We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Network Meta-Analysis Using R for Diabetes Data

Nilgün Yildiz

Abstract

The objective of a meta-analysis is usually to estimate the overall treatment effect and make inferences about the difference between the effects of the two treatments. Meta-analysis is a quantitative method commonly used to combine the results of multiple studies in the medical and social sciences. There are three common types of meta-analysis. Pairwise, Multivariate and Network Meta-analysis. In general, network meta-analysis (NMA) offers the advantage of enabling the combined assessment of more than two treatments. Statistical approaches to NMA are largely classified as frequentist and Bayesian frameworks. Because part of NMA has indirect, multiple comparisons, As reports of network meta-analysis become more common, it is essential to introduce the approach to readers and to provide guidance as to how to interpret the results. In this chapter, the terms used in NMA are defined, relevant statistical concepts are summarized, and the NMA analytic process based on the frequentist and Bayesian framework is illustrated using the R program and an example of a network involving diabetes treatments. The aim of the article is to compare the basic concepts and analyzes of network meta-analysis using diabetes data and the treatment methods used.

Keywords: Network meta-analysis, fixed effect model, random-effects model, forest plot, network graph, direct evidence plot

1. Introduction

Meta-analysis is used to synthesize the results of more than one study and overall effect size is considered to be valid only when some required assumptions are satisfied [1]. An increasing number of options for alternative medical treatment has given rise to the need for comparative effectiveness research [2, 3]. A randomized, controlled trials used to compare different treatment options are generally seen to be infeasible, there is a need for other methodological approaches. Since it makes it possible to combine data from many different studies so that a total estimate of treatment effect can be provided, a meta-analysis integrated into a systematic review is generally seen to be a useful statistical tool. On the other hand, there is an important limitation of standard meta-analysis; only two interventions can be compared at a time. When you have several treatment options to capitalize on, only partial information can be provided by a series of individual meta-analysis since only the questions about pairs of treatments can be answered in this way, which leads to difficulties in making optimal clinical decisions since each meta-analysis is just one constituent of the whole picture.

There is an increasing need for a method to be used to summarize evidence across many interventions [4]. In order to assess a number of interventions in terms of their relative effectiveness and to synthesize evidence from a set of randomized trials, network meta-analysis (or multiple treatments meta-analysis or mixed-treatment comparison) was created [5–7]. This method is built on the analysis of direct evidence (coming from research that directly randomizes treatments of interest) and indirect evidence (coming from research that compares treatments of interest with a common comparator) [8]. The benefits incurred by network analysis becoming increasingly popular have been reported in some applications and methodological articles [2, 9, 10]. **Figure 1** shows the number of network meta-analysis (NMA) studies that have been published.

Despite the fact that network meta-analysis shares many underlying assumptions with pairwise meta-analysis, it is not so much accepted as pairwise meta-analysis and thus criticized more [11].

The assumptions required by NMA about similarity, transitivity, and consistency [12–17] are methodologically, logically, and statistically more strict [18, 19] because it should be examined whether each of these is satisfied or not [15, 20, 21].

For NMA, there are some methods to calculate the contribution of direct (and indirect) evidence of each comparison to its own NMA estimate, but how to define the contribution of each study to another estimate of treatment effect is an issue of greater ambiguity. There are a number of proposals made in the literature, each of which is based on a different approach but many of them are not without limitations and generally, there are contradictions between their results [20, 22–24]. There are some investigations having been conducted on the proportions of direct and indirect evidence in the past. One of these is the method of “back calculation” [21] introduced by Dias and some others have been proposed within a Bayesian framework [25]. There is even one proposed within a frequentist context [13]. In inverse variance method-based NMA, NMA estimates refer to linear combinations of treatment effect estimates from primary studies having coefficients that make up the rows of the hat matrix. It is easy to obtain the direct evidence proportion of a study or a comparison from the diagonal elements that the respective hat matrix has [13]. Dias and others proposed “node splitting” as an alternative. Node splitting refers to the estimation of the indirect evidence for comparison by modeling out all studies providing direct information for this comparison [25]. Additions were made to this method [26] and called “side splitting” by others [9]. There are different

Figure 1.
Number of network meta-analysis publications (search PubMed until January 2020).

interpretations of the term “side” in the literature; for example, it was interpreted as an edge in the network graph by White [9] while it was interpreted as SIDE, an abbreviation of “Separating Indirect and Direct Evidence” by others [27, 28]. There is another way of quantifying the indirect evidence proposed by Noma and others, including the factorization of the total likelihood into separate component likelihoods [14]. Yet, none of these authors have attempted to make a definition or estimation of the contribution of each study to a given comparison in the network.

There are six basic steps that every NMA should follow, regardless of the analytic model chosen. These steps include

1. Understand network geometry,
2. Understand key concepts and assumptions,
3. Conduct analysis and present results,
4. Examine model assumptions through local and global tests,
5. Create a hierarchy of competing interventions (ranking),
6. Conduct heterogeneity and sensitivity analyses.

The network plot is fundamental to an NMA because it helps visualize the available studies and few of evidence across the multiple comparisons. In such a plot, each treatment/comparator identified in the review is represented by a node, and direct evidence comparing two interventions (i.e., studies which directly compared these two interventions) are represented via edges, connecting the respective nodes. The network plot of our example is presented in **Figure 2**.

In **Figure 2**, a network of treatments for type 2 diabetes is shown. The function served by the lines between the treatment nodes is to show which comparisons have been made in trials that are randomized. The absence of a line between two nodes means that there are no studies (that is, no direct evidence) comparing the two

Figure 2.
A graph of the network generated by using the net graph function for the diabetes data.

drugs. A network meta-analysis refers to an analysis of the data from all of these randomized trials at the same time. By means of a network meta-analysis, it is possible to estimate the relative effectiveness of two treatments even if they are not compared by any studies. For example, no comparison has been made between rosi and acar in any study but by using a common comparator (placebo), an indirect comparison can be made between them. After denoting rosi, acar, and placebo as treatments A, B, and C, respectively, it is possible to have an indirect comparison (AB) by subtracting the meta-analytic estimates of all studies of acar versus placebo (BC) from the estimate of all studies of rosi versus placebo (AC): $AB \text{ indirect meta-analysis} = AC \text{ direct meta-analysis} - BC \text{ direct meta-analysis}$. If there is direct evidence (such as metf vs. sulf in **Figure 2**), direct and indirect estimates can be combined by the network meta-analysis and mixed effect size can be calculated as the weighted average of the direct evidence (studies comparing metf and sulf directly) and the indirect evidence (for example, studies comparing metf and acar via placebo). The network constructed by studies of metf versus acar, metf versus placebo, and acar versus placebo is often named as a loop of evidence. By using indirect estimates, information can be provided on comparisons for which there are no trials. In this way, the accuracy of the direct estimate can be enhanced through the reduction of the width of the CIs in comparison with the direct evidence alone [9].

In a network meta-analysis, all the direct and indirect evidence can be utilized. Empirical studies have concluded that compared to a single direct or indirect estimate, it can produce more precise estimates of the intervention effects [2, 29]. Moreover, network meta-analysis has the potential of yielding data for comparisons made between pairs of interventions having never been evaluated within individual randomized trials. The comparison of all interventions of interest simultaneously in the same analysis makes it possible to estimate their ranking relatively for a given result. The purpose of this study is to show how analysis can be done with the network meta-analysis method using the R package program. Network meta-analysis as a functional method. It is to show that it can be done flexibly and easily with the R program to help researchers interested in this subject.

