis across multiple smaller RCTs is more valuable than one large RCT (137). As there are always confounding factors in studies that can affect the results, the variation in treatment effect between trials gives a better estimate of the mean effect than an RCT. A simulation study showed that when treatment effects are truly additive, the ‘conventional’ NMA model does not outperform them (138).

A notable venture in NMA that has also taken place very recently and is steadily gaining ground is the incorporation of non-randomized data to assess relative treatment effects, especially in cases with limited randomized data to avoid disconnected network phenomena. By incorporating real-world evidence from non-randomized studies can confirm findings from RCTs, thereby increasing the accuracy of results and empowering the decision-making process (139). Because quality meta-analysis is highly dependent on the availability of individual study data, the use of IPD in NMA is increasingly recognized in the scientific community today. More specifically, the benefits of integrating various proportions of individual patient data (IPD) studies into one NMA and aggregate data (AD) and IPD into the same NMA have been explored. This is because standard NMA methods combine aggregate data from multiple studies, assuming that effect modifiers are balanced across populations (95). New methods such as population fitting methods relax this assumption. One such approach is to analyze IPD from each study in a meta-regression model. IPD-based NMA can lead to increased precision of estimated treatment effects. Additionally, it can help improve network coherence and account for heterogeneity across studies by adjusting participant-level effect modifiers and adopting more advanced models to deal with missing response data. Although the availability of such data is not always feasible, an increased IPD rate has been shown to lead to more accurate estimates for most models (140, 141) and these methods need further evaluation. A typical example is the multilevel network meta-regression (ML-NMR) method as the most recent application, which in this case, is the generalization of NMA for synthesizing data from a mixture of IPD and AD studies that provide estimates for a population decision target (95, 96, 103, 142). This use of meta-analysis, which is also the future of population adjustment, including individual studies, can be extended to areas such as prognostic models and prognostic factors that are particularly important in medical disciplines such as oncology.

Conclusion

As NMA becomes more and more popular and therefore more influential in the scientific community, familiarity with statistical network concepts will be a one-way street as the demands for transparency and more reliable synthesis of the original data increase. Enriching these data belonging to databases for meta-analysis combined with the opinion of

experienced researchers can improve the construction of more reliable predictive models for the desired outcome. But this should be done on the assumption that the construction and study of an NMA should always be based on detailed protocols, as this is the only way to protect against decisions such as the selective use of circumstantial evidence. In any case, NMA as a statistical tool is undoubtedly very useful for evaluating the comparative results of multiple competing interventions in clinical practice and is the ‘next step’ in meta-analysis for further health technology development. However, more specialized training is needed to ensure that the basic methodologies underlying NMAs are understood by health researchers to maximize their ability to interpret and validate these results.

Conflicts of Interest

The Authors declared no potential conflicts of interest in relation to this study.

Authors’ Contributions

Conceptualization: GB; IP. Literature search: GB. Network analysis: GB. Writing original draft: GB. Critically revised the work: GB; IP. Supervised the study: IP.

