What is ...?

For further titles in the series, visit: **www.whatisseries.co.uk**

What is indirect comparison?

Kelly Fleetwood MSc Principal Statistician, Quantics Consulting Ltd
Reviewer: Ann Yellowlees PhD Director, Quantics Consulting Ltd

- **Indirect comparison** can be used to compare treatments that have not been directly compared with each other in a head-to-head trial. It is often used when there is no evidence or insufficient evidence from head-to-head trials, or when more than two treatments are of interest.

- Indirect comparisons are usually conducted using **network meta-analysis**, an extension of meta-analysis that includes more than two treatments. Network meta-analysis is also referred to as **multiple-treatments meta-analysis**. Network meta-analysis includes **indirect treatment comparison** and **mixed treatment comparison**, although all of these terms are often used interchangeably.

- Like meta-analysis, indirect comparison combines data from different studies (usually randomised controlled trials) in order to produce **overall estimates of treatment effects**. Basic assumptions required for indirect comparisons include a **homogeneity assumption** as per standard meta-analysis, a **similarity assumption** for indirect comparison and a **consistency assumption** for the combination of direct and indirect evidence. It is essential to fully understand these basic assumptions in order to produce valid indirect comparisons.

- Indirect comparison is often part of the **systematic review process**. The validity of any indirect comparison also **depends on the studies** on which it is based.

- The use of indirect comparison has increased rapidly in recent years, and indirect comparisons are now **accepted by many health technology assessment agencies**.

What is ...?

What is **indirect comparison?**

Why use indirect comparison?

Meta-analysis is used to summarise the relative efficacy (or safety) of two treatments that have been directly compared in multiple head-to-head trials. However, policy-makers, physicians and patients must often choose between more than two treatments, or between treatments that have not been directly compared in head-to-head trials. In such cases, indirect comparison is often possible.

Indirect comparison is possible when the treatments of interest form part of a connected network of similar trials (Figure 1; the similarity of the trials is discussed further in the section 'Basic assumptions'). Part (a) illustrates a standard pairwise meta-analysis – the two treatments, A and B, are connected by a line because they have been compared in head-to-head trials. In part (b), we are also interested in comparing treatments A and B; however, there are no head-to-head trials that include both A and B. Instead, both A and B have been compared with a third treatment, C. As A and B share a mutual comparator, it is possible to conduct an indirect comparison of A and B. In part (c), we are interested in comparing five treatments with each other. As the treatments form a connected network, it is possible to estimate the relative efficacy for any pair of treatments within the network.

In part (d), we are again interested in comparing treatments A and B. However, in this case, it is not possible to conduct a valid indirect treatment comparison. Treatment A has been compared with treatment C, and treatment B has been compared with treatment D, but the network is not connected.

Networks of treatments can range from very simple networks, such as that shown in Figure 1, part (b), to complex networks that include many treatments. The shape of the network shown in part (b) is very common because often new treatments will have been compared with a placebo or standard treatment but not with each other. Placebo-controlled trials are often sufficient for regulatory approval.[1,2] Complex networks may include many treatments and trials; for example, a recent network meta-analysis of antipsychotic drugs in schizophrenia involved 16 treatments and 212 trials with a total of 43,049 participants;[3] and a network meta-analysis of treatments for multiple sclerosis involved 145 treatments across 109 trials and 26,828 participants.[4]

The earliest and simplest method for indirect comparison was introduced by Bucher *et al* in 1997.[5] A **Bayesian method**, suitable for more complex networks, was introduced by Lu and Ades in 2004.[6] Since then, the use of indirect comparisons has become much more common. A systematic review by Lee identified indirect