This chapter is organized as follows, In the next sections, we present a review of the methods for NMA as identified in our literature search. In Section 2, we present key concepts and the basic methodology for NMA. In Section 3 Diabetes treatments data is used as an example. The last section presents conclusions about our research and results found by network meta-analysis of diabetes data using the R program.

2. Conceptual issues and underlying assumptions in network meta-analysis

There may be different alternatives for the treatment of the same health condition and what makes NMA special is that through the synthesis of direct and indirect estimates for their relative effects, it allows the selection of the best treatment. Head-to-head studies can be conducted to directly compare two treatments A and B (AB studies). It is also possible to get an indirect estimate from studies in which these two treatments are compared with a common comparator treatment C, namely, AC and BC studies [9]. If we have both direct and indirect estimates, then we can combine them to estimate a mixed-treatment effect, as you can see in the left panel of **Figure 3**. In practice, there are numerous interventions for most health conditions that have been compared in various randomized trials and build a network of evidence. For the comparison of treatments within such a network, there may be direct and many different indirect estimates obtained through many different comparators, as illustrated in the example in the right panel of **Figure 2**.

Figure 3.
Each circle represents an intervention, and lines represent direct comparisons.

Using NMA, all these different pieces of information can be compared so that an internally consistent overall estimate of the relative effects of all treatments can be produced. Researchers are still disputing about how valid it is to use indirect treatment comparisons (indirect evidence) while making decisions. There are strong arguments against using such evidence especially when there are direct treatment comparisons (direct evidence) [11, 30–32]. A focus of criticism is the nature of the evidence provided by NMA. Although patients in a randomized clinical trial (RCT) are randomly assigned to each of the treatments that are compared, it cannot be argued that the treatments are randomized across the included trials.

Thus, indirect comparisons can be defined as non-randomized comparisons and correspondingly they provide observational evidence rather than randomized evidence. Consequently, indirect treatment comparisons may be more susceptible to biased treatment effect estimates, due, for example, to confounding (for example, when randomized AB and AC studies are systematically different from BC; [2] and selection bias (e.g., when the selection of comparator in the study is based on the relative treatment effect [33].

2.1 Indirect comparisons

Consider trial 1, a two-arm trial of the comparison “B–A”, and trial 2, a two-arm trial of the comparison “C–B”. If the estimated effect sizes in these trials are $\hat{\delta}_1^{AB}$ in trial 1 and $\hat{\delta}_2^{BC}$ in trial 2, then an indirect comparison of “C–A” may be obtained as $\hat{\delta}_{indirect}^{AC} = \hat{\delta}_1^{AB} + \hat{\delta}_2^{BC}$.

Through indirect comparison, the benefits of randomization can be maintained in each trial, and differences across the trials are allowed (e.g., in baseline risk) if only the prognosis of the participants but not their response to treatment is affected by these differences (in whichever metric is chosen as a measure of effect size). However, the indirect comparison is based on the assumption that the treatment named as B is the same in both trials so that its effects are nullified when “B–A” and “C–B” are added together. It is not possible to test whether the difference between A and C is truly reflected by an indirect comparison without having further information. The comparison of the indirect comparison with a direct comparison would be allowed by a third trial of “C–A” (yielding result $\hat{\delta}_3^{AC}$). The network of these three trials can be said to be consistent only if the underlying treatment effects are related to each other as follows:

$$\hat{\delta}_3^{AC} = \hat{\delta}_1^{AB} + \hat{\delta}_2^{BC} \tag{1}$$

Here δ_1^{AB} , δ_2^{BC} and δ_3^{AC} represent the actual effects that underlay the three studies. In practice, it is not very likely for Eq. (1) to hold for a particular set of three trials such as the ones described earlier. The reason behind this may be discussed either in terms of heterogeneity (because, within each treatment comparison, each individual study may not fully represent all studies in this particular comparison) or in terms of inconsistency (because, across treatment comparisons, important differences in the types of studies contributing to the comparisons exist). We will give more detailed information on these two concepts in subsequent sections [34].

2.2 Heterogeneity

The existing research has widely investigated heterogeneity in meta-analysis, referring to the situation where multiple studies focused on the same research question have different underlying values regarding the effect measure that is being estimated. The way of understanding heterogeneity in the network meta-analysis scenario is to keep the treatment comparison constant while changing the study index. In particular, the existence of heterogeneity can be argued for comparison ‘B–A’ if $\hat{\delta}_i^{AB} \neq \hat{\delta}_j^{AB}$ for some pair of studies i and j . It has been claimed that heterogeneity is an inevitable part of a meta-analysis [35] indicating that it is not likely that two trials of the same pairwise comparison are to have equal underlying treatment effects. Hence, within the context of Eq. (1), it is unlikely that the equality holds since the particular instance of “C–A”, which is examined in trial 3, will probably not represent all instances of “C–A” comparisons (and this holds true for trials 1 and 2 for their respective treatment comparisons). A random-effects model is a common way of allowing for heterogeneity. This assumes that the main effects in multiple runs of the same comparison arise from a common distribution, usually a normal distribution; namely,

$$\delta_i^{JK} \sim N(\delta^{JK}, \tau_{JK}^2) \quad (2)$$

for pairwise comparison JK (taking values AB, AC, or BC in the running example) [34].

2.3 Consistency

Consistency is the statistical manifestation of transitivity [12]. An additional way of making implicit inferences about the plausibility of the transitivity assumption is to check the network for consistency. What is meant by consistency is the statistical agreement between observed direct and (possibly many) indirect sources of evidence. A simple network can only contain treatments A, B, and C.

A consistency equation is generally used to express the relationship that is desirable between direct and indirect sources of evidence for a single comparison

$$\delta^{AC} = \delta^{AB} + \delta^{BC} \quad (3)$$

where the mean effect size across all studies of comparison JK is represented by δ^{JK} . (Under a fixed-effect meta-analysis model where the absence of heterogeneity is assumed, δ^{JK} represents a fixed (common) treatment effect for comparison JK). We refer to evidence that satisfies the consistency equation as showing consistency. We show this in **Figure 4(a)** as a three (non-touching) solid-edge relationship triangle in a network with only two-arm trials. Each edge represents one or more

Figure 4.
Graphical representation of consistency, loop inconsistency and design inconsistency.

two-arm trials that compare two treatments identified at either end of the edge. Using the same line style (a solid line), we draw all three edges to describe the situation where there is no contradiction (inconsistency) between them, that is, Eq. (2) is valid [34].

2.4 Loop inconsistency

When studies focused on various treatment comparisons are highly different in such a way that their effect sizes are affected, the consistency Eq. (2) might not be valid; thus, the effect sizes are not “added up” around the loop in the figure. This is called loop inconsistency and is shown by drawing edges using different line styles (**Figure 4(b)**). Loop inconsistency may only result from when there are different comparisons made in at least three separate study groups (e.g., studies “B–A”, “C–A” and “C–B”). Equivalently, it can only occur when we have both indirect and direct estimates of effect size (e.g., when “C–B” is measured both directly and through “A” indirectly) [34]. Some examples showing the causes of loop consistency are given below:

2.5 Multi-arm trials

Generally, some studies having more than two treatment arms are included in a network meta-analysis. In fact, about a quarter of randomized trials involve more

than two arms [36], so it is important to select appropriate methods while dealing with the condition.

When there is the presence of multi-arm trials in an evidence network, the definition of loop inconsistency becomes more complicated. It is not possible to loop inconsistency in a multi-arm trial. As a result, consistency can occur for a network either structurally (because all studies include all treatments) or through observation (when assumptions about equality of direct and various indirect comparisons hold across studies), or by means of a combination of the two.