References

- 1 Ter Veer E, van Oijen MGH and van Laarhoven HWM: The use of (network) meta-analysis in clinical oncology. *Front Oncol* 9: 822, 2019. PMID: 31508373. DOI: 10.3389/fonc.2019.00822
- 2 Gyawali B: Meta-analyses and RCTs in oncology-what is the right balance? *The Lancet Oncology* 19(12): 1565-1566, 2020. DOI: 10.1016/S1470-2045(18)30655-7
- 3 Ge L, Tian JH, Li XX, Song F, Li L, Zhang J, Li G, Pei GQ, Qiu X and Yang KH: Epidemiology characteristics, methodological assessment and reporting of statistical analysis of network meta-analyses in the field of cancer. *Sci Rep* 6: 37208, 2016. PMID: 27848997. DOI: 10.1038/srep37208
- 4 Murad MH, Montori VM, Ioannidis JP, Jaeschke R, Devereaux PJ, Prasad K, Neumann I, Carrasco-Labra A, Agoritsas T, Hatala R, Meade MO, Wyer P, Cook DJ and Guyatt G: How to read a systematic review and meta-analysis and apply the results to patient care: users’ guides to the medical literature. *JAMA* 312(2): 171-179, 2014. PMID: 25005654. DOI: 10.1001/jama.2014.5559
- 5 Dias S and Caldwell DM: Network meta-analysis explained. *Arch Dis Child Fetal Neonatal Ed* 104(1): F8-F12, 2019. PMID: 30425115. DOI: 10.1136/archdischild-2018-315224
- 6 Riley RD, Jackson D, Salanti G, Burke DL, Price M, Kirkham J and White IR: Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 358: j3932, 2017. PMID: 28903924. DOI: 10.1136/bmj.j3932
- 7 Lumley T: Network meta-analysis for indirect treatment comparisons. *Stat Med* 21(16): 2313-2324, 2002. PMID: 12210616. DOI: 10.1002/sim.1201
- 8 Phillippo DM: multinma: Bayesian network meta-analysis of individual and aggregate data. R package version 0.5.0, 2022. DOI: 10.5281/zenodo.3904454
- 9 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 Synth Methods* 3(2): 80-97, 2012. PMID: 26062083. DOI: 10.1002/jrsm.1037
- 10 Jansen JP, Crawford B, Bergman G and Stam W: Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health* 11(5): 956-964, 2008. PMID: 18489499. DOI: 10.1111/j.1524-4733.2008.00347.x
- 11 Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L and Cappelleri JC: 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 Health* 14(4): 417-428, 2011. PMID: 21669366. DOI: 10.1016/j.jval.2011.04.002
- 12 Papakonstantinou T, Nikolakopoulou A, Egger M and Salanti G: In network meta-analysis, most of the information comes from indirect evidence: empirical study. *J Clin Epidemiol* 124: 42-49, 2020. PMID: 32302680. DOI: 10.1016/j.jclinepi.2020.04.009
- 13 Lu G and Ades AE: Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 23(20): 3105-3124, 2004. PMID: 15449338. DOI: 10.1002/sim.1875
- 14 Song F, Altman DG, Glenny AM and Deeks JJ: Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ* 326(7387): 472, 2003. PMID: 12609941. DOI: 10.1136/bmj.326.7387.472
- 15 Mills EJ, Bansback N, Ghement I, Thorlund K, Kelly S, Puhan MA and Wright J: Multiple treatment comparison meta-analyses: a step forward into complexity. *Clin Epidemiol* 3: 193-202, 2011. PMID: 21750628. DOI: 10.2147/CLEPS.16526
- 16 Tonin FS, Rotta I, Mendes AM and Pontarolo R: Network meta-analysis: a technique to gather evidence from direct and indirect comparisons. *Pharm Pract (Granada)* 15(1): 943, 2017. PMID: 28503228. DOI: 10.18549/PharmPract.2017.01.943
- 17 Sutton A, Ades AE, Cooper N and Abrams K: Use of indirect and mixed treatment comparisons for technology assessment. *Pharmacoeconomics* 26(9): 753-767, 2008. PMID: 18767896. DOI: 10.2165/00019053-200826090-00006
- 18 Mavridis D, Giannatsi M, Cipriani A and Salanti G: A primer on network meta-analysis with emphasis on mental health. *Evid Based Ment Health* 18(2): 40-46, 2015. PMID: 25908686. DOI: 10.1136/eb-2015-102088
- 19 Sutton AJ and Higgins JP: Recent developments in meta-analysis. *Stat Med* 27(5): 625-650, 2008. PMID: 17590884. DOI: 10.1002/sim.2934
- 20 Caldwell DM: An overview of conducting systematic reviews with network meta-analysis. *Syst Rev* 3: 109, 2014. PMID: 25267336. DOI: 10.1186/2046-4053-3-109
- 21 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* 9(1): e86754, 2014. PMID: 24466222. DOI: 10.1371/journal.pone.0086754