Figure 1. Networks of treatments

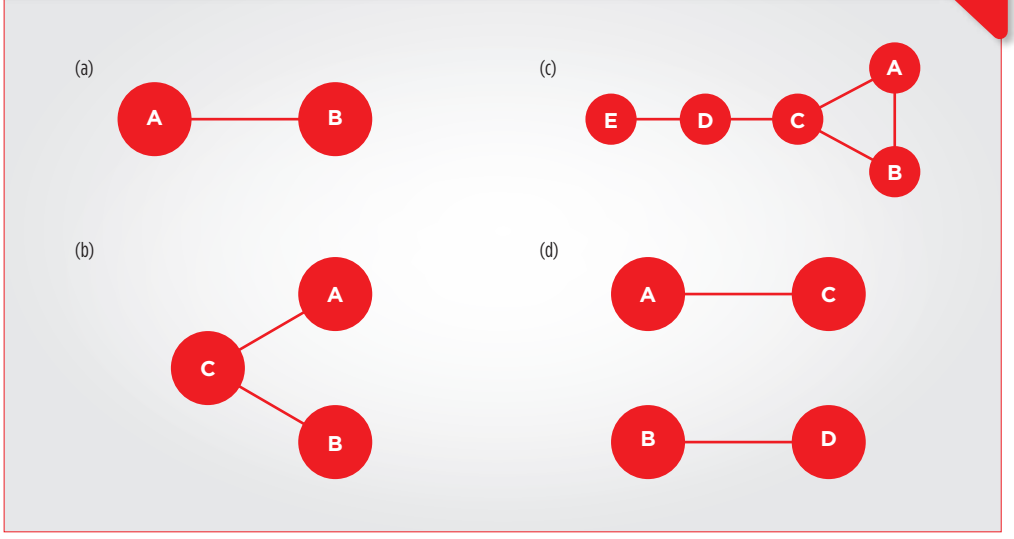


Table 1. Meta-analyses of risperidone versus haloperidol for schizophrenia: number of patients without clinical improvement[9]

Comparison	Number of trials	Log odds ratio (SE)	Odds ratio (95% CI)	I^2%
Placebo-controlled trials				
Risperidone vs placebo	3	–0.909 (0.218)	0.40 (0.26, 0.62)	37%
Haloperidol vs placebo	9	–1.707 (0.318)	0.18 (0.10, 0.34)	11%
Risperidone vs haloperidol				
Direct comparison	10	–0.262 (0.142)	0.77 (0.58, 1.02)	14%
Adjusted indirect comparison	3/9	0.798 (0.386)	2.22 (1.04, 4.72)	
Combination of direct and indirect estimates	10+(3/9)	0.207 (0.527)	1.23 (0.44, 3.45)	

NB A random-effects model was used in meta-analyses of trials and for the combination of the direct and indirect estimates. Odds ratio = EXP(log odds ratio)
CI: confidence interval; SE: standard error
Reprinted from J Clin Epidemiol, Vol 61, Song F, Harvey I, Lilford R, Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions, Pages 455–463, Copyright (2008), with permission from Elsevier

comparisons published up to June 2012.[7] He noted a rapid increase in their use from 2009 onwards. Indirect comparisons were used to compare treatments for many indications, including heart and circulation conditions, rheumatology conditions, cancer, and endocrine and metabolic conditions.[7]

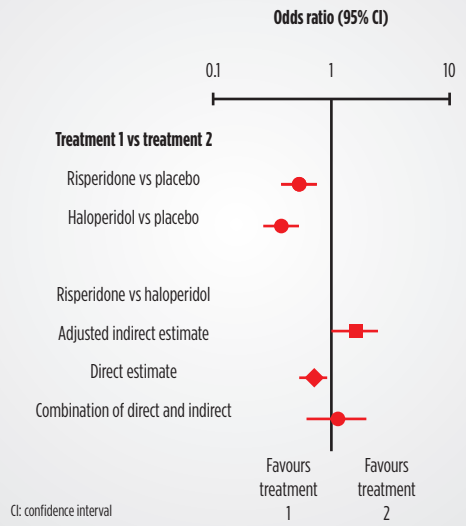
Methods for indirect comparison are now accepted by many health technology assessment agencies, including the National Institute for Health and Care Excellence (NICE) in the UK, the Canadian Agency for Drugs and Technologies in Health (CADTH), the Pharmaceutical Benefits Advisory Committee in Australia, the Institute for Quality and Efficiency in Health Care in Germany and the Haute Autorité de Santé in France.[8]

Motivating example

To illustrate these methods, we use an example of **risperidone versus haloperidol** for the treatment of schizophrenia, which includes three randomised, placebo-controlled trials of risperidone, nine randomised, placebo-controlled trials of haloperidol, and ten randomised controlled trials that directly compared risperidone and haloperidol (Table 1 and Figure 2).[9] The primary outcome in this example is the proportion of patients without clinical improvement.[9]

