Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews

Fujian Song, reader in research synthesis,[1] Yoon K Loke, senior lecturer in clinical pharmacology,[1] Tanya Walsh, lecturer in dental statistics,[2] Anne-Marie Glenny, lecturer in evidence based oral care,[2] Alison J Eastwood, senior research fellow,[3] Douglas G Altman, professor and director[4]

[1]Faculty of Health, University of East Anglia, Norwich NR4 7TJ

[2]School of Dentistry, University of Manchester, Manchester

[3]Centre for Reviews and Dissemination, University of York, York

[4]Centre for Statistics in Medicine, Oxford

Correspondence to: F Song
fujian.song@uea.ac.uk

ABSTRACT

Objective To investigate basic assumptions and other methodological problems in the application of indirect comparison in systematic reviews of competing healthcare interventions.

Design Survey of published systematic reviews.

Inclusion criteria Systematic reviews published between 2000 and 2007 in which an indirect approach had been explicitly used.

Data extraction Identified reviews were assessed for comprehensiveness of the literature search, method for indirect comparison, and whether assumptions about similarity and consistency were explicitly mentioned.

Results The survey included 88 review reports. In 13 reviews, indirect comparison was informal. Results from different trials were naively compared without using a common control in six reviews. Adjusted indirect comparison was usually done using classic frequentist methods (n=49) or more complex methods (n=18). The key assumption of trial similarity was explicitly mentioned in only 40 of the 88 reviews. The consistency assumption was not explicit in most cases where direct and indirect evidence were compared or combined (18/30). Evidence from head to head comparison trials was not systematically searched for or not included in nine cases.

Conclusions Identified methodological problems were an unclear understanding of underlying assumptions, inappropriate search and selection of relevant trials, use of inappropriate or flawed methods, lack of objective and validated methods to assess or improve trial similarity, and inadequate comparison or inappropriate combination of direct and indirect evidence. Adequate understanding of basic assumptions underlying indirect and mixed treatment comparison is crucial to resolve these methodological problems.

Appendix 1 PubMed search strategy
Appendix 2 Characteristics of identified reports
Appendix 3 Identified studies
References of included studies

INTRODUCTION

The number of available healthcare interventions increases with time, reflecting advances in science and technology. For many clinical indications clinicians may have to choose among several competing interventions. In this era of evidence based decision making, relative effectiveness and cost effectiveness of different interventions need to be objectively and accurately assessed in clinical studies. It has been accepted generally that well designed and implemented head to head randomised controlled trials provide the most rigorous and valid research evidence on the relative effects of different interventions.[1] Evidence from head to head comparison trials is often limited or unavailable, however, and indirect comparison may therefore be necessary.[2 3]

Indirect comparison may be done narratively—for example, by discussing the results of separate systematic reviews of different interventions for a given condition. A simple but inappropriate statistical method is to compare the results of individual arms from different trials as if they were from the same randomised controlled trial. This naive or unadjusted indirect comparison has been criticised for discarding the within trial comparison, increasing liability to bias and over-precise estimates.[2] In contrast to such a naive indirect comparison, the adjusted indirect comparison can take advantage of the strength of randomised controlled trials in making unbiased comparisons.[4 5] Here the indirect comparison of different interventions is adjusted by comparing the results of their direct comparisons with a common control group.

Box 1 provides an example of adjusted indirect comparison. The case study involves a comparison of bupropion and nicotine replacement therapy patch for smoking cessation, based on a direct comparison trial[6] and an adjusted indirect comparison using 21 placebo controlled trials.[7]