Also, loop inconsistency cannot be properly defined using Eq. (2) anymore, since average effect sizes, δ^{JK} , refer to pairwise comparisons made from a combination of possibly inconsistent loops (e.g., from the two-arm trial) and naturally consistent loops (i.e., from multi-arm trials). In our drawings, multi-arm trials are shown using a closed (merged) polygon (**Figure 4(d)**) [34].

2.6 Transitivity

The purpose of an NMA is to improve the decision-making process for making choices between alternative treatments for a specific health condition and a target population. Hence, the estimates intended to be estimated in an NMA are the mean relative treatment effect sizes among the treatments competing with each other as they are expected to be present in the target population. If unbiased estimates are yielded by studies involved in the dataset and if a representative sample of the population addressed is constituted by these studies, then estimates generated by an NMA model for these parameters will be unbiased and consistent. The same set of assumptions is adopted by NMA as a pairwise meta-analysis [37], but there is also another assumption adopted by it which can be difficult to assess [38] and is called transitivity [39], (also called similarity [40, 41], or exchangeability [42]). Transitivity means that information for comparison between treatments A and B can be attained through another treatment C using comparisons A to C and B to C. It is not possible to test this assumption statistically, but it is possible to evaluate its validity in a conceptual and epidemiological way [21].

What is meant by the transitivity assumption is that direct evidence from studies AC and BC can be combined to gain insights (indirectly) about AB comparison. However, this will be open to questioning if there are significant differences in the distribution of effect modifiers (variables or characteristics that alter the observed relative effects, e.g., the mean age of participants and treatment dose) across the AC and BC trials, which yield insights about the indirect comparison [24, 39]. An effect modifier might have different effects across studies of the same comparison (e.g., the mean age of participants may differ across AC trials), but if its distribution across comparisons (AC and BC) is similar, the assumption of transitivity may still hold [21]. As a consequence, how plausible the transitivity assumption is can be assessed by reviewing the collection of studies for significant differences in the distribution of effect modifiers. Assuming that the studies are similar, the assumption of transitivity may be realistic, on the condition that there aren't any unknown modifiers of the relative treatment effect [43]. It is clear that such an assessment of transitivity may not be possible when the effect modifiers are not reported or when the number of studies per treatment comparison is low [12]. If there are significant differences identified and sufficient data is available, the transitivity of the network can be enhanced by using a network meta-regression. This might indicate, for example, that it is necessary for the common comparator treatment C to be similar in the AC and BC studies in terms of dose, modes of administration, duration, etc.

In an NMA of studies conducted to compare fluoride treatments administered to prevent dental caries, the definition of placebo differed between fluoride toothpaste

studies and fluoride rinse studies [44], casting doubt on how plausible the transitivity assumption is and thus challenging the reliability of the NMA results. In another example, Julious and Wang [45] focused on how the use of placebo as an intermediate comparator can result in the distortion of the results of indirect comparisons due to changes in the population's placebo response over the years; for instance, there might be a bias in the indirect estimate for A versus B when studies that compare treatment A versus placebo are older than studies that compare B versus placebo. Other ways used to formulate the transitivity assumption is to suppose that the true relative effect of A versus B is the same in the fixed-effects model or may vary across studies in the random-effects model, regardless of the treatments compared in each study [42, 46], that "missing" treatments in each trial are randomly missing [5] or, equivalently, that the choice of treatment comparisons in trials is not related directly or indirectly to the relative efficacy of the interventions. Finally, arguing that the patients included can be randomly distributed to any of the treatments in the network is an alternative way of postulating this assumption [21].

However, this does not mean that the assumption of transitivity will necessarily be valid. It should be stated that the absence of statistical inconsistency does not offer any evidence to prove the validity of the transitivity assumption that is essentially an assumption that cannot be tested as discussed in the previous section. Therefore, the conduct of an NMA should be preceded by a conceptual/theoretical evaluation of the transitivity assumption besides statistical tests for inconsistency [12] and the studies that are included in an NMA should always be reviewed for important differences that can be seen in patients, interventions, outcomes, study design, methodological characteristics, and reporting biases [2, 9, 14, 32, 43].

2.7 Design inconsistency

What is meant with the "design" of a study is a set of treatments that are compared within the study, recognizing that it is different from traditional interpretations made for the term. Then, differences in effect sizes among studies including different sets of treatments are referred to by design inconsistency. While allowing for this variation, it is implicitly assumed that different designs (i.e., different treatment sets included) can serve the function of a proxy for one or more important modifiers of effect [47]. Design inconsistency is depicted in **Figure 4(e)**, in which different line styles represent possible contradictions between study designs. The AC effect size depicted with a solid line in the three-arm trial is different from the AC effect size in the two-arm trial depicted with a dashed line. It is possible to see design inconsistency as a special case of heterogeneity since study designs correspond to a study-level covariate that has the potential to change effect sizes in the study, as can occur in a standard meta-regression analysis. It should be noted that in a network of only two-arm studies, additional insights provided by loop inconsistency cannot be provided by the concept of design inconsistency. In the case of a multi-arm trial, loop inconsistency in two-arm trials means design inconsistency (**Figure 4(f)**). The reason for this is that the multi-arm trial must be self-consistent, so the effect sizes of the multi-arm trial should be different from those of at least one of the two-arm trials: our definition of design inconsistency. Nevertheless, what is implied by design inconsistency for loop inconsistency is less clear. Design consistency with one three-arm trial and two two-arm trials is shown in **Figure 4(g)**. It is possible to create a loop by subtracting the pairwise BC comparison from the three-arm trial and then by comparing it to the two-arm trials. But, in this way, the existence of a consistent loop in the three-arm experiment is overlooked and thus it is unclear whether this network should be defined as exhibiting loop inconsistency. Also, it is seen in **Figure 4(h)** that the two-arm trials are consistent among themselves, but the effect sizes are different

from the effect sizes of the multi-arm trial. Does this show design inconsistency without loop inconsistency? [34].

2.8 Similarity

In order to make a comparison among the clinical trial studies used for analysis, it must be assumed that there is a similarity in the methodology used in the studies [12, 44]. The assessment of similarity is qualitatively performed on each of the selected articles from a methodological point of view and is not a hypothesis that can be tested statistically. The technique used to investigate similarity is the population, intervention, comparison, and outcome (PICO) technique [17]. Examination of similarity among the studies used for analysis is based on the following four items: clinical characteristics of study subjects, treatment interventions, comparison treatments, and outcome measures. In cases where the similarity assumption is not satisfied, the other two assumptions are also negatively affected [24] and moreover, there is also a need to check for the heterogeneity error [18, 21].

2.8.1 Network diagrams

One way of graphically depicting the structure of a network of interventions is a network diagram [12]. Such a graph is comprised of nodes that represent the interventions in the network and lines that show the available direct comparisons between pairs of interventions. An example of a network diagram including four interventions is given in **Figure 3**. In this example, in order to show the presence of a three-arm study, distinct lines that form a closed triangular loop have been added. It should be noted that complex and useless network diagrams may be yielded by such presentation of multi-arm studies; in this case, a tabular format can be preferred to depict multi-arm studies (**Figure 5**).

Figure 5.
Example of network diagram with four competing interventions and information on the presence of multi-arm randomized trials.

3. Illustrating example

The estimation of the relative effects on HbA1c change, of adding different oral glucose-lowering agents to a baseline sulfonylurea therapy in patients with type 2 diabetes, was the aim of the network meta-analysis in Diabetes. Systemic literature research was carried out on all relevant articles that were published from January 1993 to June 2009 in Medline and Embase. The search strategy was restricted to “randomized controlled 170 Statistical Methods in Medical Research 22(2) trials”, “sulfonylurea or sulphonylurea” and “humans”. This initial search was confirmed by combining each of the Medical Subject Headings key words “chlopropamide”, “glibenclamide”, “glyburide”, “gliclazide”, “glimepiride”, “glipizide”, “gliquidone”, “tolbutamide” on the one hand and ‘RCT’ on the other hand. No language restriction was applied. R program was used to analyze the data (Figure 6).

