ISSUE NINE August 2017

# Evidence Notes

Bridge Medical
evidence that matters

Welcome to the 9th issue of Evidence Notes. This article is a status update on the role of Network Meta-Analyses (NMAs) in the context of Health Technology Assessment (HTA) submissions, with a specific focus on critiques by HTA bodies on manufacturer NMA submissions. As with all our Evidence Notes, the aim is to provide a brief readable summary, rather than a lot of technical information. References are provided for further information.

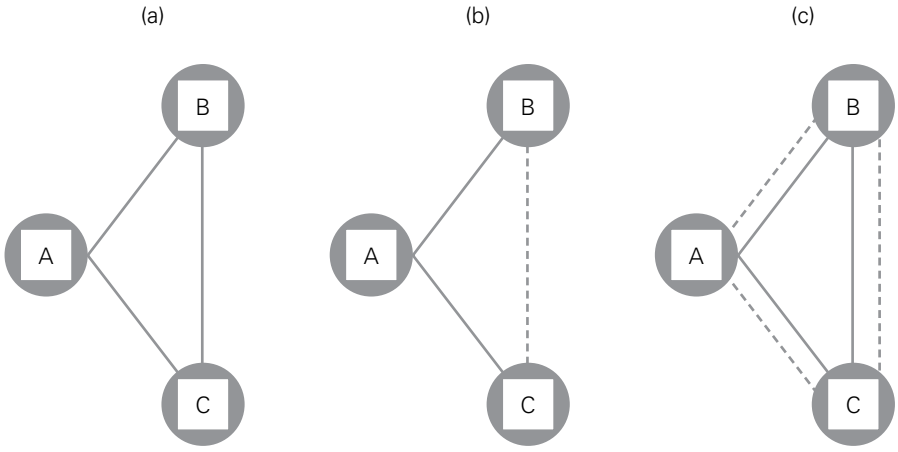
## The Role of Network Meta-Analysis in Health Technology Assessment

The need for robust comparative effectiveness data to enable good decision making by Health Technology Assessment (HTA) bodies, for example, is driving new methodologies for evidence synthesis such as network meta-analysis (NMA), sometimes called multiple or mixed treatment comparison or meta-analysis.

Traditionally, data from multiple studies (usually randomised controlled trials [RCTs]) have been synthesised using standard pairwise meta-analyses of studies which directly compare two interventions. In practice, new treatments are rarely compared against all available marketed therapies, and it is even less likely that they would be compared with *all* relevant treatments together in one large scale, high-quality study. Therefore, it is becoming ever more important to provide decision makers (e.g. clinicians, HTA bodies and reimbursement authorities) with analyses that compare the potential plethora of treatment options available for many conditions.

NMA enables the *simultaneous* analysis of all available comparative evidence to provide estimates of the relative treatment effects between competing interventions. It allows treatments to be compared where direct comparisons do not exist, and combines evidence from direct and indirect comparisons where both exist.[1–4] Instances where direct and indirect evidence are consistent will lead to an increase in precision, while inconsistencies can lead to a better understanding of heterogeneity. Overall, this approach allows decision makers to reach conclusions by comparing all relevant comparators based on a statistical synthesis of all relevant evidence. For example, in its simplest form, if there are studies that compare A vs B and A vs C but not B vs C, NMA can be used to make an indirect estimate of the latter (see **Figure 1b**, dotted line).[5–7]

*Figure 1:* **Direct and Indirect Comparisons in NMA: A Simple Network**




Circles or nodes represent different interventions (A, B or C). Solid lines represent direct comparison(s) in ≥ 1 study per comparison; dotted lines represent an indirect comparison (IC; e.g. where there are no studies that directly compare the two interventions): (a) all interventions have been directly compared; (b) A vs B and A vs C have been directly compared and B vs C can be indirectly estimated based on the direct comparisons; (c) even in the presence of a direct comparisons, indirect estimates can be used to supplement the overall data for all comparisons (e.g. an indirect estimate of B vs C is obtained from direct comparison of B vs A and C vs A; an indirect estimate of A vs B is obtained from direct comparison of A vs C and B vs C; and so on)

Graphical illustrations of more complex networks with more treatments than shown in **Figure 1** can be seen in various guidelines or technical support documents (e.g. European Network for Health Technology Assessment [EUNetHTA];[8] National Institute for Health and Care Excellence [NICE])[9] and in the literature.[10–12]

Based on these benefits, NMA is increasingly being used to synthesise data. From 1997 when the first examples were reported up to March 2015, there were 456 published NMAs comparing ≥4 interventions.[13] The approach is also increasingly being accepted around the world by healthcare decision makers,[4] legal frameworks and in the development of treatment guidelines.[14] Indeed, some have argued that NMA should be the highest level of evidence in treatment guidelines.[15]

There are, however, several methodological issues and challenges. A summary of strengths and limitations specific to NMA is provided in **Table 1**.