Indirect comparison
The results of placebo controlled trials suggested that both bupropion and nicotine replacement therapy are more effective than placebo for smoking cessation. The results of the two sets of placebo controlled trials can also be used to indirectly compare bupropion with nicotine replacement therapy patch. The magnitude
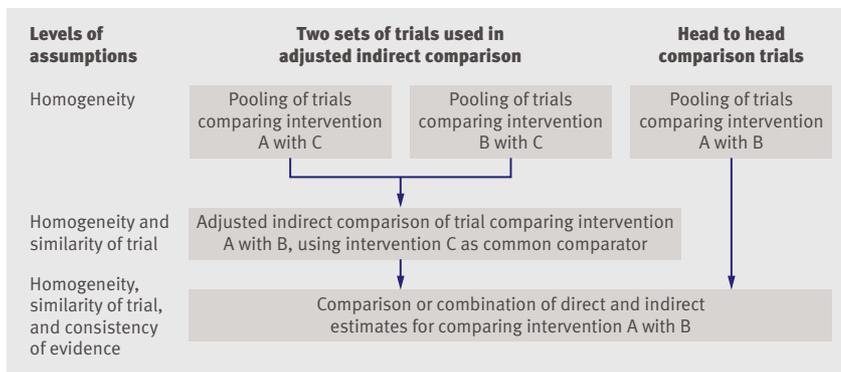
Fig 1| Assumptions underlying adjusted indirect and mixed treatment comparison

of treatment effect of bupropion compared with placebo (odds ratio 0.51, 95% confidence interval 0.36 to 0.73) was similar to that of nicotine replacement therapy compared with placebo (0.57, 0.48 to 0.67). Therefore it could be indirectly concluded that bupropion was as effective as nicotine replacement therapy. The adjusted indirect comparison can also be formally done, using one of several methodologically sound methods. The result of adjusted indirect comparison suggests that bupropion was as effective as nicotine replacement therapy for smoking cessation (0.90, 0.61 to 1.34), although the confidence interval is wide. The validity of the adjusted indirect comparison depends on a similarity assumption, which assumes that the two sets of placebo controlled trials are sufficiently similar for moderators of relative treatment effect.

Comparison of direct and indirect estimates
The result of the head to head comparison trial suggested that bupropion was more effective than nicotine replacement therapy for smoking cessation (0.48, 0.28 to 0.82), which is different from the result of adjusted indirect comparison (0.90, 0.61 to 1.34). The discrepancy between the direct and indirect estimate was marginally statistically significant (I^2=71%, P=0.06). Statistical methods are available to combine the results of direct and indirect evidence (combined odds ratio 0.68, 95% confidence interval 0.37 to 1.25). A consistency assumption is, however, required to combine the direct and indirect estimate. The combination of inconsistent evidence from different sources may provide invalid and misleading results.

To improve statistical power evidence generated by indirect comparison can be combined with evidence from head to head trials.[8-10] The combination of direct and indirect evidence has been facilitated by the development of network meta-analysis[11] and Bayesian hierarchical models for mixed treatment comparisons.[12] These methods are especially helpful when the relevant trials have considered many different treatments.

Available empirical evidence indicates that the results of an adjusted indirect comparison usually but not always agree with the results of direct comparison trials.[4] Recently, conflicting evidence has emerged about the validity of indirect comparison. For

example, a report of three case studies suggested that adjusted indirect comparison may be less biased than head to head trials for evaluating new drugs.[7] One trial, however, concluded that "indirect comparisons could be unreliable for complex and rapidly evolving interventions" such as initial highly active antiretroviral therapy.[13] Therefore the potential usefulness of adjusted indirect comparison is still overshadowed by concern about bias resulting from its misuse.

Existing statistical methods for adjusted indirect comparison and mixed treatment comparison are unbiased, but only if some assumptions are fulfilled.[2] One study described the assumption for indirect comparison as the constant treatment effect "across differences in the populations' baseline characteristics."[5] The only requirement mentioned by another study was that the populations of the individual studies contain some subpopulation in common.[14] Therefore the description of important assumptions underlying indirect comparison may not be clear in some methodological studies. For mixed treatment comparison it was noted that "the only additional assumption is that the similarity of the relative effects of treatment holds across the entire set of trials, irrespective of which treatments were actually evaluated."[8] The additional assumption may hold to a subset of trials, although it may not be the case across the entire set of trials. For instance, an indirect comparison may be valid but its result may not be consistent with the result of a direct comparison.[7] Therefore we suggest a framework to delineate the main assumptions related to indirect and mixed treatment comparison (fig 1).