At first, we consider only the indirect evidence – that is, the trials that compare placebo with risperidone and the trials that compare placebo with haloperidol. It may be tempting to evaluate the active treatments by comparing the proportion of patients in the risperidone arms without clinical improvement with the equivalent proportion in the haloperidol arms. However, this method is not advised and is often referred to as **naive indirect comparison**.[10] Naive indirect comparison treats different arms as though they were from the same trial and does not account for

Figure 2. Results of different methods of comparing risperidone and haloperidol for schizophrenia. Outcome: not clinically improved[9]



Reprinted from J Clin Epidemiol, Vol 61, Song F, Harvey I, Lilford R, Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions, Pages 455–463, Copyright (2008), with permission from Elsevier

What is indirect comparison?

the randomisation within trials. This method is equivalent to using evidence from observational studies and is susceptible to bias. It may result in flawed recommendations for clinical practice. Therefore, naive indirect comparison should not be used to synthesise data from randomised controlled trials.[10]

The simplest valid method is **adjusted indirect comparison**.[5] Valid methods preserve the randomisation within trials. Instead of comparing the results from individual arms within trials, adjusted indirect comparison compares the relative results versus a common comparator. In this case, the relative efficacy of the active treatments versus placebo can be measured by the odds ratio (OR). The first stage uses standard pairwise meta-analysis to estimate the overall OR for each active treatment versus placebo (Table 1).[9] The overall OR for risperidone versus haloperidol (OR_{RH}) can then be estimated by dividing the OR for risperidone versus placebo (OR_{RP}) by the OR for haloperidol versus placebo (OR_{HP}).

$$OR_{RH} = \frac{OR_{RP}}{OR_{HP}} = \frac{0.40}{0.18} = 2.22$$

It is also possible to calculate the variance of this estimate and hence its confidence interval (CI): 1.04, 4.72 (Table 1).[9] The result suggests that the odds of no clinical improvement are higher in patients treated with risperidone. Hence, the result is favourable to haloperidol.

Basic assumptions

The following three basic assumptions are necessary for valid indirect comparison. The assumptions all concern the homogeneity of the studies included in the network: studies should be similar in terms of patient characteristics, the treatments administered (for example, doses and formulations) and the outcomes measured.[11,12]

Homogeneity

In standard meta-analysis, it is assumed that different trials estimate the same single effect (**fixed-effect model**) or different effects are distributed around a typical value (**random-effects model**).[12] For each pairwise comparison within a network, the trials should be sufficiently homogeneous in terms of their characteristics.[11,12] For example, if we consider the trials that compared risperidone versus placebo, all of these trials should have recruited patients with a similar baseline severity of disease, and evaluated similar doses of risperidone. Heterogeneity in results across studies can be quantified using I^2.[13] The I^2 values are included in Table 1 and indicate low to moderate heterogeneity.[9]

Similarity (or transitivity)

For an indirect comparison to be valid, it is necessary to extend the homogeneity assumption. All the trials that contribute to an indirect comparison must be similar enough to enable a fair estimate of the treatment effect to be calculated. This assumption is referred to as the similarity (or transitivity) assumption.[11,12]

Although it is possible to measure heterogeneity, it is not possible to measure similarity. Similarity must be carefully evaluated by examining the characteristics of each trial to be included in the indirect comparison.[11,12] All of the patients should have been eligible to participate in any of the trials.[11] Systematic differences between A versus B trials and A versus C trials could lead to invalid and misleading comparisons of B and C.[11,12]

Relative effects estimated in trials may also be associated with the methodological quality of the trials. Empirical evidence indicates that poor-quality trials may report greater treatment effects than well-conducted trials, particularly when outcomes are subjectively measured. Such evidence showed that the results of adjusted indirect comparison will be unbiased when the two sets of trials are similarly biased.[9]

Consider the indirect comparison of risperidone and haloperidol via placebo (Table 1).[9] Patient characteristics, dose of drug and treatment duration were similar between trials. However, the clinical improvement was defined differently between the two sets of placebo-controlled trials. In placebo-controlled trials of risperidone, clinical improvement was predefined as a 20% or greater reduction in

the total score on the Positive and Negative Syndrome Scale or Brief Psychiatric Rating Scale. In placebo-controlled trials of haloperidol, clinical improvement was rated by clinicians using the Clinical Global Impression or other scales.[9] Therefore, the results of indirect comparison should be interpreted cautiously.

