



OPEN ACCESS

# Simple tool implementing Bucher's method for indirect treatment comparisons in meta-analyses

Samuel James Tingle ,<sup>1</sup> Georgios Kourounis,<sup>1</sup> Felix Liu,<sup>1</sup> Colin H Wilson,<sup>1</sup> Rachel Richardson ,<sup>2</sup> Josie Sandercock<sup>3</sup>

10.1136/bmjebm-2024-113103

► Additional supplemental material is published online only. To view, please visit the journal online (<https://doi.org/10.1136/bmjebm-2024-113103>).

<sup>1</sup>Translational and Clinical Research Unit, Newcastle University, Newcastle upon Tyne, UK

<sup>2</sup>Cochrane Evidence Production and Methods Directorate, Cochrane, London, UK

<sup>3</sup>Systematic Review Initiative, NHS Blood and Transplant, Bristol, UK

Correspondence to:  
**Dr Samuel James Tingle;**  
[samjamestingle@gmail.com](mailto:samjamestingle@gmail.com)



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

**To cite:** Tingle SJ, Kourounis G, Liu F, *et al.* *BMJ Evidence-Based Medicine* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjebm-2024-113103

## Standfirst

Healthcare practitioners and patients are often required to make choices between two treatment options which may have never been compared directly. When performing standard 'pairwise' meta-analysis, it is common to find trials comparing one of two experimental treatments against a common control (standard of care or placebo) but no head-to-head trials. Statistical methods are available to perform indirect comparisons between two treatments when direct evidence is lacking. Here, we provide a simple, easy-to-use, Excel spreadsheet capable of performing these calculations.

## Introduction

The statistical foundation for indirect comparisons in the setting of meta-analysis was first described by Bucher *et al.*<sup>1</sup> They discussed the common situation where two treatments have been compared with the same control treatment but have not been compared directly to each other. Figure 1 displays this simple type of 'network' where a common control (treatment A) is compared with treatments B and C. The key to Bucher's technique is that the variance (the square root of which is used to calculate the confidence interval of the indirect comparison B versus C is equal to the variance of the direct comparison A versus B plus the variance of the comparison A versus C. For ratio effect estimates, such as the OR, risk ratio (relative risk) and HR, calculations are done on a logarithmic scale.

The direct comparisons can be results from a single clinical trial or, ideally, the results of a systematic review and meta-analysis if multiple trials have compared the same two interventions in a similar clinical context.

## Key assumptions

When performing indirect comparisons, the assumption of transitivity must be met.<sup>2</sup> For the indirect comparison to be valid, the A versus B and A versus C trials should not differ with respect to potential effect modifiers, for example, the age of the participants. Their eligibility criteria should be comparable, with the treatment regimens in the shared arm (A), and arms B and C, being comparable across all studies.

This transitivity assumption is measured by statistical tests when conducting larger network meta-analyses (NMAs); however, this is not possible with a small network containing three treatments and two direct comparisons.<sup>3</sup> Instead,

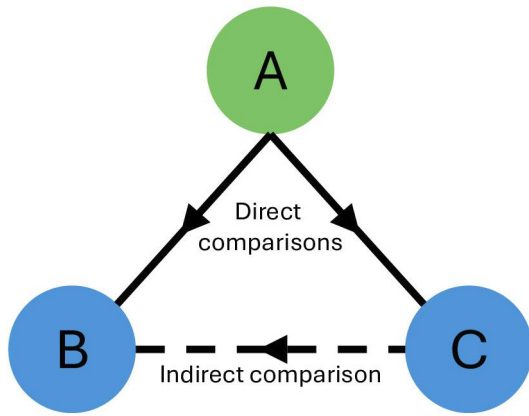
we can assess for violations in the transitivity assumption by checking that the study characteristics of the three trials are broadly similar.<sup>2</sup> If the direct comparisons come from pairwise meta-analysis, heterogeneity in these meta-analyses might be caused by violation of the transitivity assumption. If there is no statistical heterogeneity in a meta-analysis with a non-negligible sample size from a properly conducted systematic review, it is reasonable to conclude that all the trials were attempting to measure the same effect. One should also check that the mean outcomes for a specific intervention are similar across all trials reporting on that intervention, and that adherence to assigned treatments is similar across studies (differences in either suggesting that the assumption of transitivity is violated). However, differences in the baseline risk do not necessarily mean that there will be differences in the effect size.