In standard meta-analysis of randomised trials it is assumed that different trials are sufficiently (not necessarily completely) homogeneous and that they estimate the same single treatment effect (fixed effect model) or different treatment effects distributed around a typical value (random effects model).[15 16] We refer to this assumption for standard meta-analysis as the homogeneity assumption, to distinguish it from other related assumptions. In adjusted indirect comparison, the homogeneity assumption for conventional meta-analysis should be fulfilled when multiple trials are involved. For the example in box 1, the nine placebo controlled trials of bupropion should be homogeneous enough to be pooled in meta-analysis; and the same for the 19 placebo controlled trials of nicotine replacement therapy.

In addition to the homogeneity assumption, a similarity assumption is required for adjusted indirect comparison—namely, that trials are similar for moderators of relative treatment effect.[2 5] That is, for the indirect comparison of bupropion compared with nicotine replacement therapy based on the common placebo control (box 1), the average relative effect estimated by placebo controlled trials of bupropion should be generalisable to patients in placebo controlled trials of nicotine replacement therapy, and vice versa. Trial similarity could be considered from two perspectives: clinical similarity and methodological similarity. Clinical similarity refers to similarity in patients'

characteristics, interventions, settings, length of follow-up, and outcomes measured. Methodological similarity refers to aspects of trials associated with the risk of bias. It is mathematically proved that adjusted indirect comparison may counterbalance bias in trials and provide an unbiased estimate, if the two sets of trials are similarly biased.[7]

A further assumption of consistency is required to combine results of direct and adjusted indirect comparison using fixed effect or random effects model.[2 11 17] Even when the adjusted indirect comparison is valid, the indirect evidence may not be consistent with evidence from head to head trials because of clinically meaningful heterogeneity.[5 7] For the case study of bupropion compared with nicotine replacement therapy for smoking cessation box 1), the result of the direct comparison was different from that of the adjusted indirect comparison (I^2=71%, P=0.06). The discrepancy between the direct and indirect estimate may result from several possible causes, including the play of chance, invalid indirect comparison, bias in the head to head comparison trial, and clinically meaningful heterogeneity across trials. For example, trials that evaluate new drugs may include patients who had responded poorly to old drugs. If the response to the new drug is not affected by the poor response to the old drug, the results of direct comparison of new and old drugs will be applicable only to patients who had a poor response to the old drugs, whereas the adjusted indirect comparison may be more generalisable to patients in general.

Comparisons and related assumptions may need to be further expanded when multiple sets of trials are involved in indirect comparison. For instance, the adjusted indirect comparison of intervention A with B could be done using trials that compared intervention A with D and those that compared intervention B with D as well as trials that compared
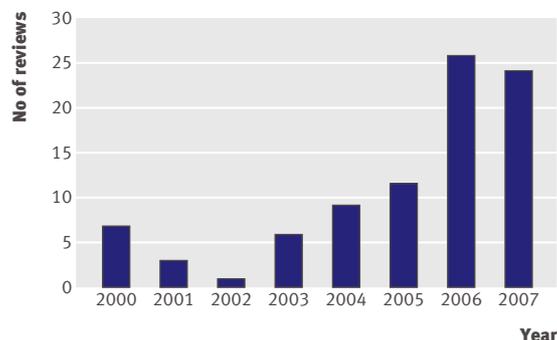
Fig 2 | Publication year of 88 review reports that explicitly used indirect comparison

intervention A with C and those that compared intervention B with C. A consistency assumption is necessary to pool the results from different indirect comparisons. Because of increased complexity of the data structure, assumptions additional to those outlined in figure 1 may be required for many mixed treatment comparisons.[2]

In summary, assumptions concerning adjusted indirect comparison and mixed treatment comparison are similar to but more complex than the underlying assumption for standard meta-analysis. At least three issues of comparability need to be considered: a homogeneity assumption for each meta-analysis, a similarity assumption for individual adjusted indirect comparison, and a consistency assumption for the combination of evidence from different sources (fig 1). The three issues of comparability concern the different levels of decisions for a research synthesis of clinical trials. The trial similarity assumption for adjusted indirect comparison is relevant only if the homogeneity assumption is valid; and the consistency assumption needs the prerequisite of both the homogeneity assumption and the similarity assumption (fig 1).