Consistency

Consistency can be evaluated when both direct and indirect evidence is available. Consistency refers to the agreement between the two sets of evidence.[11,12] In our example, the indirect comparison via placebo led to an OR of 2.22 (95% CI: 1.04, 4.72) in favour of haloperidol. However, the direct comparison leads to an OR of 0.77 (95% CI: 0.58, 1.02) in favour of risperidone (Table 1 and Figure 2).[9] The discrepancy between the direct and indirect estimates is statistically significant (p=0.01). Although we can quantitatively combine the two estimates, the pooled estimate (OR 1.23; 95% CI: 0.44, 3.45) may be invalid and misleading.[9]

When both direct and indirect evidence is available, consistency should always be evaluated.[11,12] If inconsistency is detected, then it is important to investigate possible causes of the discrepancy. In our example, it has already been noted that the trials used different definitions of clinical improvement.[9] This may have contributed to the inconsistency. The doses and length of follow-up in all 22 trials in the network should also be reviewed, with assistance from a clinician in the field.

When the direct comparison differs from the adjusted indirect comparison, we should usually give more credibility to evidence from head-to-head comparative trials. However, evidence from direct comparative trials may not always be valid. Some case studies have indicated that adjusted indirect comparison may provide less biased results than head-to-head comparative trials under certain circumstances.[9]

Methods for indirect comparison

An indirect comparison, like a meta-analysis, should be based on the results of a well-conducted systematic review. A systematic review of more than two treatments is sometimes called a **comparative effectiveness review** or comparing multiple interventions review.[11] The review should follow a clear protocol and include an extensive search that is designed to identify all relevant studies. Further details about the systematic review process are provided in *What is meta-analysis?*[14] and *What is a systematic review?*[15] in this series.

Once the review is completed, the results must be carefully assessed in order to determine what indirect comparisons are possible. This will depend on whether the above assumptions are reasonable and whether connected networks can be formed. This can be a time-consuming process, especially if there are many treatments, studies or outcomes.

In the motivating example, the primary outcome was binary (the proportion of patients without clinical improvement) and the results were summarised on the OR scale.[9] For binary outcomes, the OR scale is most common, but indirect comparison can also be conducted on the relative risk or risk difference scales. It is also possible to use indirect comparison to evaluate other types of outcomes, including:

- Continuous outcomes; for example, the mean change from baseline in some measurements
- Survival outcomes; for example, overall survival or progression-free survival on the hazard ratio scale.

The same general methods for indirect comparison apply regardless of the type of outcome.[16]

In the motivating example, we used adjusted indirect comparison to compare risperidone versus haloperidol. This was a two-stage process: first, pairwise meta-analysis was conducted for each treatment versus placebo; and, second, the results of the two meta-analyses were combined. In practice, indirect comparison is usually conducted using more complex methods that cover both of these stages simultaneously.[17]

The complex methods are referred to as **network meta-analysis** (or **multiple-treatments meta-analysis**). Network

What is **indirect comparison?**

meta-analysis includes **indirect treatment comparison** and **mixed treatment comparison**. The use of the terminology varies across publications. We follow the definitions set out by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).[1] It uses 'indirect treatment comparison' to describe the synthesis of evidence across networks that do not contain any loops; for example, indirect treatment comparison could be applied to the network shown in Figure 1, part (b). 'Mixed treatment comparison' is used to describe the synthesis of evidence across networks that include loops; for example, mixed treatment comparison could be applied to the network shown in Figure 1, part (c). Most network meta-analysis is conducted using Bayesian methods, and WinBUGS is a common choice of software.[7,11] An introduction to Bayesian statistics is provided in *What is Bayesian statistics?*[18] in this series.

Like standard pairwise meta-analyses, network meta-analyses can be summarised by forest plots. When Bayesian methods are used, it is common to report an estimate of the treatment effect (for example, an OR), along with a 95% credible interval. Credible intervals are the Bayesian analogue of CIs. Bayesian methods also allow the treatments to be ranked, by estimating the probability that each treatment is best.[16]

Network meta-analysis is a developing field and new methodology is continuously being published. Recent developments include methods that incorporate different types of outcomes and methods that incorporate both individual patient data and aggregate data to adjust for differences between studies in patient characteristics.[19,20]

Appropriate use of indirect comparisons

A recent survey of published systematic reviews found that methodological problems in the use of indirect comparisons for evaluating healthcare interventions include the use of inappropriate or flawed methods, the lack of objective or validated methods to assess trial similarity, and inadequate comparison or inappropriate combination of direct and indirect evidence. In particular, mixed treatment comparison has often been used to combine direct and indirect evidence without explicit assessment of evidence consistency. To use indirect and mixed treatment comparisons appropriately, we need to adequately understand and fully appreciate the basic assumptions underlying valid indirect approaches.[12]