When performing any meta-analysis or indirect comparison, it is critical to assess the certainty of the evidence.<sup>4</sup> There are some specific points to highlight for indirect comparisons. First, the indirect comparison is subject to all of the internal biases of the individual trials, which should be formerly assessed with risk of bias tools, and these must be considered when assessing the certainty of the indirect effect estimate.<sup>5</sup> If the outputs of pairwise meta-analyses are used as the direct effect estimates, it is important to assess heterogeneity between individual studies contributing to these meta-analyses. Heterogeneity reduces the certainty of evidence from pairwise meta-analyses and will therefore reduce the certainty of evidence for an indirect estimate which uses these as 'inputs'.<sup>6</sup> Where there is heterogeneity, a random effects model should be used to pool treatment effects to avoid underestimating the uncertainty in the direct estimates.<sup>7</sup>

## Case study

This example is taken from a Cochrane review on techniques to preserve donated livers for transplantation.<sup>8</sup> Both cold and warm machine perfusion have been compared with standard of care ('ice-box storage') in several randomised trials. However, there are no trials directly comparing cold and warm machine perfusion.

The authors of this review deemed there to be no important differences between trials with respect to potential effect modifiers (specifically study and participant characteristics were sufficiently similar) and there was no evidence of



**Figure 1** Standard of care treatment A is compared with treatments B and C in separate clinical trials. There are no trials which directly compare B with C, making indirect comparison necessary.

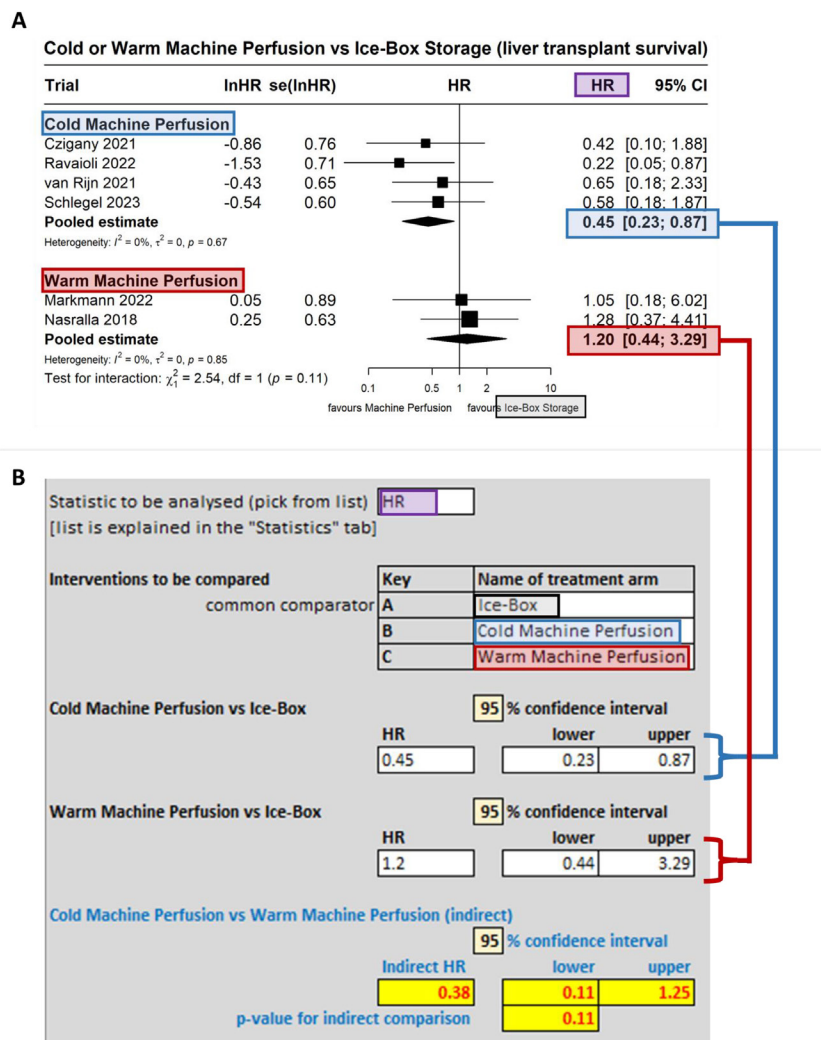
statistical heterogeneity in either of the pairwise meta-analyses. This suggests that all of the trials were asking the same or very similar questions with respect to the HRs they set out to estimate. As there was no evidence of violation of the transitivity

assumption, the Excel tool (online supplemental materials) can be used to generate an indirect comparison.