We report findings from a survey of methodological problems in the application of indirect and mixed treatment comparison. We determined which methods have been used for indirect comparison and what methodological problems can be identified in these applications of indirect comparison.

METHODS

We searched PubMed for systematic reviews or meta-analyses published between 2000 and 2007 in which indirect comparison had been explicitly used according to the titles or abstracts. The search strategy (see web extra appendix 1) was developed on the basis of our previous work[2] and modified after comparing the results of a preliminary search with some known references. Key terms used were "adjusted indirect comparison", "indirect" or "indirectly", "network meta-analysis", "mixed treatment", "multiple treatment", "multiple comparison", "cross-trial", and "cross-study". We limited the search to systematic reviews by combining the above terms with "meta-analysis" or "systematic[sb]". The literature search was done first in May 2007 and then updated in October 2008,

---

Box 1: A simple example of indirect comparison

The case study involves a comparison of bupropion with nicotine replacement therapy patch for smoking cessation.[7] The outcome was the number of smokers who failed to quit at 12 months (table). Indirect comparison can be made using two sets of randomised controlled trials: nine trials that compared bupropion with placebo and 19 that compared nicotine replacement therapy with placebo. One trial also directly compared bupropion with nicotine replacement therapy.[6]

Number of smokers failing to quit at 12 months, according to treatment group

Comparison	No of trials	Odds ratio (95% CI)	I^2 (%)
Bupropion *v* placebo	9	0.51 (0.36 to 0.73)	54
NRT patch *v* placebo	19	0.57 (0.48 to 0.67)	12
Bupropion *v* NRT patch:			
Direct comparison	1	0.48 (0.28 to 0.82)	—
Adjusted indirect comparison	9+19	0.90 (0.61 to 1.34)	—
Combined (direct +indirect)	1+(9+19)	0.68 (0.37 to 1.25)	71

NRT=nicotine replacement therapy.
See Bucher et al[5] and Song et al[4] for indirect comparison methods. Random effects model was used in meta-analyses of trials and for combination of direct and indirect estimates.[16]

Table 1 | Methods used for indirect comparison and availability of direct comparison evidence

Indirect comparison method	No (%) of included reviews	Direct comparison available	Comparison of indirect and direct evidence	Combination of indirect and direct evidence
Simple adjusted	49 (56)	15	11	2
Network or Bayesian approaches	18 (20)	16	9	15
Informal indirect	13 (15)	6	1	0
Naive indirect	6 (7)	2	1	0
Unclear	2 (2)	1	1	0
Total	88 (100)	40	24	17

and 745 references were retrieved. The titles and abstracts of the retrieved references were independently assessed by two reviewers to identify systematic reviews or meta-analyses of randomised controlled trials in which indirect comparison was explicitly used in the comparison of different healthcare interventions.

Full publications were obtained for systematic reviews identified on the basis of the titles or abstracts. After examining full publications we excluded reports in which primary studies were not randomised controlled trials, indirect comparison had not been done, or the articles were editorial style reviews. We also excluded duplicate publications of the same reports or considered them as one. For Cochrane reviews we used the most recent update, although the earliest version since 2000 is also cited.

From the included reports we extracted data on clinical indications, interventions compared, comprehensiveness of the literature search for trials used in indirect comparison, methods for indirect comparison, and whether direct evidence from head to head comparison trials was also available. We examined whether the assumption of similarity was explicitly mentioned and whether any efforts were made to investigate or improve the similarity of trials for indirect comparison. One reviewer (FS) extracted data and another reviewer (YL, A-MG, or TW) checked each study. Extracted data were summarised in tables and narratively described.