- 22 Petropoulou M, Nikolakopoulou A, Veroniki AA, Rios P, Vafaei A, Zarin W, Giannatsi M, Sullivan S, Tricco AC, Chaimani A, Egger M and Salanti G: Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. *J Clin Epidemiol* 82: 20-28, 2017. PMID: 27864068. DOI: 10.1016/j.jclinepi.2016.11.002
- 23 Dias S, Sutton AJ, Ades AE and Welton NJ: Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 33(5): 607-617, 2013. PMID: 23104435. DOI: 10.1177/0272989X12458724
- 24 Fanelli D, Costas R and Ioannidis JP: Meta-assessment of bias in science. *Proc Natl Acad Sci U S A* 114(14): 3714-3719, 2017. PMID: 28320937. DOI: 10.1073/pnas.1618569114
- 25 Efthimiou O and White IR: The dark side of the force: Multiplicity issues in network meta-analysis and how to address them. *Res Synth Methods* 11(1): 105-122, 2020. PMID: 31476256. DOI: 10.1002/jrsm.1377
- 26 Shokraneh F and Adams CE: A simple formula for enumerating comparisons in trials and network meta-analysis. *F1000Res* 8: 38, 2019. PMID: 30863537. DOI: 10.12688/f1000research.17352.2
- 27 Salanti G, Kavvoura FK and Ioannidis JP: Exploring the geometry of treatment networks. *Ann Intern Med* 148(7): 544-553, 2008. PMID: 18378949. DOI: 10.7326/0003-4819-148-7-200804010-00011
- 28 Mills EJ, Thorlund K and Ioannidis JP: Demystifying trial networks and network meta-analysis. *BMJ* 346: f2914, 2013. PMID: 23674332. DOI: 10.1136/bmj.f2914
- 29 Rouse B, Chaimani A and Li T: Network meta-analysis: an introduction for clinicians. *Intern Emerg Med* 12(1): 103-111, 2017. PMID: 27913917. DOI: 10.1007/s11739-016-1583-7
- 30 Mbuagbaw L, Rochweg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L and Guyatt GH: Approaches to interpreting and choosing the best treatments in network meta-analyses. *Syst Rev* 6(1): 79, 2017. PMID: 28403893. DOI: 10.1186/s13643-017-0473-z
- 31 Chiocchia V, Nikolakopoulou A, Papakonstantinou T, Egger M and Salanti G: Agreement between ranking metrics in network meta-analysis: an empirical study. *BMJ Open* 10(8): e037744, 2020. PMID: 32819946. DOI: 10.1136/bmjopen-2020-037744
- 32 Veroniki AA, Straus SE, Rucker G and Tricco AC: Is providing uncertainty intervals in treatment ranking helpful in a network meta-analysis? *J Clin Epidemiol* 100: 122-129, 2018. PMID: 29432861. DOI: 10.1016/j.jclinepi.2018.02.009
- 33 Salanti G, Ades AE and Ioannidis JP: Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 64(2): 163-171, 2011. PMID: 20688472. DOI: 10.1016/j.jclinepi.2010.03.016
- 34 Ahn E and Kang H: Concepts and emerging issues of network meta-analysis. *Korean J Anesthesiol* 74(5): 371-382, 2021. PMID: 34551467. DOI: 10.4097/kja.21358
- 35 Catalá-López F, Tobías A, Cameron C, Moher D and Hutton B: Network meta-analysis for comparing treatment effects of multiple interventions: an introduction. *Rheumatol Int* 34(11): 1489-1496, 2014. PMID: 24691560. DOI: 10.1007/s00296-014-2994-2
- 36 Efthimiou O, Debray TP, van Valkenhoef G, Trelle S, Panayidou K, Moons KG, Reitsma JB, Shang A, Salanti G and GetReal Methods Review Group: GetReal in network meta-analysis: a review of the methodology. *Res Synth Methods* 7(3): 236-263, 2016. PMID: 26754852. DOI: 10.1002/jrsm.1195
- 37 Phillips MR, Steel DH, Wykoff CC, Busse JW, Bannuru RR, Thabane L, Bhandari M, Chaudhary V and Retina Evidence Trials International Alliance (R.E.T.I.N.A.) Study Group: A clinician's guide to network meta-analysis. *Eye (Lond)* 36(8): 1523-1526, 2022. PMID: 35145277. DOI: 10.1038/s41433-022-01943-5
- 38 Jansen JP and Naci H: Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med* 11: 159, 2013. PMID: 23826681. DOI: 10.1186/1741-7015-11-159
- 39 Fernández-Castilla B and Van den Noortgate W: Network meta-analysis in psychology and educational sciences: A systematic review of their characteristics. *Behav Res Methods*, 2022. PMID: 35821493. DOI: 10.3758/s13428-022-01905-5
- 40 Le-Rademacher J and Wang X: Time-to-event data: an overview and analysis considerations. *J Thorac Oncol* 16(7): 1067-1074, 2021. PMID: 33887465. DOI: 10.1016/j.jtho.2021.04.004
- 41 Spruance SL, Reid JE, Grace M and Samore M: Hazard ratio in clinical trials. *Antimicrob Agents Chemother* 48(8): 2787-2792, 2004. PMID: 15273082. DOI: 10.1128/AAC.48.8.2787-2792.2004
- 42 Cox D: Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)* 34(2): 187-202, 2022. DOI: 10.1111/j.2517-6161.1972.tb00899.x
- 43 Spineli LM: Modeling missing binary outcome data while preserving transitivity assumption yielded more credible network meta-analysis results. *J Clin Epidemiol* 105: 19-26, 2019. PMID: 30223064. DOI: 10.1016/j.jclinepi.2018.09.002
- 44 Cipriani A, Higgins JP, Geddes JR and Salanti G: Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 159(2): 130-137, 2013. PMID: 23856683. DOI: 10.7326/0003-4819-159-2-201307160-00008
- 45 Baker SG and Kramer BS: The transitive fallacy for randomized trials: if A bests B and B bests C in separate trials, is A better than C? *BMC Med Res Methodol* 2: 13, 2002. PMID: 12429069. DOI: 10.1186/1471-2288-2-13
- 46 Xiong T, Parekh-Bhurke S, Loke YK, Abdelhamid A, Sutton AJ, Eastwood AJ, Holland R, Chen YF, Walsh T, Glenny AM and Song F: Overall similarity and consistency assessment scores are not sufficiently accurate for predicting discrepancy between direct and indirect comparison estimates. *J Clin Epidemiol* 66(2): 184-191, 2013. PMID: 23186991. DOI: 10.1016/j.jclinepi.2012.06.022
- 47 Katsanos K: Appraising Inconsistency between Direct and Indirect Estimates (5th Section, Chapter 12). In: *Network Meta-Analysis: Evidence Synthesis with Mixed Treatment Comparison*. Biondi-Zoccai G (ed.). Hauppauge, NY, USA, Nova Science Publishers, pp. 191-210, 2014.
- 48 Lu G and Ades A: Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association* 101(474): 447-459, 2020. DOI: 10.1198/016214505000001302