Figure 2 displays how the output of the standard pairwise meta-analyses for graft survival are entered into the tool. The resulting indirect HR for cold versus warm machine perfusion is 0.38 (95% CI 0.11 to 1.25, p=0.11). As discussed above, it is important to assess the certainty of evidence of this result, which cannot be higher than the certainty of evidence for each of two direct comparisons. The referenced review deemed this indirect comparison low certainty evidence.

In this example, both direct comparisons are the output of a meta-analysis. However, if one of the direct comparisons only had a single trial, then the effect estimate from that single trial (rather than a meta-analysis) would be used. Many trials report a p value instead of a 95% CI and so we have included a calculator to derive CIs from p values (online supplemental material-statics worksheet) along with other resources for calculating effect estimates and CIs from other data provided in trial reports.<sup>9 10</sup>

If the data are taken from different sources, it is critical to ensure that the direction of the comparison is the same for both reported comparisons. Continuous effect measures are typically calculated as intervention minus control, and ratio statistics as intervention divided by control, but these may be calculated the



**Figure 2** Case study demonstrating the use of Excel tool to perform indirect comparisons. Treatments, effect estimates and 95% CIs are taken from standard pairwise meta-analysis (panel A), and entered into the white cells of the Excel tool (panel B). The results are displayed in the yellow cells.

other way around. For continuous statistics, the effect estimate for B versus A is the negative of the effect of A versus B. For ratio statistics, the effect estimate for B versus A is the inverse of the effect of A versus B. A worked example where the two treatment effects are calculated in different directions is provided within the Excel tool (online supplemental material; worked example 2 'inverted').

Note that the level of precision for the indirect comparison (represented by the width of the CI) is less than for the direct comparisons. Despite the high-certainty evidence that cold machine perfusion is superior to standard ice-box storage, while warm machine perfusion appears no better than equivalent, we still cannot be certain that cold machine perfusion is superior to warm.

Statistically, the Bucher method is equivalent to a test for interaction, often used as a test for subgroup differences in meta-analysis. This can be seen from the identical p values reported in the spreadsheet for the indirect comparison and the test for interaction reported on the forest plot. The test for interaction has very low power compared with main effects and so requires much larger sample sizes to draw reliable conclusions.<sup>2 11</sup> Effect estimates from meta-analyses in systematic reviews will have larger sample sizes than individual trials and provide a less-biased summary of the evidence than any single trial, and so should always be preferred if available.

### More complex networks and NMA

For the simple network displayed in figure 1, the techniques described above are effective and are recommended in the Cochrane Handbook (chapter 11).<sup>12</sup> However, for more complex networks of treatments (those with more than two comparisons with direct evidence), more complicated statistical methods, collectively referred to as NMA, should be applied.

Such methods are seeing a surge in use, with methods implemented in Stata, a number of R packages (such as GeMTC, BUGSnet and netmeta), and user-friendly open-source applications such as MetaInsight.<sup>13-17</sup> These techniques are recommended for more complex networks.<sup>15</sup> For the simple network in figure 1, frequentist NMA will generate identical results to the Bucher technique we describe (the techniques are equivalent for this network). This is demonstrated for our case study in online supplemental figure 1. We also performed Bayesian NMA (BUGSnet) using the case study example, with the default uninformative priors. As expected, the central estimate is identical to our analysis and the credible intervals were a little bit wider than the frequentist CIs.

### Conclusions

In contrast to NMA methods, the techniques of Bucher *et al* are less frequently used. However, for the network shown in figure 1, Bucher indirect comparisons are recommended by Cochrane and generate identical results to frequentist NMA.<sup>12</sup> Such networks occur frequently, where two new treatments or procedures are developed and assessed against placebo or standard of care, rather than each other, or where two different treatments are considered suitable options by clinicians and patients but the potential benefits and harms need to be weighed up carefully. Previous reports of Bucher's methods have been written for an audience of statisticians; our aim here is to make the methods more accessible.