RESULTS

Overall, 88 review reports (91 publications) in which an indirect comparison was explicitly done (see web extra appendix 2 for the main characteristics of the included reports[w1-w91]) were included. Fifty nine were reviews of effectiveness of interventions published in journals, 19 were reports of health technology assessment or cost effectiveness analysis, six were Cochrane systematic reviews, and four were reviews used to illustrate methods for indirect comparisons.

Indirect comparison has become increasingly or more explicitly used in research syntheses for the evaluation of a wide range of healthcare interventions (figs 2 and 3). Indirect comparison was used to evaluate drug interventions in 72 of the 88 reviews. Of the 72 drug assessments, 43 compared drugs of different classes, 17 compared drugs of the same class, and 10

compared different formats or modes of delivery of the same drug. Two reviews compared the relative efficacy of an active drug with placebo. Non-drug interventions, including counselling, devices, and surgical or diagnostic procedures, were indirectly compared in 16 reviews.

The most commonly used approach (49/88) was the adjusted indirect comparison using classic frequentist methods (table 1). More complex methods (including network or Bayesian hierarchical meta-analysis and mixed treatment comparison methods) were used in 18 reviews. In 13 reviews, indirect comparison was informal, without calculation of relative effects or testing for statistical significance. In six reviews results from different trials were naively compared without using a common treatment control.

Most indirect comparisons (n=67) used outcome measures for categorical data, including odds ratio, risk difference, relative risk, or hazard ratio. Outcome measures for continuous data (mean difference or standardised mean difference) were used in 13 reviews, and eight reviews measured both categorical and continuous outcomes.

Direct evidence from head to head comparison trials was available in 40 of the 88 reviews (see web extra appendix 3), including 15 reviews that used simple adjusted methods, 16 that used more complex methods, and six that used informal methods. As compared with simple adjusted methods, complex methods were more likely to be used to combine the direct and indirect evidence. Where direct comparison was available, direct and indirect evidence were combined in 15 of the 16 reviews that used complex methods and in only two of the 15 reviews that used simple methods (table 1). Furthermore, direct and indirect evidence were less likely to be explicitly compared in reviews that used complex methods than in those that used simple methods (9/16 v 11/15).

The assumption of trial similarity was explicitly mentioned or discussed in only 40 of the 88 reviews (table 2). Explicit mention of the similarity assumption was associated with efforts to examine or improve the similarity between trials for indirect comparisons (30/40 v 19/48). Methods to investigate or improve trial
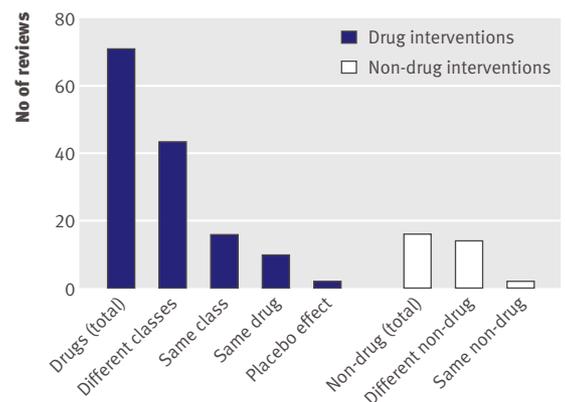
Fig 3 | Types of interventions indirectly compared in 88 review reports

Table 2 | Similarity assumptions and efforts to examine or improve similarity of trials used in indirect comparison

Explicit similarity assumption	Efforts reported		Efforts not reported	Total
	Subjective judgment	Subgroup or metaregression analysis		
Yes	16	14	10	40
No	10	9	29	48
Total	26	23	39	88

similarity included subjective judgment by a comparison of study characteristics (n=26) and subgroup and metaregression analysis to identify or adjust for possible moderators of treatment effects (n=23). The assumption of consistency was not explicit in most cases where direct and indirect evidence were compared or combined (18/30; table 3).