An original dataset offered by Senn [48] will be used in our first network meta-analysis. In this dataset, there are effect size data obtained from randomized controlled trials that compare different medications for diabetes. The effect size obtained for all comparisons represents the mean difference (MD) of diabetic patients’ HbA1c value in the posttest. What is represented by this value is the concentration of glucose found in the blood, which is aimed to be decreased with diabetic medication. As can be seen, there are 28 rows that represent the treatment comparisons and seven columns in the data. In the first column, TE, there is the effect size of each comparison, and the respective standard error is contained in se TE. In case effect size data that have already been calculated for each comparison might not be possessed.

	TE	seTE	treat1.long	treat2.long	treat1	treat2	studlab
1	-1.90	0.1414	Metformin	Placebo	metf	plac	DeFronzo1995
2	-0.82	0.0992	Metformin	Placebo	metf	plac	Lewin2007
3	-0.20	0.3579	Metformin	Acarbose	metf	acar	Willms1999
4	-1.34	0.1435	Rosiglitazone	Placebo	rosi	plac	Davidson2007
5	-1.10	0.1141	Rosiglitazone	Placebo	rosi	plac	Wolffenbuttel1999
6	-1.30	0.1268	Pioglitazone	Placebo	piog	plac	Kipnes2001
7	-0.77	0.1078	Rosiglitazone	Placebo	rosi	plac	Kerenyi2004
8	0.16	0.0849	Pioglitazone	Metformin	piog	metf	Hanefeld2004
9	0.10	0.1831	Pioglitazone	Rosiglitazone	piog	rosi	Derosa2004
10	-1.30	0.1014	Rosiglitazone	Placebo	rosi	plac	Baksi2004
11	-1.09	0.2263	Rosiglitazone	Placebo	rosi	plac	Rosenstock2008
12	-1.50	0.1624	Rosiglitazone	Placebo	rosi	plac	Zhu2003
13	-0.14	0.2239	Rosiglitazone	Metformin	rosi	metf	Yang2003
14	-1.20	0.1436	Rosiglitazone	Sulfonylurea	rosi	sulf	Vongthavaravat2002
15	-0.40	0.1549	Acarbose	Sulfonylurea	acar	sulf	Oyama2008
16	-0.80	0.1432	Acarbose	Placebo	acar	plac	Costa1997
17	-0.57	0.1291	Sitagliptin	Placebo	sita	plac	Hermansen2007
18	-0.70	0.1273	Vildagliptin	Placebo	vild	plac	Garber2008
19	-0.37	0.1184	Metformin	Sulfonylurea	metf	sulf	Alex1998
20	-0.74	0.1839	Miglitol	Placebo	migl	plac	Johnston1994
21	-1.41	0.2235	Miglitol	Placebo	migl	plac	Johnston1998a
22	0.00	0.2339	Rosiglitazone	Metformin	rosi	metf	Kim2007
23	-0.68	0.2828	Miglitol	Placebo	migl	plac	Johnston1998b
24	-0.40	0.4356	Metformin	Placebo	metf	plac	Gonzalez-Ortiz2004
25	-0.23	0.3467	Benfluorex	Placebo	benf	plac	Stuccil1996
26	-1.01	0.1366	Benfluorex	Placebo	benf	plac	Moulin2006
27	-1.20	0.3758	Metformin	Placebo	metf	plac	Willms1999
28	-1.00	0.4669	Acarbose	Placebo	acar	plac	Willms1999

Figure 6.
Diabetes example and view the data.

The two treatments that are compared are represented by treat1. long, treat2. long, treat1, and treat2. As a shortened name of the original treatment name is contained in the variables treat1 and treat2, they are redundant.

We can now move forward by fitting our initial network meta-analysis model using the net metafunction. Now, we can look at the results of our first model, for now assuming a fixed-effects model.

Results (fixed effects model):

	treat1	treat2	MD	95%-CI	Q	leverage
DeFronzo1995	metf	plac	-1.1141	[-1.2309; -0.9973]	30.89	0.18
Lewin2007	metf	plac	-1.1141	[-1.2309; -0.9973]	8.79	0.36
Willms1999	acar	metf	0.2867	[0.0622; 0.5113]	0.05	.
Davidson2007	plac	rosi	1.2018	[1.1084; 1.2953]	0.93	0.11
Wolffenbuttel1999	plac	rosi	1.2018	[1.1084; 1.2953]	0.80	0.17
Kipnes2001	piog	plac	-1.0664	[-1.2151; -0.9178]	3.39	0.36
Kerenyi2004	plac	rosi	1.2018	[1.1084; 1.2953]	16.05	0.20
Hanefeld2004	metf	piog	-0.0477	[-0.1845; 0.0891]	1.75	0.68
Derosa2004	piog	rosi	0.1354	[-0.0249; 0.2957]	0.04	0.20
Baksi2004	plac	rosi	1.2018	[1.1084; 1.2953]	0.94	0.22
Rosenstock2008	plac	rosi	1.2018	[1.1084; 1.2953]	0.24	0.04
Zhu2003	plac	rosi	1.2018	[1.1084; 1.2953]	3.37	0.09
Yang2003	metf	rosi	0.0877	[-0.0449; 0.2203]	0.05	0.09
Vongthavaravat2002	rosi	sulf	-0.7623	[-0.9427; -0.5820]	9.29	0.41
Oyama2008	acar	sulf	-0.3879	[-0.6095; -0.1662]	0.01	0.53
Costa1997	acar	plac	-0.8274	[-1.0401; -0.6147]	0.04	0.57
Hermansen2007	plac	sita	0.5700	[0.3170; 0.8230]	0.00	1.00
Garber2008	plac	vild	0.7000	[0.4505; 0.9495]	0.00	1.00
Alex1998	metf	sulf	-0.6746	[-0.8482; -0.5011]	6.62	0.56
Johnston1994	migl	plac	-0.9439	[-1.1927; -0.6952]	1.23	0.48
Johnston1998a	migl	plac	-0.9439	[-1.1927; -0.6952]	4.35	0.32
Kim2007	metf	rosi	0.0877	[-0.0449; 0.2203]	0.14	0.08
Johnston1998b	migl	plac	-0.9439	[-1.1927; -0.6952]	0.87	0.20
Gonzalez-Ortiz2004	metf	plac	-1.1141	[-1.2309; -0.9973]	2.69	0.02
Stucci1996	benf	plac	-0.9052	[-1.1543; -0.6561]	3.79	0.13
Moulin2006	benf	plac	-0.9052	[-1.1543; -0.6561]	0.59	0.87
Lewin2007	metf	plac	-1.1141	[-1.2309; -0.9973]	8.79	0.36
Willms1999	acar	metf	0.2867	[0.0622; 0.5113]	0.05	.
Davidson2007	plac	rosi	1.2018	[1.1084; 1.2953]	0.93	0.11
Wolffenbuttel1999	plac	rosi	1.2018	[1.1084; 1.2953]	0.80	0.17
Kipnes2001	piog	plac	-1.0664	[-1.2151; -0.9178]	3.39	0.36
Kerenyi2004	plac	rosi	1.2018	[1.1084; 1.2953]	16.05	0.20
Hanefeld2004	metf	piog	-0.0477	[-0.1845; 0.0891]	1.75	0.68
Derosa2004	piog	rosi	0.1354	[-0.0249; 0.2957]	0.04	0.20
Baksi2004	plac	rosi	1.2018	[1.1084; 1.2953]	0.94	0.22
Rosenstock2008	plac	rosi	1.2018	[1.1084; 1.2953]	0.24	0.04
Zhu2003	plac	rosi	1.2018	[1.1084; 1.2953]	3.37	0.09
Yang2003	metf	rosi	0.0877	[-0.0449; 0.2203]	0.05	0.09
Vongthavaravat2002	rosi	sulf	-0.7623	[-0.9427; -0.5820]	9.29	0.41
Oyama2008	acar	sulf	-0.3879	[-0.6095; -0.1662]	0.01	0.53
Costa1997	acar	plac	-0.8274	[-1.0401; -0.6147]	0.04	0.57
Hermansen2007	plac	sita	0.5700	[0.3170; 0.8230]	0.00	1.00
Garber2008	plac	vild	0.7000	[0.4505; 0.9495]	0.00	1.00
Alex1998	metf	sulf	-0.6746	[-0.8482; -0.5011]	6.62	0.56
Johnston1994	migl	plac	-0.9439	[-1.1927; -0.6952]	1.23	0.48
Johnston1998a	migl	plac	-0.9439	[-1.1927; -0.6952]	4.35	0.32
Kim2007	metf	rosi	0.0877	[-0.0449; 0.2203]	0.14	0.08
Johnston1998b	migl	plac	-0.9439	[-1.1927; -0.6952]	0.87	0.20
Gonzalez-Ortiz2004	metf	plac	-1.1141	[-1.2309; -0.9973]	2.69	0.02
Stucci1996	benf	plac	-0.9052	[-1.1543; -0.6561]	3.79	0.13
Moulin2006	benf	plac	-0.9052	[-1.1543; -0.6561]	0.59	0.87
Willms1999	metf	plac	-1.1141	[-1.2309; -0.9973]	0.04	.
Willms1999	acar	plac	-0.8274	[-1.0401; -0.6147]	0.04	.