- 49 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* 50(6): 683-691, 1997. PMID: 9250266. DOI: 10.1016/s0895-4356(97)00049-8
- 50 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* 3(2): 111-125, 2012. PMID: 26062085. DOI: 10.1002/jrsm.1045
- 51 Dias S, Welton NJ and Ades AE: Study designs to detect sponsorship and other biases in systematic reviews. *J Clin Epidemiol* 63(6): 587-588, 2010. PMID: 20434021. DOI: 10.1016/j.jclinepi.2010.01.005
- 52 Cavanaugh JE and Neath AA: The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. *Wiley Interdiscip Rev Comput Stat* 11(3): e1460, 2019. DOI: 10.1002/wics.1460
- 53 Spiegelhalter DJ, Best NG, Carlin BP and van der Linde A: Bayesian measures of model complexity and fit. *J R Statist Soc B Stat Methodol* 64(4): 583-639, 2002. DOI: 10.1111/1467-9868.00353
- 54 Freeman SC, Fisher D, White IR, Auperin A and Carpenter JR: Identifying inconsistency in network meta-analysis: Is the net heat plot a reliable method? *Stat Med* 38(29): 5547-5564, 2019. PMID: 31647136. DOI: 10.1002/sim.8383
- 55 Krahn U, Binder H and König J: A graphical tool for locating inconsistency in network meta-analyses. *BMC Med Res Methodol* 13: 35, 2013. PMID: 23496991. DOI: 10.1186/1471-2288-13-35
- 56 Dias S, Welton NJ, Caldwell DM and Ades AE: Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 29(7-8): 932-944, 2010. PMID: 20213715. DOI: 10.1002/sim.3767
- 57 van Valkenhoef G, Dias S, Ades AE and Welton NJ: Automated generation of node-splitting models for assessment of inconsistency in network meta-analysis. *Res Synth Methods* 7(1): 80-93, 2016. PMID: 26461181. DOI: 10.1002/jrsm.1167
- 58 Yu-Kang T: Node-splitting generalized linear mixed models for evaluation of inconsistency in network meta-analysis. *Value Health* 19(8): 957-963, 2016. PMID: 27987646. DOI: 10.1016/j.jval.2016.07.005
- 59 Beyene J, Bonner AJ and Neupane B: Choosing the statistical model and between fixed and random effects (Chapter 8). In: *Network Meta-Analysis: Evidence Synthesis with Mixed Treatment Comparison*. Biondi-Zoccai G (ed.). Hauppauge, NY, USA, Nova Science Publishers, pp.117-140, 2014.
- 60 Gagnier JJ: Appraising Between-Study Heterogeneity (5th Section, Chapter 11). In: *Network Meta-Analysis: Evidence Synthesis with Mixed Treatment Comparison*. Biondi-Zoccai G (ed.). Hauppauge, NY, USA, Nova Science Publishers, pp. 171-190, 2014.
- 61 Pereira TV, Patsopoulos NA, Salanti G and Ioannidis JP: Critical interpretation of Cochran's Q test depends on power and prior assumptions about heterogeneity. *Res Synth Methods* 1(2): 149-161, 2010. PMID: 26061380. DOI: 10.1002/jrsm.13
- 62 Huedo-Medina TB, Sánchez-Meca J, Marín-Martínez F and Botella J: Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychol Methods* 11(2): 193-206, 2006. PMID: 16784338. DOI: 10.1037/1082-989X.11.2.193
- 63 Patil KD: Cochran's Q test: Exact distribution. *J Am Stat Assoc* 70(349): 186-189, 1975. DOI: 10.1080/01621459.1975.10480285
- 64 Cochran WG: The combination of estimates from different experiments. *Biometrics* 10(1): 101-129, 1954. DOI: 10.2307/3001666