In eight of the 88 reviews, indirect comparison was based on data from other published systematic reviews or meta-analyses (see web extra appendix 2). Evidence from head to head comparison trials was not systematically searched for or not included in nine cases (see web extra appendix 2).

Box 2: Main methodological problems in using indirect comparison and recommendations

Methodological problems
- Unclear understanding of underlying assumptions
- Incomplete search and inclusion of relevant studies
- Use of flawed or inappropriate methods
- Lack of objective and validated methods to assess or improve trial similarity
- Inadequate comparison and inappropriate combination of direct and indirect evidence

Recommendations
- More explicit and elaborate description and discussion of underlying assumptions in methodological studies and in systematic reviews in which different interventions are indirectly compared
- Literature search needs to be systematic in order to identify all relevant studies
- The evidence from head to head comparison trials should not be excluded in reviews that use indirect comparison
- The availability of all active treatment controlled studies that are suitable for adjusted indirect comparison should be explicitly discussed, and justifications provided if only placebo controlled trials are used for adjusted indirect comparison
- Naive indirect comparison of different arms from different trials should be avoided
- Data from trials with multiple arms should be appropriately analysed, to avoid both downgrading direct evidence and using the same control group more than once in adjusted indirect comparison
- Methods for investigating heterogeneity in standard meta-analysis can be adopted to assess trial similarity in adjusted indirect comparison, including subgroup analysis, metaregression, and experts' subjective judgment. However, further research is needed
- Direct and indirect evidence should be separately presented and explicitly compared, whether or not the two sets of data are subsequently combined
- Possible reasons for any observed discrepancies between direct and indirect evidence should be investigated
- The consistency assumption should be explicitly assessed before direct evidence is combined with indirect evidence

DISCUSSION

Indirect comparison is being increasingly (or more explicitly) used for the evaluation of a wide range of healthcare interventions. The need for indirect comparison is more apparent in the assessment of effectiveness and cost effectiveness of healthcare interventions to support clinical and policy decision making. Sixteen of the 88 included reviews were health technology assessment reports. In many such reports, indirect comparison had not been done in relation to clinical effectiveness but was used in the economic evaluation. Researchers may have to use whatever data are available to estimate the incremental cost effectiveness of competing interventions.

In the existing literature, several related but different assumptions underlying adjusted indirect comparison (fig 1) have not been clearly distinguished, resulting in some methodological and practical problems in the use and interpretation of indirect or mixed treatment comparison. The identified problems include unclear understanding of underlying assumptions, inappropriate search and selection of relevant trials, use of inappropriate or flawed methods, lack of objective and validated methods to assess or improve trial similarity, and inadequate comparison or inappropriate combination of direct and indirect evidence.

What methods should be used for indirect comparison?

Indirect comparison was explicit but informal in 13 review reports. Informal indirect comparison means that neither relative effects nor statistical significance were formally calculated. Since the use of indirect comparison is often inevitable, a more explicit and formal approach is preferable to implicit and informal approaches. In six reviews of randomised controlled trials, the results from individual arms of different trials were compared naively as if they were from a single controlled trial. This naive indirect comparison is methodologically flawed because the strength of randomisation is totally disregarded.[2]

The strength of randomisation could be preserved in adjusted indirect comparison. The methods for simple adjusted indirect comparison have been described in several articles.[5 14 18 19] The most common scenario was the indirect comparison of two competing interventions adjusted by common comparators using classic frequentist methods (including simple metaregression). The advantages of the simple methods include ease of use and transparency. However, when there are several alternative interventions to be compared, the simple adjusted indirect comparison may become rather tedious and inconvenient. More complex methods, including network meta-analysis and Bayesian hierarchical models, which can be used to make simultaneous comparisons of multiple interventions, have been increasingly used.[9 11 12] These methods treat all included interventions equally rather than focusing on one particular comparison of two interventions.