Number of studies: k = 26
 Number of treatments: n = 10
 Number of pairwise comparisons: m = 28
 Number of designs: d = 15


```
Treatment estimate (sm = 'MD', comparison: other treatments vs 'plac'):  
      MD      95%-CI      z p-value  
acar -0.8274 [-1.0401; -0.6147] -7.62 < 0.0001  
benf -0.9052 [-1.1543; -0.6561] -7.12 < 0.0001  
metf -1.1141 [-1.2309; -0.9973] -18.69 < 0.0001  
mig1 -0.9439 [-1.1927; -0.6952] -7.44 < 0.0001  
piog -1.0664 [-1.2151; -0.9178] -14.06 < 0.0001  
plac . . . . .  
rosi -1.2018 [-1.2953; -1.1084] -25.22 < 0.0001  
sita -0.5700 [-0.8230; -0.3170] -4.42 < 0.0001  
sulf -0.4395 [-0.6188; -0.2602] -4.80 < 0.0001  
vild -0.7000 [-0.9495; -0.4505] -5.50 < 0.0001  
  
Quantifying heterogeneity / inconsistency:  
tau^2 = 0.1087; tau = 0.3297; I^2 = 81.4% [72.0%; 87.7%]  
  
Tests of heterogeneity (within designs) and inconsistency (between designs):  
      Q d.f. p-value  
Total      96.99   18 < 0.0001  
Within designs  74.46   11 < 0.0001  
Between designs 22.53    7  0.0021
```

As we have created our network meta-analysis model, we can go ahead and draw our network graph (Figure 2). Several types of information are conveyed by this network graph.

- First, there is the overall structure of comparisons in our network, which makes it possible for us to understand which treatments were compared with each other in the original data.
- Second, there are the edges having a different thickness, indicating how often this specific comparison can be found in our network. We see that there are many trials comparing Rosiglitazone with Placebo.

There is also one multiarm trial in our network, represented by the triangle shown in blue in our network. As a next step, our attention can be shifted towards the direct and indirect evidence in our network by looking at the rate of direct and indirect contribution to each comparison. A function has been prepared to this end with the name of `direct.evidence.plot`. As can be seen in Figure 7, there are many estimates included in our network model that needed to be inferred by indirect evidence only. We are also provided with two additional metrics by the plot: The Minimal Parallelism and the Mean Path Length

Figure 7.
Direct evidence proportion for each network estimate.

Length of each comparison. It is noted by König [49] that lower values of minimal parallelism and Mean Path Length >2 means that care should be taken while interpreting results for specific comparison.

Then we can look at our network’s estimates for all possible combinations of treatments. In order to be able to do this, result matrices stored in our net meta results object under the fixed effects model can be used. Through a few preprocessing steps, the matrix can be made easier to read. First, the matrix is extracted from our data and the numbers in the matrix are rounded to three digits.

	acar	benf	metf	migl	piog	plac	rosi	sita	sulf	vild
acar	0.000	0.078	0.287	0.117	0.239	-0.827	0.374	-0.257	-0.388	-0.127
benf	-0.078	0.000	0.209	0.039	0.161	-0.905	0.297	-0.335	-0.466	-0.205
metf	-0.287	-0.209	0.000	-0.170	-0.048	-1.114	0.088	-0.544	-0.675	-0.414
migl	-0.117	-0.039	0.170	0.000	0.123	-0.944	0.258	-0.374	-0.504	-0.244
piog	-0.239	-0.161	0.048	-0.123	0.000	-1.066	0.135	-0.496	-0.627	-0.366
plac	0.827	0.905	1.114	0.944	1.066	0.000	1.202	0.570	0.439	0.700
rosi	-0.374	-0.297	-0.088	-0.258	-0.135	-1.202	0.000	-0.632	-0.762	-0.502
sita	0.257	0.335	0.544	0.374	0.496	-0.570	0.632	0.000	-0.131	0.130
sulf	0.388	0.466	0.675	0.504	0.627	-0.439	0.762	0.131	0.000	0.261
vild	0.127	0.205	0.414	0.244	0.366	-0.700	0.502	-0.130	-0.261	0.000

When the fact that a “triangle” in our matrix includes too much redundant information is considered, it seems to be possible to replace the lower triangle with an empty value.

	acar	benf	metf	migl	piog	plac	rosi	sita	sulf	vild
acar	"0"	"0.078"	"0.287"	"0.117"	"0.239"	"-0.827"	"0.374"	"-0.257"	"-0.388"	"-0.127"
benf	"."	"0"	"0.209"	"0.039"	"0.161"	"-0.905"	"0.297"	"-0.335"	"-0.466"	"-0.205"
metf	"."	"."	"0"	"-0.17"	"-0.048"	"-1.114"	"0.088"	"-0.544"	"-0.675"	"-0.414"
migl	"."	"."	"."	"0"	"0.123"	"-0.944"	"0.258"	"-0.374"	"-0.504"	"-0.244"
piog	"."	"."	"."	"."	"0"	"-1.066"	"0.135"	"-0.496"	"-0.627"	"-0.366"
plac	"."	"."	"."	"."	"."	"0"	"1.202"	"0.57"	"0.439"	"0.7"
rosi	"."	"."	"."	"."	"."	"."	"0"	"-0.632"	"-0.762"	"-0.502"
sita	"."	"."	"."	"."	"."	"."	"."	"0"	"-0.131"	"0.13"
sulf	"."	"."	"."	"."	"."	"."	"."	"."	"0"	"0.261"
vild	"."	"."	"."	"."	"."	"."	"."	"."	"."	"0"

The net league() function offers an extremely convenient way of exporting all estimated effect sizes. A matrix similar to the one given above can be generated by this function. Yet, in the matrix created by this function, only the pooled effect sizes belonging to the direct comparisons available in our network will be shown by the upper triangle, like the ones to be attained if a conventional meta-analysis had been conducted for each comparison. As there is no direct evidence for all comparisons, we will see some fields in the upper triangle empty. In this case, the network meta-analysis effect sizes for each comparison are contained by the lower triangle. The biggest advantage of this function is that it allows effect size estimates and confidence intervals to be shown together in each cell; the only thing that we need to tell the function is how the brackets for the confidence intervals should look like and how many digits we want our estimates to have behind the comma.