- 65 Piepho HP, Madden LV, Roger J, Payne R and Williams ER: Estimating the variance for heterogeneity in arm-based network meta-analysis. *Pharm Stat* 17(3): 264-277, 2018. PMID: 29676023. DOI: 10.1002/pst.1857
- 66 Higgins JP, Thompson SG, Deeks JJ and Altman DG: Measuring inconsistency in meta-analyses. *BMJ* 327(7414): 557-560, 2003. PMID: 12958120. DOI: 10.1136/bmj.327.7414.557
- 67 von Hippel PT: The heterogeneity statistic I(2) can be biased in small meta-analyses. *BMC Med Res Methodol* 15: 35, 2015. PMID: 25880989. DOI: 10.1186/s12874-015-0024-z
- 68 Kao YS, Ma KS, Wu MY, Wu YC, Tu YK and Hung CH: Topical prevention of radiation dermatitis in head and neck cancer patients: a network meta-analysis. *In Vivo* 36(3): 1453-1460, 2022. PMID: 35478163. DOI: 10.21873/invivo.12851
- 69 Jansen JP and Cope S: Meta-regression models to address heterogeneity and inconsistency in network meta-analysis of survival outcomes. *BMC Med Res Methodol* 12: 152, 2012. PMID: 23043545. DOI: 10.1186/1471-2288-12-152
- 70 Baker WL, White CM, Cappelleri JC, Kluger J, Coleman CI and Health Outcomes, Policy, and Economics (HOPE) Collaborative Group: Understanding heterogeneity in meta-analysis: the role of meta-regression. *Int J Clin Pract* 63(10): 1426-1434, 2009. PMID: 19769699. DOI: 10.1111/j.1742-1241.2009.02168.x
- 71 Nikolakopoulou A, Higgins JPT, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M and Salanti G: CINeMA: An approach for assessing confidence in the results of a network meta-analysis. *PLoS Med* 17(4): e1003082, 2020. PMID: 32243458. DOI: 10.1371/journal.pmed.1003082
- 72 Turner RM, Domínguez-Islas CP, Jackson D, Rhodes KM and White IR: Incorporating external evidence on between-trial heterogeneity in network meta-analysis. *Stat Med* 38(8): 1321-1335, 2019. PMID: 30488475. DOI: 10.1002/sim.8044
- 73 Rucker G and Schwarzer G: Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol* 15: 58, 2015. PMID: 26227148. DOI: 10.1186/s12874-015-0060-8
- 74 Antoniou SA, Koelemay M, Antoniou GA and Mavridis D: A practical guide for application of network meta-analysis in evidence synthesis. *Eur J Vasc Endovasc Surg* 58(1): 141-144, 2019. PMID: 30528457. DOI: 10.1016/j.ejvs.2018.10.023
- 75 Van de Schoot R, Depaoli S, King R, Kramer B, Märtens K, Tadesse MG, Vannucci M, Gelman A, Veen D, Willemsen J and Yau C: Bayesian statistics and modelling. *Nature Reviews Methods Primers* 1(1), 2022. DOI: 10.1038/s43586-020-00001-2
- 76 López Puga J, Krzywinski M and Altman N: Points of significance: Bayes' theorem. *Nat Methods* 12(4): 277-278, 2015. PMID: 26005726. DOI: 10.1038/nmeth.3335
- 77 Hackenberger BK: Bayesian meta-analysis now - let's do it. *Croat Med J* 61(6): 564-568, 2020. PMID: 33410305. DOI: 10.3325/cmj.2020.61.564
- 78 Stone JV: *Bayes' rule: A tutorial introduction to Bayesian analysis*. Sebtel Press, 2013. DOI: 10.13140/2.1.1371.6801
- 79 Van Valkenhoef G, Tervonen T, de Brock B and Hillege H: Algorithmic parameterization of mixed treatment comparisons. *Stat Comput* 22(5): 1099-1111, 2012. DOI: 10.1007/s11222-011-9281-9