Table 3 | Consistency assumption when direct and indirect evidence were compared or combined

Compared or combined direct and indirect evidence	Consistency assumption		Total
	Explicit	Not explicit	
Yes	12	18	30
No	0	10	10
Total	12	28	40

Trial similarity in adjusted indirect comparison

Awareness of the assumption of trial similarity is associated with increased efforts to investigate or improve trial similarity for adjusted indirect comparison. Subgroup analysis and metaregression are commonly used methods to assess or improve trial similarity for adjusted indirect comparison (table 2). Trials included in subgroup analyses are homogeneous for certain variables at study level, whereas indirect comparison could be further adjusted with selected study level variables in metaregression analysis. However, the usefulness of subgroup analysis and metaregression may be rather limited because the number of trials involved in adjusted indirect comparison was usually small and it was uncertain whether the important study level variables were reported in all relevant trials.

Trial similarity was often assessed by examining heterogeneity across trials and by a narrative comparison of trial characteristics for the different treatment comparisons being included, which may be deemed informal and subjective. One study noted that the assumptions underlying mixed treatment comparison "are unlikely to be statistically verifiable, and it seems reasonable to rely on expert clinical and epidemiological judgment."[8] Because of lack of transparency, however, it is unclear to what extent we could trust the reported judgment that trials were similar enough for an adjusted indirect comparison. Further research is needed in this area.

When both direct and indirect evidence are available

When data from head to head comparison trials are available, several important things need to be considered. Firstly, whether the use of indirect comparison is justified when direct comparison trials are available. Secondly, any discrepancies between direct and indirect evidence need to be interpreted sensibly. Thirdly, according to answers to the second question, whether direct evidence could be combined with the results of indirect comparison.

It is controversial whether indirect evidence needs to be considered when there is evidence from direct comparison trials.[5 8] Indirect comparison was seemingly considered still to be helpful by authors of the 40 reviews in which both direct and indirect evidence were available. Handling of the second and the third questions seemed to reflect the choice of methods used for indirect comparison. Direct and indirect evidence were less likely to be explicitly compared and more likely to be combined in review reports that used complex methods (network meta-analysis or mixed treatment comparison), as compared with simple adjusted indirect comparisons (table 1). Since the assumptions underlying mixed treatment comparison are much more arduous to verify than simple adjusted indirect comparison (fig 1), it seems odd that the comparison of direct and indirect evidence has not been more explicitly done in many published mixed treatment comparisons or network meta-analyses (table 1). The importance of investigating incoherence or inconsistency in the evidence network has been emphasised.[11 17] Since the evidence consistency is usually assessed informally and subjectively,[8] transparency is important to allow others to make their own judgment according to explicitly presented discrepancies. In the practice of research synthesis, the use of complex methods needs to focus more on the comparison of evidence from different sources.

Review reports may include trials with three or more arms. For example, the trial that directly compared bupropion and nicotine replacement therapy (see box 1) also has a placebo arm.[6] Some review reports separately compared the two active treatments with placebo within the same trial, and then the results of two separate comparisons were used in adjusted indirect comparison. This is methodologically inappropriate because it downgrades direct evidence to indirect evidence, reduces precision, and uses data from the same placebo arm twice. Evidence from head to head comparison trials should be clearly separated from evidence based on indirect approaches.

Literature search and study inclusion

Reviewed reports of indirect comparison often mentioned that no head to head comparison trials were available, but in some cases the availability of direct evidence was uncertain. In nine of the 88 reviews, direct comparison trials were excluded or not searched for systematically. It is worrying that direct evidence has sometimes been disregarded without good justification.