In a network meta-analysis, the most interesting question desired to be answered is: which intervention works the best? Such an ordering of treatments from most to least useful can be performed by the net rank() function implemented in net meta. The net rank() function is also built on a method of frequentist treatment ranking that uses P-scores. With these P-scores, the certainty that one treatment is better than another treatment is measured. It has been shown that this P-score is equivalent to the SUCRA score [50]. Our net meta object is needed as input by the function. Moreover, the small values parameter used to define whether smaller effect sizes in comparison are an indicator of a beneficial (“good”) or harmful (“bad”) effect should be specified. Now we will look at the output for our example:

	P-score
rosi	0.9789
metf	0.8513
piog	0.7686
migl	0.6200
benf	0.5727
acar	0.4792
vild	0.3512
sita	0.2386
sulf	0.1395
plac	0.0000

As can be seen, the Rosiglitazone treatment has the highest P-score, which indicates that this treatment may be particularly helpful. Contrarily, the P-score of Placebo is zero, supporting our intuition that placebo may not be the best treatment decision. It should be noted, however, that treatment should never be automatically concluded to be the best just because it has the highest score [51]. One of the good ways to be used to visualize the uncertainty in our network is to generate network forest plots with the “weakest” treatment as a comparison. The forest plot can also be used to do this. The reference group for the forest plot can be specified by using the reference group argument (**Figures 8 and 9**).

Now it can be seen that the results are more ambiguous than they seemed before; it is seen that several high-performing treatments having overlapping confidence intervals are available. This means we cannot make a firm judgment about which treatment is actually the best, but rather we see that there are a number of treatments that are more effective compared to placebo.

3.1 Decomposition of heterogeneity statistics

It is possible to decompose the Q total statistic (of the “whole network”) into a Q statistic to assess heterogeneity between studies having the same design (“within designs”) and a Q statistic to assess design inconsistency (“between designs”). The subsets of treatments that are compared with each other in a study are used to define designs.

Figure 8.
Forest plot for fixed effect model with placebo as reference.

Figure 9.
Forest plot for random-effects model with placebo as reference.

	Q	df	pval
Total	96.98555	18	7.871369e-13
Within designs	74.45528	11	1.723944e-11
Between designs	22.53027	7	2.057014e-03

For this analysis, the fixed-effect model has been used and it is seen that there is considerable heterogeneity/inconsistency within as well as between designs. The total within-design heterogeneity can be further decomposed into the contribution from each design.

	design	Q	df	pval
1	acar vs sulf	0.000	0	NA
2	metf vs piog	0.000	0	NA
3	metf vs rosi	0.187	1	6.65e-01
4	metf vs sulf	0.000	0	NA
5	piog vs rosi	0.000	0	NA
6	plac vs acar	0.000	0	NA
7	plac vs benf	4.381	1	3.63e-02
8	plac vs metf	42.164	2	6.98e-10
9	plac vs migl	6.449	2	3.98e-02
10	plac vs piog	0.000	0	NA
11	plac vs rosi	21.273	5	7.19e-04
12	plac vs sita	0.000	0	NA
13	plac vs vild	0.000	0	NA
14	rosi vs sulf	0.000	0	NA
15	plac vs acar vs metf	0.000	0	NA

As can be seen, the network meta-analysis includes 26 studies and these 26 studies use 15 different designs. Because only five designs for which more than one study exist, the remaining Q statistics specific to design are equal to zero and do not have any degrees of freedom. Except for design metf:rosi (p value = 0.67), heterogeneity is higher than would be expected between the contributing studies for all the other four designs; in the case of metf:plac a substantial amount more ($p < 0.0001$). Sources of this could be identified in a substantive application and thus the analysis could be updated appropriately.

	Q	df	pval	tau.within
Between designs	2.194	7	0.948	0.38

Now the net heat plot, put forward by Krahn, König, and Binder [49] will be introduced. This is a graphical presentation where two types of information are shown in a single plot. These are:

1. For each network estimate, the contribution of each design to this estimate, and
2. For each network estimate, the extent of inconsistency due to each design.

Net heat plot is very useful in terms of evaluating the inconsistency in our network model, and what contributes to it (**Figure 10**).

A quadratic matrix is produced by the function so that each element in a row can be compared to all other elements in the columns. It should be noted here that rows and columns do not refer to all treatment comparisons in our network rather to specific designs. Thus, we also have rows and columns for the multiarm study, which had a design that compares “Plac”, “Metf” and Acar. Comparison of treatments with only one type of evidence (i.e., indirect or indirect evidence) is not included in this chart, as we are dealing with cases of inconsistency between direct and indirect evidence. Moreover, the net heat plot has also two important properties:

1. Gray boxes. The Gray boxes for each design comparison show the extent to which one treatment comparison is important in terms of estimating another treatment comparison. The increasing size of the box indicates the increasing importance of comparison. This can be easily analyzed by going through the rows of the plot one after another, and then by checking for each row in which columns the gray boxes are the largest. In rows where the row comparison and the column comparison intersect, the boxes are large, which is a common finding and means that direct evidence was employed. For instance, it is possible to see a big gray box at the point where the “Plac vs Rosi2” row and the “Plac vc Rosi” column intersect [52].

The colored backgrounds which range from blue to red indicate the inconsistency of the comparison in a row, which can be attributed to the design in a column. Inconsistent fields are shown in the upper-left corner in red. For instance, it is seen that the entry in column “Metf vs. Sulf” is shown with red in the row for “Rosi vs. Sulf”. This indicates that the evidence that “Metf vs. Sulf” provides for the “Metf vs. Sulf” estimation is not consistent with the other evidence. We can now remember

Figure 10.
Net heat plot of the Senn data example based on a fixed-effect model.

Figure 11.
Net heat plot of the Senn data example from a random-effects model.

that the fixed effects model that we initially used for our network analysis forms the basis of these results. On the basis of the things we have seen so far, we can reach the conclusion that due to too much unexpected heterogeneity, justification is not provided for the fixed effects model. How the net heat graph changes when a random-effects model is assumed can be controlled by changing the random argument of the net heat function to TRUE. It is seen that this results in a significant reduction of inconsistency in our network (**Figure 11**).