Fortunately, tools are available to help researchers evaluate the quality of published indirect comparisons. The **NICE Decision Support Unit** provides a series of technical support documents for network meta-analysis (available from: www.nicedsu.org.uk/Evidence-Synthesis-TSD-series(2391675).htm). The seventh document in the series provides a checklist for reviewers of indirect comparisons.[21] ISPOR has also recently produced guidelines for assessing the relevance and credibility of indirect comparisons.[8] The guidelines cover 22 aspects of indirect comparison – from systematic review, to the methods of analysis and interpretation of results.

Conclusions

Meta-analysis is a powerful tool for evaluating pairs of treatments that have been directly compared in head-to-head trials. However, in many situations, multiple treatments are of interest, and no trials compare all of the treatments. Yet, policy-makers, physicians and patients must make informed decisions about the most appropriate treatments. Indirect comparison can be used to synthesise evidence from multiple treatments that have not all been compared directly. It provides decision-makers with a means to select the most appropriate treatments ●

References
1. Jansen JP, Fleurence R, Devine B *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 Health* 2011; **14:** 417–428.
2. 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.
3. Leucht S, Cipriani A, Spineli L *et al*. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet* 2013; **382:** 951–962.
4. Zintzaras E, Doxani C, Mprotsis T, Schmid CH, Hadjigeorgiou GM. Network analysis of randomized controlled trials in multiple sclerosis. *Clin Ther* 2012; **34:** 857–869.
5. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct

and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997; **50:** 683–691.
6. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004; **23:** 3105–3124.
7. Lee AW. Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. *J Clin Epidemiol* 2014; **67:** 138–143.
8. Jansen JP, Trikalinos T, Cappelleri JC *et al*. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force Report. *Value Health* 2014; **17:** 157–173.
9. Song F, Harvey I, Lilford R. Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *J Clin Epidemiol* 2008; **61:** 455–463.
10. Glenny AM, Altman DG, Song F *et al*. Indirect comparisons of competing interventions. . *Health Technol Assess* 2005; **9:** 1–134, iii–iv.
11. 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* 2012; **3:** 80–97.
12. Song F, Loke YK, Walsh T *et al*. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009; **338:** b1147.
13. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med* 2002; **21:** 1539–1558.
14. Hodgson R. *What is meta-analysis?,* 3rd edn. London: Hayward Medical Communications, 2014.
15. McCool R, Glanville J. *What is a systematic review?,* 3rd edn. London: Hayward Medical Communications, 2014.
16. Welton NJ, Sutton AJ, Cooper NJ, Abrams KR, Ades AE. *Evidence Synthesis for Decision Making in Healthcare.* John Wiley & Sons, 2012.
17. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005; **331:** 897–900.
18. Stevens JW. *What is Bayesian statistics?* London: Hayward Medical Communications, 2014.
19. Woods BS, Hawkins N, Scott DA. Network meta-analysis on the log-hazard scale, combining count and hazard ratio statistics accounting for multi-arm trials: a tutorial. *BMC Med Res Methodol* 2010; **10:** 54.
20. Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. *Stat Med* 2012; **31:** 3516–3536.
21. Ades AE, Caldwell DM, Reken S *et al*. NICE DSU Technical Support Document 7: Evidence synthesis of treatment efficacy in decision making: a reviewer's checklist (2012). www.nicedsu.org.uk/TSD7%20reviewer%20checklist.final.08.05.12.pdf (last accessed 17 July 2014)

Further reading

The references listed above are suggested as useful further reading. In particular, Jansen *et al* (1) provide a general introduction to network meta-analysis; Salanti (11) provides an excellent discussion of the assumptions of indirect comparison.

What is **indirect comparison?**

First published 2009
Revised edition 2014
This publication, along with the others in the series, is available online at
www.whatisseries.co.uk

Published by Hayward Medical Communications, a division of Hayward Group Ltd, The Pines, Fordham Road, Newmarket CB8 7LG.
Tel: 01638 723560. Fax: 01638 723561.
email: hmc@hayward.co.uk

For reprints, permissions or to discuss training packages, email: **hmc@hayward.co.uk**