- 80 Lunn DJ, Thomas A, Best N and Spiegelhalter D: WinBUGS - A Bayesian modeling framework: Concepts, structure, and extensibility. *Stat Comput* 10(4): 325-337, 2000. DOI: 10.1023/a:1008929526011
- 81 Hoff PD: A first course in Bayesian statistical methods. Springer Texts in Statistics, 2009. DOI: 10.1007/978-0-387-92407-6
- 82 Lunn DJ, Best N, Thomas A, Wakefield J and Spiegelhalter D: Bayesian analysis of population PK/PD models: general concepts and software. *J Pharmacokinet Pharmacodyn* 29(3): 271-307, 2002. PMID: 12449499. DOI: 10.1023/a:1020206907668
- 83 Harrer M, Cuijpers P, Furukawa T and Ebert D: Doing Meta-Analysis with R. 2021. DOI: 10.1201/9781003107347
- 84 R Core Team: R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2014. Available at: <http://www.R-project.org> [Last accessed on March 31, 2023]
- 85 Owen RK, Bradbury N, Xin Y, Cooper N and Sutton A: MetaInsight: An interactive web-based tool for analyzing, interrogating, and visualizing network meta-analyses using R-shiny and netmeta. *Res Synth Methods* 10(4): 569-581, 2019. PMID: 31349391. DOI: 10.1002/jrsm.1373
- 86 Plummer M: JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria, pp. 1-10, 2003. Available at: <https://www.r-project.org/conferences/DSC-2003/Proceedings/Plummer.pdf> [Last accessed on March 31, 2023]
- 87 van Valkenhoef G, Lu G, de Brock B, Hillege H, Ades AE and Welton NJ: Automating network meta-analysis. *Res Synth Methods* 3(4): 285-299, 2012. PMID: 26053422. DOI: 10.1002/jrsm.1054
- 88 Plummer M: rjags: Bayesian Graphical Models Using MCMC. R package version 4-6, 2016. Available at: <https://CRAN.R-project.org/package=rjags> [Last accessed on March 31, 2023]
- 89 Lunn D, Spiegelhalter D, Thomas A and Best N: The BUGS project: Evolution, critique and future directions. *Stat Med* 28(25): 3049-3067, 2009. PMID: 19630097. DOI: 10.1002/sim.3680
- 90 Béliveau A, Boyne DJ, Slater J, Brenner D and Arora P: BUGSnet: an R package to facilitate the conduct and reporting of Bayesian network Meta-analyses. *BMC Med Res Methodol* 19(1): 196, 2019. PMID: 31640567. DOI: 10.1186/s12874-019-0829-2
- 91 Van Valkenhoef G and Kuiper J: gemtc: Network Meta-Analysis Using Bayesian Methods. R package version 0.8-2, 2016. Available at: <https://CRAN.R-project.org/package=gemtc> [Last accessed on March 31, 2023]
- 92 Spiegelhalter D, Thomas A, Best N and Lunn D: OpenBUGS user manual, version 3.2.3. MRC Biostatistics Unit, Cambridge, 2014. Available at: https://www.mrc-bsu.cam.ac.uk/wp-content/uploads/2021/06/OpenBUGS_Manual.pdf [Last accessed on March 31, 2023]
- 93 Lawson A: Using R for Bayesian spatial and spatio-temporal health modeling. 2021. DOI: 10.1201/9781003043997
- 94 Sturtz S, Ligges U and Gelman A: R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software* 12(3): 1-16, 2015. DOI: 10.18637/jss.v012.i03
- 95 Phillippo DM, Dias S, Ades AE, Belger M, Brnabic A, Schacht A, Saure D, Kadziola Z and Welton NJ: Multilevel network meta-regression for population-adjusted treatment comparisons. *J R Stat Soc Ser A Stat Soc* 183(3): 1189-1210, 2020. PMID: 32684669. DOI: 10.1111/rssa.12579
- 96 Singh J, Gsteiger S, Wheaton L, Riley RD, Abrams KR, Gillies CL and Bujkiewicz S: Bayesian network meta-analysis methods for combining individual participant data and aggregate data from single arm trials and randomised controlled trials. *BMC Med Res Methodol* 22(1): 186, 2022. PMID: 35818035. DOI: 10.1186/s12874-022-01657-y
- 97 Pandey S, Sharma A and Siddiqui M: NM3 extending the network meta-analysis (NMA) framework to multilevel network meta-regression (ML-NMR): a worked example of ML-NMR vs. standard NMA. *Value in Health* 23: S407, 2021. DOI: 10.1016/j.jval.2020.08.058
- 98 Freeman SC: Individual patient data meta-analysis and network meta-analysis. *Methods Mol Biol* 2345: 279-298, 2022. PMID: 34550597. DOI: 10.1007/978-1-0716-1566-9_17
- 99 Phillippo DM: Calibration of treatment effects in network meta-analysis using individual patient data. Ph.D. thesis, University of Bristol, UK, 2019. Available at: <https://research-information.bris.ac.uk/en/studentTheses/calibration-of-treatment-effects-in-network-meta-analysis-using-i> [Last accessed on March 31, 2023]
- 100 Jansen JP: Network meta-analysis of individual and aggregate level data. *Res Synth Methods* 3(2): 177-190, 2012. PMID: 26062089. DOI: 10.1002/jrsm.1048
- 101 Kanters S, Karim ME, Thorlund K, Anis AH, Zoratti M and Bansback N: Comparing the use of aggregate data and various methods of integrating individual patient data to network meta-analysis and its application to first-line ART. *BMC Med Res Methodol* 21(1): 60, 2021. PMID: 33784981. DOI: 10.1186/s12874-021-01254-5
- 102 Annis J, Miller BJ and Palmeri TJ: Bayesian inference with Stan: A tutorial on adding custom distributions. *Behav Res Methods* 49(3): 863-886, 2017. PMID: 27287444. DOI: 10.3758/s13428-016-0746-9
- 103 Carpenter B, Gelman A, Hoffman M, Lee D, Goodrich B, Betancourt M, Brubaker M, Guo J, Li P and Riddell A: Stan: a probabilistic programming language. *Journal of Statistical Software* 76(1): 1-32, 2017. DOI: 10.18637/jss.v076.i01
- 104 Hoffman MD and Gelman A: The no-U-turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res* 15(47): 1593-1623, 2014. DOI: 10.48550/arXiv.1111.4246
- 105 Egidio L, Hansson A and Wahlberg B: Learning the step-size policy for the limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm. 2021 International Joint Conference on Neural Networks (IJCNN), 2022. DOI: 10.1109/IJCNN52387.2021.9534194
- 106 Martin A, Quinn K and Park J: MCMCpack: Markov Chain Monte Carlo in R. *Journal of Statistical Software* 42(9): 1-21, 2015. DOI: 10.18637/jss.v042.i09
- 107 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* 9(12): e115065, 2014. PMID: 25541687. DOI: 10.1371/journal.pone.0115065
- 108 Rucker G, Petropoulou M and Schwarzer G: Network meta-analysis of multicomponent interventions. *Biom J* 62(3): 808-821, 2020. PMID: 31021449. DOI: 10.1002/bimj.201800167
- 109 Shim SR, Kim SJ, Lee J and Rucker G: Network meta-analysis: application and practice using R software. *Epidemiol Health* 41: e2019013, 2019. PMID: 30999733. DOI: 10.4178/epih.e2019013