3.1.1 Net splitting

Net splitting, also known as node splitting, is another method for checking consistency in our network. With this method, our network estimates are split into the contribution of direct and indirect evidence and in this way, we can control for inconsistency in specific comparisons in our network. To generate a net split and compare the results.

```

Fixed effects model:

  comparison k prop      nma    direct   indir.    Diff      z    p-value
acar vs benf 0   0  0.0778      .    0.0778      .      .      .
acar vs metf 1 0.10 0.2867  0.2000  0.2966 -0.0966 -0.26  0.7981
acar vs migl 0   0  0.1166      .    0.1166      .      .      .
acar vs piog 0   0  0.2391      .    0.2391      .      .      .
acar vs plac 2 0.63 -0.8274 -0.8172 -0.8446  0.0274  0.12  0.9030
acar vs rosi 0   0  0.3745      .    0.3745      .      .      .
acar vs sita 0   0 -0.2574      .   -0.2574      .      .      .
acar vs sulf 1 0.53 -0.3879 -0.4000 -0.3740 -0.0260 -0.11  0.9088
acar vs vild 0   0 -0.1274      .   -0.1274      .      .      .
benf vs metf 0   0  0.2089      .    0.2089      .      .      .
benf vs migl 0   0  0.0387      .    0.0387      .      .      .
benf vs piog 0   0  0.1612      .    0.1612      .      .      .
benf vs plac 2 1.00 -0.9052 -0.9052      .      .      .
benf vs rosi 0   0  0.2967      .    0.2967      .      .      .
benf vs sita 0   0 -0.3352      .   -0.3352      .      .      .
benf vs sulf 0   0 -0.4657      .   -0.4657      .      .      .
benf vs vild 0   0 -0.2052      .   -0.2052      .      .      .
metf vs migl 0   0 -0.1702      .   -0.1702      .      .      .
metf vs piog 1 0.68 -0.0477 -0.1600  0.1866 -0.3466 -2.32  0.0201
metf vs plac 4 0.58 -1.1141 -1.1523 -1.0608 -0.0915 -0.76  0.4489
metf vs rosi 2 0.18  0.0877  0.0731  0.0908 -0.0178 -0.10  0.9204

```

metf vs sita	0	0	-0.5441	.	-0.5441	.	.	.
metf vs sulf	1	0.56	-0.6746	-0.3700	-1.0611	0.6911	3.88	0.0001
metf vs vild	0	0	-0.4141	.	-0.4141	.	.	.
migl vs piog	0	0	0.1225	.	0.1225	.	.	.
migl vs plac	3	1.00	-0.9439	-0.9439	.	.	.	.
migl vs rosi	0	0	0.2579	.	0.2579	.	.	.
migl vs sita	0	0	-0.3739	.	-0.3739	.	.	.
migl vs sulf	0	0	-0.5044	.	-0.5044	.	.	.
migl vs vild	0	0	-0.2439	.	-0.2439	.	.	.
piog vs plac	1	0.36	-1.0664	-1.3000	-0.9363	-0.3637	-2.30	0.0215
piog vs rosi	1	0.20	0.1354	0.1000	0.1442	-0.0442	-0.22	0.8289
piog vs sita	0	0	-0.4964	.	-0.4964	.	.	.
piog vs sulf	0	0	-0.6269	.	-0.6269	.	.	.
piog vs vild	0	0	-0.3664	.	-0.3664	.	.	.
rosi vs plac	6	0.83	-1.2018	-1.1483	-1.4665	0.3182	2.50	0.0125
sita vs plac	1	1.00	-0.5700	-0.5700	.	.	.	.
sulf vs plac	0	0	-0.4395	.	-0.4395	.	.	.
vild vs plac	1	1.00	-0.7000	-0.7000	.	.	.	.
rosi vs sita	0	0	-0.6318	.	-0.6318	.	.	.
rosi vs sulf	1	0.41	-0.7623	-1.2000	-0.4575	-0.7425	-3.97	< 0.0001
rosi vs vild	0	0	-0.5018	.	-0.5018	.	.	.
sita vs sulf	0	0	-0.1305	.	-0.1305	.	.	.
sita vs vild	0	0	0.1300	.	0.1300	.	.	.
sulf vs vild	0	0	0.2605	.	0.2605	.	.	.

Legend:

comparison	-	Treatment comparison
k	-	Number of studies providing direct evidence
prop	-	Direct evidence proportion
nma	-	Estimated treatment effect (MD) in network meta-analysis
direct	-	Estimated treatment effect (MD) derived from direct evidence
indir.	-	Estimated treatment effect (MD) derived from indirect evidence
Diff	-	Difference between direct and indirect treatment estimates
z	-	z-value of test for disagreement (direct versus indirect)

Here, the important information is found in the p-value column. Any value that is $p < 0.05$ in this column is an indicator of a significant discrepancy (inconsistency) between the direct and indirect estimates. In the output, it is seen that there are indeed few comparisons showing significant discrepancies between direct and indirect evidence when the fixed effects model is used. Net split results can be visualized with a forest chart showing all comparisons for which both direct and indirect evidence are present in **Figure 12**.

4. Conclusions

For the estimation and comparison of treatment effects in a particular area, network meta-analysis can be used as a potentially powerful tool for using all the evidence. This approach has been depicted through an example from diabetes [48], which shows how to graph the network and explore a range of analyses. The results of our first model (fixed-effect model) Q value of DeFronzo1995 is highest with $Q = 30.89$. As a network model, the effects of all treatments are displayed in comparison to the placebo condition, which is why there is no effect shown for placebo. We can say heterogeneity/inconsistency in our network model is high, with $I^2 = 84\%$. The heterogeneity between treatment designs reflects the actual inconsistency in our network, and is highly significant ($p = 0.0021$). In **Figure 2**, looking at the network graph, it is seen that Rosiglitazone has been compared to Placebo in many trials. The only multi-arm trial in our network is that of Willms 2003. We see that it is the Rosiglitazone treatment with the highest P score. It is necessary to look at network forest plots with the “weakest” treatment, as it can be misleading to conclude that a treatment is best just because it has the highest score.

Figure 12.
Net split plot of the Senn data example from a fixed-effect model.

Looking at the forest network plot, we see that there are several high-performance treatments with overlapping confidence intervals. From here, we looked at the net heat plot as we could not make a definitive decision.

The extent of the information obtained in a given treatment comparison by means of indirect evidence and the extent of heterogeneity can be defined as two important aspects of network meta-analysis. The net heat graph communicates information about both of these and the software allows for the decomposition of heterogeneity within and between designs. If there is clinically relevant heterogeneity, it is worth being explored further. Looking at **Figure 10**, a particularly large

gray box is seen where the “Plac vs. Rosi2 row and the “Plac vs. Rosi” column intersect. Using the random-effects model in **Figure 11**, we see that the inconsistency is significantly reduced.

Since it is not possible to conduct covariate adjustment at present with the software, one approach is to conduct study-specific (ideally individual participant data) analyses with appropriate covariate adjustment before the software presented here is used to perform network meta-analysis.