In review reports that included only placebo controlled trials, it was often unclear whether there were other active treatment controlled trials that could also be used for adjusted indirect comparison. For

---

WHAT IS ALREADY KNOWN ON THIS TOPIC

Indirect comparisons can be valid if some basic assumptions are fulfilled

The related but different methodological assumptions have not been clearly distinguished

WHAT THIS STUDY ADDS

Certain methodological problems may invalidate the results of evaluations using indirect comparison approaches

Understanding basic assumptions underlying indirect and mixed treatment comparison is crucial to resolve these problems

A framework can help clarify homogeneity, similarity, and consistency assumptions underlying adjusted indirect comparisons

---

completeness in utilising all available evidence, indirect comparison should include not only placebo controlled trials but also active treatment controlled trials when available. As a minimum, the exclusion of other active treatment controlled trials should be explicit and with adequate justification.

Some indirect comparisons seemed to be done on an ad hoc basis, using data from existing systematic reviews and meta-analyses. The use of indirect comparison should be more systematic in adhering to generally accepted standards for systematic reviews.

Limitations and implications of the study

Review reports could be included in this survey only if the indirect comparison was explicit in their titles and abstracts, and if they were indexed in PubMed. Thus we may have missed many review reports of indirect comparisons. The use of indirect comparison in possibly missed review reports may have been less explicit and less formal than review reports included in this survey, thus not meriting a mention in the abstract. It is possible that methodological problems would be more frequent and severe in review reports that were not identified.

Findings from this study can be used to promote valid indirect comparisons and to reduce invalid ones. Box 2 summarises identified methodological problems in the use of indirect approaches and corresponding recommendations to resolve these problems. However, empirical evidence on the validity of indirect and mixed treatment comparison is still limited and many questions remain unanswered. For example, it is unclear whether the assessment of trial similarity and evidence consistency was appropriately done in the included review reports. In addition, there is only limited empirical evidence to show that improved trial similarity is associated with improved validity of indirect and mixed treatment comparison. Further research is required to answer these questions.

Conclusion

We have identified certain methodological problems that may invalidate the results of evaluations using indirect comparison approaches. Adequate understanding of basic assumptions underlying indirect and mixed treatment comparison is crucial to resolve these methodological problems.

1. Pocock SJ. *Clinical trials: a practical approach*. New York: Wiley, 1996.
2. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technol Assess* 2005;9:1-134.
3. Ioannidis JP. Indirect comparisons: the mesh and mess of clinical trials. *Lancet* 2006;368:1470-2.
4. 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;326:472-5.
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;50:683-91.
6. Jorenby DE, Leischow SJ, Nides MA, Rennard SI, Johnston JA, Hughes AR, et al. A controlled trial of sustained-release bupropion, a nicotine patch, or both for smoking cessation. *N Engl J Med* 1999;340:685-91.
7. 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;61:455-63.
8. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897-900.
9. Higgins JP, Whitehead A. Borrowing strength from external trials in a meta-analysis. *Stat Med* 1996;15:2733-49.
10. Song F, Glenny AM, Altman DG. Indirect comparison in evaluating relative efficacy illustrated by antimicrobial prophylaxis in colorectal surgery. *Control Clin Trials* 2000;21:488-97.
11. Lumley T. Network meta-analysis for indirect treatment comparisons. *Stat Med* 2002;21:2313-24.
12. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23:3105-24.
13. Chou R, Fu R, Huffman LH, Korthuis PT. Initial highly-active antiretroviral therapy with a protease inhibitor versus a non-nucleoside reverse transcriptase inhibitor: discrepancies between direct and indirect meta-analyses. *Lancet* 2006;368:1503-15.
14. Hasselblad V, Kong DF. Statistical methods for comparison to placebo in active-control trials. *Drug Inf J* 2001;35:435-49.
15. Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954;3:101-29.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177-88.
17. Lu GA, Ades AE. Assessing evidence inconsistency in mixed treatment comparisons. *JASA* 2006;101:447-59.
18. Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;326:219.
19. Fisher LD, Gent M, Buller HR. Active-control trials: how would a new agent compare with placebo? A method illustrated with clopidogrel, aspirin, and placebo. *Am Heart J* 2001;141:26-32.