- 110 Zhang J, Carlin BP, Neaton JD, Soon GG, Nie L, Kane R, Virnig BA and Chu H: Network meta-analysis of randomized clinical trials: reporting the proper summaries. *Clin Trials* 11(2): 246-262, 2014. PMID: 24096635. DOI: 10.1177/1740774513498322
- 111 Lin L, Zhang J, Hodges JS and Chu H: Performing arm-based network meta-analysis in R with the pnetmeta package. *J Stat Softw* 80: 5, 2017. PMID: 28883783. DOI: 10.18637/jss.v080.i05
- 112 Denwood M: runjags: an R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *Journal of Statistical Software* 71(9): 1-25, 2016. DOI: 10.18637/jss.v071.i09
- 113 Senn S, Gavini F, Magrez D and Scheen A: Issues in performing a network meta-analysis. *Stat Methods Med Res* 22(2): 169-189, 2013. PMID: 22218368. DOI: 10.1177/0962280211432220
- 114 Schwarzer G, Carpenter JR and Rücker G: *Meta-analysis with R*. Use R! Cham, Springer, 2015. Available at: https://zums.ac.ir/files/socialfactors/files/Meta-Analysis_with_R-2015.pdf [Last accessed on March 30, 2023]
- 115 Thom H, White IR, Welton NJ and Lu G: Automated methods to test connectedness and quantify indirectness of evidence in network meta-analysis. *Res Synth Methods* 10(1): 113-124, 2019. PMID: 30403829. DOI: 10.1002/jrsm.1329
- 116 Salanti G, Higgins JP, Ades AE and Ioannidis JP: Evaluation of networks of randomized trials. *Stat Methods Med Res* 17(3): 279-301, 2008. PMID: 17925316. DOI: 10.1177/0962280207080643
- 117 Edwards SJ, Clarke MJ, Wordsworth S and Borrill J: Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *Int J Clin Pract* 63(6): 841-854, 2009. PMID: 19490195. DOI: 10.1111/j.1742-1241.2009.02072.x
- 118 Gartlehner G and Moore CG: Direct *versus* indirect comparisons: a summary of the evidence. *Int J Technol Assess Health Care* 24(2): 170-177, 2008. PMID: 18400120. DOI: 10.1017/S0266462308080240
- 119 Ioannidis JP: Indirect comparisons: the mesh and mess of clinical trials. *Lancet* 368(9546): 1470-1472, 2006. PMID: 17071265. DOI: 10.1016/S0140-6736(06)69615-3
- 120 Kibret T, Richer D and Beyene J: Bias in identification of the best treatment in a Bayesian network meta-analysis for binary outcome: a simulation study. *Clin Epidemiol* 6: 451-460, 2014. PMID: 25506247. DOI: 10.2147/CLEP.S69660
- 121 Phillippo DM, Dias S, Ades AE, Didelez V and Welton NJ: Sensitivity of treatment recommendations to bias in network meta-analysis. *J R Stat Soc Ser A Stat Soc* 181(3): 843-867, 2018. PMID: 30449954. DOI: 10.1111/rssa.12341
- 122 Davies AL and Galla T: Degree irregularity and rank probability bias in network meta-analysis. *Res Synth Methods* 12(3): 316-332, 2021. PMID: 32935913. DOI: 10.1002/jrsm.1454
- 123 Trinquart L, Attiche N, Bafeta A, Porcher R and Ravaud P: Uncertainty in treatment rankings: Reanalysis of network meta-analyses of randomized trials. *Ann Intern Med* 164(10): 666-673, 2016. PMID: 27089537. DOI: 10.7326/M15-2521
- 124 Lin L, Chu H and Hodges JS: On evidence cycles in network meta-analysis. *Stat Interface* 13(4): 425-436, 2020. PMID: 32742550. DOI: 10.4310/sii.2020.v13.n4.a1
- 125 Li T, Puhan MA, Vedula SS, Singh S, Dickersin K and Ad Hoc Network Meta-analysis Methods Meeting Working Group: Network meta-analysis-highly attractive but more methodological research is needed. *BMC Med* 9: 79, 2011. PMID: 21707969. DOI: 10.1186/1741-7015-9-79
- 126 Sturtz S and Bender R: Unsolved issues of mixed treatment comparison meta-analysis: network size and inconsistency. *Res Synth Methods* 3(4): 300-311, 2012. PMID: 26053423. DOI: 10.1002/jrsm.1057
- 127 Zhang J, Fu H and Carlin BP: Detecting outlying trials in network meta-analysis. *Stat Med* 34(19): 2695-2707, 2015. PMID: 25851533. DOI: 10.1002/sim.6509