IntechOpen

IntechOpen

Author details

Nilgün Yildiz
Marmara University, Istanbul, Turkey

*Address all correspondence to: ncelebi@marmara.edu.tr

IntechOpen

© 2022 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] Shim SR, Yoon BY, Shin IS, Bae JM. Network meta-analysis: Application and practice using Stata. Korean Society of Epidemiology. 2017;**39**:e2017047. DOI: 10.4178/epih.e2017047
- [2] Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: Combining direct and indirect evidence. *BMJ*. 2005;**331**:897-900
- [3] Li T, Vedula SS, Scherer R, Dickersin K. What comparative effectiveness research is needed? A framework for using guidelines and systematic reviews to identify evidence gaps and research priorities. *Annals of Internal Medicine*. 2012;**156**:367-377
- [4] Mitka M. US government kicks off program for comparative effectiveness research. *Journal of the American Medical Association*. 2010;**304**:2230-2231
- [5] Lu G, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*. 2006;**101**: 447-459
- [6] Salanti G, Higgins JP, Ades AE, Ioannidis JP. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research*. 2008;**17**:279-301
- [7] Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Statistics in Medicine*. 1996;**15**: 2733-2749
- [8] Mills EJ, Ioannidis JP, Thorlund K, Schünemann HJ, Puhan MA, Guyatt GH. How to use an article reporting a multiple treatment comparison metaanalysis. *Journal of the American Medical Association*. 2012; **308**:1246-1253
- [9] Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology*. 1997;**50**:683-691
- [10] Mills EJ, Ghement I, O'Regan C, Thorlund K. Estimating the power of indirect comparisons: A simulation study. *PLoS One*. 2011;**6**(1):e16237
- [11] Ioannidis JP. Indirect comparisons: The mesh and mess of clinical trials. *Lancet*. 2006;**368**:1470-1472
- [12] Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Annals of Internal Medicine*. 2013;**159**: 130-137
- [13] Tonin FS, Rotta I, Mendes AM, Pontarolo R. Network meta-analysis: A technique to gather evidence from direct and indirect comparisons. *Pharmacy Practice (Granada)*. 2017;**15**:943
- [14] Hoaglin DC, Hawkins N, Jansen JP, Scott DA, Itzler R, Cappelleri JC, et al. Conducting indirect-treatment-comparison and network meta-analysis studies: Report of the ISPOR Task Force on indirect treatment comparisons good research practices: Part 2. *Value in Health*. 2011;**14**:429-437
- [15] Li T, Puhan MA, Vedula SS, Singh S, Dickersin K, Ad Hoc Network Meta-analysis Methods Meeting Working Group. Network meta-analysis-highly attractive but more methodological research is needed. *BMC Medicine*. 2011;**9**:79
- [16] Mills EJ, Bansback N, Ghement I, Thorlund K, Kelly S, Puhan MA, et al. Multiple treatment comparison meta-analyses: A step forward into complexity. *Clinical Epidemiology*. 2011;**3**:193-202
- [17] Reken S, Sturtz S, Kiefer C, Böhler YB, Wieseler B. Assumptions of

mixed treatment comparisons in health technology assessments: Challenges and possible steps for practical application. *PLoS One*. 2016;**11**:e0160712

[18] Veroniki AA, Vasiliadis HS, Higgins JP, Salanti G. Evaluation of inconsistency in networks of interventions. *International Journal of Epidemiology*. 2013;**42**:332-345

[19] Bhatnagar N, Lakshmi PV, Jeyashree K. Multiple treatment and indirect treatment comparisons: An overview of network meta-analysis. *Perspectives in Clinical Research*. 2014; **5**:154-158

[20] Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ*. 2013;**346**:f2914

[21] 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. *Research Synthesis Methods*. 2012;**3**:80-97

[22] Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine*. 2004;**23**:3105-3124

[23] Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, et al. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: Report of the ISPOR Task Force on indirect treatment comparisons good research practices: Part 1. *Value in Health*. 2011; **14**:417-428

[24] 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 Medicine*. 2013;**11**:159

[25] Dakin HA, Welton NJ, Ades AE, Collins S, Orme M, Kelly S. Mixed

treatment comparison of repeated measurements of a continuous endpoint: An example using topical treatments for primary openangle glaucoma and ocular hypertension. *Statistics in Medicine*. 2011;**30**:2511-2535

[26] Schmitz S, Adams R, Walsh CD, Barry M, FitzGerald O. A mixed treatment comparison of the efficacy of anti-TNF agents in rheumatoid arthritis for methotrexate non-responders demonstrates differences between treatments: A Bayesian approach. *Annals of the Rheumatic Diseases*. 2012; **71**:225-230

[27] Jones B, Roger J, Lane PW, Lawton A, Fletcher C, Cappelleri JC, et al. Statistical approaches for conducting network meta-analysis in drug development. *Pharmaceutical Statistics*. 2011;**10**:523-531

[28] White IR. Network meta-analysis. *The Stata Journal*. 2015;**15**:951-985

[29] Cooper NJ, Peters J, Lai MC, et al. How valuable are multiple treatment comparison methods in evidence-based health-care evaluation? *Value in Health*. 2011;**14**(2):371-380

[30] Edwards SJ, Clarke MJ, Wordsworth S, Borriall J. Indirect comparisons of treatments based on systematic reviews of randomised controlled trials. *International Journal of Clinical Practice*. 2009;**63**:841-854. DOI: 10.1111/ j.1742-1241.2009.02072

[31] Gartlehner G, Moore CG. Direct versus indirect comparisons: A summary of the evidence. *The International Journal of Technology Assessment in Health Care*. 2008;**24**: 170-177. DOI: 10.1017/S0266462308080240

[32] Efthimiou O, Debray TPA, vanValkenhoef G, Trelle S, Panayidou K, Moons KGM, et al. GetReal in network meta-analysis: A

- p>review of the methodology.
- Research Synthesis Methods*
- . 2016;
- 7**
- :236-263. DOI: 10.1002/jrsm.1195
- [33] Salanti G, Kavvoura FK, Ioannidis JP. Exploring the geometry of treatment networks. *Annals of Internal Medicine*. 2008;**148**:544-553
- [34] Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. *Research Synthesis Methods*. 2012;**3**:98-110
- [35] Higgins JPT. Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International Journal of Epidemiology*. 2008;**37**:1158-1160
- [36] Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *The Lancet*. 2005;**365**:1159-1162
- [37] Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Statistics in Medicine*. 2010b;**29**:932-944. DOI: 10.1002/sim.3767
- [38] Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: Survey of published systematic reviews. *BMJ*. 2009;**338**:b1147
- [39] Baker SG, 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 Medical Research Methodology*. 2002;**2**:13
- [40] Donegan S, Williamson P, Gamble C, Tudur SC. Indirect comparisons: A review of reporting and methodological quality. *PLoS One*. 2010;**5**:e11054. DOI: 10.1371
- [41] Song F, Altman DG, Glenny AM, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*. 2003;**326**:472. DOI: 10.1136/bmj.326.7387.472
- [42] Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: Inconsistency in networks of evidence based on randomized controlled trials. *Medical Decision Making*. 2013d;**33**:641-656. DOI: 10.1177/0272989X12455847
- [43] Donegan S, Williamson P, D'Alessandro U, Tudur SC. Assessing key assumptions of network meta-analysis: A review of methods. *Research Synthesis Methods*. 2013b;**4**:291-323. DOI: 10.1002/jrsm.1085
- [44] Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *Journal of Clinical Epidemiology*. 2009;**62**:857-864. DOI: 10.1016/j.jclinepi.2008.10.001
- [45] Julious SA, Wang SJ. How biased are indirect comparisons, particularly when comparisons are made over time in controlled trials. *Drug Information Journal*. 2008;**42**:625
- [46] Lu G, Ades A. Modeling between-trial variance structure in mixed treatment comparisons. *Biostatistics*. 2009;**10**:792-805. DOI: 10.1093/biostatistics/kxp032
- [47] Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine*. 2002;**21**:2313-2324. DOI: 10.1002/sim.1201
- [48] Senn S, Gavini F, Magrez D, Scheen A. Issues in performing a network meta-analysis. *Statistical Methods in Medical Research*. 2013;**22**(2):169-189. DOI: 10.1177/0962280211432220

[49] König J, Krahn U, Binder H.
Visualizing the flow of evidence in
network meta-analysis and
characterizing mixed treatment
comparisons. *Statistics in Medicine*.
2013;**32**(30):5414-5429

[50] Rücker G, Schwarzer G. Ranking
treatments in frequentist network meta-
analysis works without resampling
methods. *BMC Medical Research
Methodology*. 2015;**15**(1):58

[51] Mbuagbaw L, Rochwerg B,
Jaeschke R, Heels-Andsell D,
Alhazzani W, Thabane L, et al.
Approaches to interpreting and
choosing the best treatments in network
meta-analyses. *Systematic Reviews*.
2017;**6**(1):79

[52] Schwarzer G, Carpenter JR,
Rücker G. *Meta-Analysis with R*.
Switzerland: Springer International
Publishing; 2015