The Excel spreadsheet which we provide here is easy to use and should facilitate the practical application of these techniques. We envisage this tool being used by researchers performing

meta-analyses where a network analogous to figure 1 is found. This should be performed under the guidance of someone with meta-analysis expertise, preferably a statistician.

X Samuel James Tingle @SamJTingle and Rachel Richardson @RacheRichardson

**Acknowledgements** Thank you to Kim Pearce (Newcastle University) for their statistical support which laid the initial foundations for this project. Without this, the published tool would likely never have been created.

**Contributors** SJT conceived the paper. SJT and JS created the first version of the excel tool which was then adapted following testing and suggestions from GK, FL, CW and RR. SJT and JS drafted the first version of the manuscript which was changed following critical review from all other authors. All authors reviewed and approved the final version. JS is the guarantor of this work.

**Funding** SJT worked on this project during an MRC Clinical Research Training Fellowship (MR/Y000676/1) at Newcastle University. This study was supported by the National Institute for Health and Care Research (NIHR) Blood and Transplant Research Unit in Organ Donation and Transplantation (NIHR203332), a partnership between NHS Blood and Transplant, University of Cambridge and Newcastle University. The views expressed are those of the author(s) and not necessarily those of the NIHR, NHS Blood and Transplant or the Department of Health and Social Care.

**Competing interests** All authors have completed the ICMJE uniform disclosure form at <http://www.icmje.org/disclosure-of-interest/> and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; internally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iDs

Samuel James Tingle <http://orcid.org/0000-0001-5529-7815>

Rachel Richardson <http://orcid.org/0000-0003-2848-6260>

## References

- 1 Bucher HC, Guyatt GH, Griffith LE, *et al.* The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50:683–91.
- 2 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.
- 3 Donegan S, Williamson P, D'Alessandro U, *et al.* Assessing key assumptions of network meta-analysis: a review of methods. *Res Synth Methods* 2013;4:291–323.
- 4 Guyatt G, Oxman AD, Akl EA, *et al.* GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
- 5 Nikolakopoulou A, Higgins JPT, Papakonstantinou T, *et al.* CINEMA: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Med* 2020;17:e1003082.
- 6 Guyatt GH, Oxman AD, Kunz R, *et al.* GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol* 2011;64:1294–302.
- 7 Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med* 1999;18:2693–708.
- 8 Tingle SJ, Dobbins JJ, Thompson ER, *et al.* Machine perfusion in liver transplantation. *Cochrane Database Syst Rev* 2023;9:CD014685.
- 9 Tierney JF, Stewart LA, Ghersi D, *et al.* Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials* 2007;8:16.
- 10 MedCalc. MedCalc - free statistical calculators: MedCalc. 2023. Available: [www.medcalc.org/calc/index.php](http://www.medcalc.org/calc/index.php)
- 11 Sommet N, Weissman DL, Cheutin N, *et al.* How many participants do I need to test an interaction? Conducting an appropriate power analysis and achieving sufficient power to detect an interaction. *Adv Meth Pract Psychol Sci* 2023;6:25152459231178728.
- 12 Chaimani A, Caldwell D, Li T, *et al.* Chapter 11: undertaking network meta-analyses. In: Higgins JT, Chandler J, Cumpston M, *et al.*, eds. *Cochrane handbook for systematic reviews of interventions version 64*. Cochrane, 2023. Available: [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook)
- 13 Béliveau A, Boyne DJ, Slater J, *et al.* BUGSnet: an R package to facilitate the conduct and reporting of bayesian network meta-analyses. *BMC Med Res Methodol* 2019;19:196.
- 14 Owen RK, Bradbury N, Xin Y, *et al.* MetaInsight: an interactive web-based tool for analyzing, interrogating, and visualizing network meta-analyses using R-shiny and netmeta. *Res Synth Methods* 2019;10:569–81.
- 15 Rouse B, Chaimani A, Li T. Network meta-analysis: an introduction for clinicians. *Intern Emerg Med* 2017;12:103–11.
- 16 Shim SR, Kim S-J, Lee J, *et al.* Network meta-analysis: application and practice using R software. *Epidemiol Health* 2019;41:e2019013.
- 17 Owen RK, Bradbury N, Xin Y, *et al.* MetaInsight web app. 2024. Available: [crsu.shinyapps.io/MetaInsight/](https://crsu.shinyapps.io/MetaInsight/)