- 128 Mills EJ, Kanter S, Thorlund K, Chaimani A, Veroniki AA and Ioannidis JP: The effects of excluding treatments from network meta-analyses: survey. *BMJ* 347: f5195, 2013. PMID: 24009242. DOI: 10.1136/bmj.f5195
- 129 Naudet F, Schuit E and Ioannidis JPA: Overlapping network meta-analyses on the same topic: survey of published studies. *Int J Epidemiol* 46(6): 1999-2008, 2017. PMID: 29040566. DOI: 10.1093/ije/dyx138
- 130 James A, Yavchitz A, Ravaud P and Boutron I: Node-making process in network meta-analysis of nonpharmacological treatment are poorly reported. *J Clin Epidemiol* 97: 95-102, 2018. PMID: 29196202. DOI: 10.1016/j.jclinepi.2017.11.018
- 131 Xing A and Lin L: Effects of treatment classifications in network meta-analysis. *Research Methods in Medicine & Health Sciences* 1(1): 12-24, 2020. DOI: 10.1177/2632084320932756
- 132 Tonin FS, Borba HH, Mendes AM, Wiens A, Fernandez-Llimos F and Pontarolo R: Description of network meta-analysis geometry: A metrics design study. *PLoS One* 14(2): e0212650, 2019. PMID: 30785955. DOI: 10.1371/journal.pone.0212650
- 133 Rücker G: Network meta-analysis, electrical networks and graph theory. *Res Synth Methods* 3(4): 312-324, 2012. PMID: 26053424. DOI: 10.1002/jrsm.1058
- 134 Jackson D, Barrett JK, Rice S, White IR and Higgins JP: A design-by-treatment interaction model for network meta-analysis with random inconsistency effects. *Stat Med* 33(21): 3639-3654, 2014. PMID: 24777711. DOI: 10.1002/sim.6188
- 135 Lunny C, Tricco AC, Veroniki AA, Dias S, Hutton B, Salanti G, Wright JM, White I and Whiting P: Methodological review to develop a list of bias items used to assess reviews incorporating network meta-analysis: protocol and rationale. *BMJ Open* 11(6): e045987, 2021. PMID: 34168027. DOI: 10.1136/bmjopen-2020-045987
- 136 Cameron C, Fireman B, Hutton B, Clifford T, Coyle D, Wells G, Dormuth CR, Platt R and Toh S: Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. *Syst Rev* 4: 147, 2015. PMID: 26537988. DOI: 10.1186/s13643-015-0133-0
- 137 Int'Hout J, Ioannidis JP and Borm GF: Obtaining evidence by a single well-powered trial or several modestly powered trials. *Stat Methods Med Res* 25(2): 538-552, 2016. PMID: 23070590. DOI: 10.1177/0962280212461098
- 138 Thorlund K and Mills E: Stability of additive treatment effects in multiple treatment comparison meta-analysis: a simulation study. *Clin Epidemiol* 4: 75-85, 2012. PMID: 22570567. DOI: 10.2147/CLEP.S29470

- 139 Efthimiou O, Mavridis D, Debray TP, Samara M, Belger M, Siontis GC, Leucht S, Salanti G and GetReal Work Package 4: Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med* 36(8): 1210-1226, 2017. PMID: 28083901. DOI: 10.1002/sim.7223
- 140 Leahy J, O'Leary A, Afdhal N, Gray E, Milligan S, Wehmeyer MH and Walsh C: The impact of individual patient data in a network meta-analysis: An investigation into parameter estimation and model selection. *Res Synth Methods* 9(3): 441-469, 2018. PMID: 29923679. DOI: 10.1002/jrsm.1305
- 141 Debray TP, Schuit E, Efthimiou O, Reitsma JB, Ioannidis JP, Salanti G, Moons KG and GetReal Workpackage: An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res* 27(5): 1351-1364, 2018. PMID: 27487843. DOI: 10.1177/0962280216660741
- 142 Kanters S, Karim ME, Thorlund K, Anis A and Bansback N: When does the use of individual patient data in network meta-analysis make a difference? A simulation study. *BMC Med Res Methodol* 21(1): 21, 2021. PMID: 33435879. DOI: 10.1186/s12874-020-01198-2

Received March 13, 2023

Revised March 27, 2023

Accepted March 31, 2023