Given the challenges described in **Table 1**, a number of tools have been developed in order to assess the degree of confidence or certainty in NMA data. Some of the main tools are summarised in **Table 2**.

ISPOR has also published best practice guidelines on the conduct of NMA, but there are currently no widely accepted international guidelines.[4,23]

Reviews of the various HTA guidance in individual jurisdictions have been conducted (with cut-off dates up to mid-2013),[24–28] and **Table 3** is an amalgamated summary of key findings. The key take home message from these reviews was that whilst head-to-head RCTs are preferred, use of ICs was recommended in their absence by many HTA agencies.

*Table 1:* **Strengths and Limitations of NMA[4,5,9,15–22]**

| STRENGTHS | LIMITATIONS |
|---|---|
| Provides simultaneous comparison of all treatments in network of studies; overcomes issue of comparing separate (sometimes conflicting) pairwise meta-analyses | Analysis is "observational" – treatment comparisons are not randomised *across trials* |
| Can indirectly compare two treatments that have not been directly compared in an RCT | Techniques not yet fully established and are mathematically complex [THERE IS NOW A LARGE BODY OF METHODOLOGICAL WORK] |
| Inclusion of indirect evidence may increase power and precision of overall estimate | Effects of heterogeneity; e.g. potential differences in patient characteristics at baseline; the more studies in the network the greater the risk of bias; results can also be distorted by small trials |
| Possible to include RWE studies which may increase the precision of the overall estimate | Inclusion of non-randomised RWE studies may create lack of balance in EMs across treatments (although there are methods to adjust for this) |

EM = Effect modifier; IC = Indirect comparison; RCT = Randomised controlled trial; RWE = Real-world evidence

*Table 2:* **Tools for the Assessment of NMA Quality**

| TOOL | DESCRIPTION |
|---|---|
| GRADE | Rates quality of evidence in direct, indirect, and NMA estimates. Widely used in guideline development (e.g. WHO, Cochrane etc.) |
| Modified GRADE | Modification to GRADE to draw a distinction between: a) effect sizes for pairwise comparisons of treatments and b) a ranking of treatments |
| ISPOR questionnaire | Consensus-based 26-item questionnaire to help healthcare decision makers assess the relevance and credibility of NMA. Used widely in guideline development |
| PRISMA extension | Modified, 32-item PRISMA extension checklist developed to address reporting of NMA |
| ROBINS-I | Evaluates risk of bias in estimates of comparative effectiveness from non-randomised studies. Particularly useful in NMAs that include non-randomised studies |

GRADE = Grading of Recommendations Assessment, Development and Evaluation; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NMA = Network meta-analyses; PRISMA = Preferred Reporting Items for Systematic Reviews and Meat-Analysis; ROBINS = Risk of Bias in Non-Randomised Studies – of Interventions; WHO = World Health Organisation

*Table 3:* **Summary of Reviews of HTA Guidance in Individual Jurisdictions[24–28]**

| ISSUE | HTA |
|---|---|
| Is the guidance comprehensive? | • Generally not comprehensive<br>• Some recommendations on the conduct of NMA provided by Australia, Belgium, Canada, France, Germany, Scotland, Spain, South Africa, and England and Wales<br>• Others (Ireland, Norway, Poland, Sweden, and US) did not provide any detailed guidance |
| Does the guidance recommend use of ICs in absence of head-to-head studies? | • Recommended by many HTA agencies (e.g. NICE and in Canada, France, Australia etc.)<br>• All reviewed agencies except Turkey confirmed use for "rapid assessments"<br>• Use of IC data sometimes limited to pharmacoeconomic evaluations (e.g. France, Germany) |
| What about use of ICs when head-to-head evidence *is* available? | • No consensus<br>• NICE: if technologies have not been compared in a single RCT, NMA can be used alongside data from a series of pairwise head-to-head RCTs for each treatment comparison of interest |
| Is IC in their hierarchy of evidence? | • Few agencies specifically list IC in their hierarchy of evidence<br>• Australia include IC of randomised data as second tier and a comparison across non-randomised data as third tier |
| Does the guidance recommend a particular NMA methodology? | • Do not typically describe or recommend specific NMA methodology (only ISPOR addresses detailed NMA methods) |
| Do they exclude the use of non-randomised studies? | • Canadian guidelines specifically exclude the use of non-randomised studies from ICs |
| What does their guidance say about use of NMA for safety data? | • Data analysis and synthesis of safety data was discussed by Canada and Australia, with only Australia encouraging extended assessment beyond direct RCTs |

HTA = Health Technology Assessment; IC = Indirect comparison; ISPOR = International Society for Pharmacoeconomics and Outcomes Research; NICE = National Institute for Health and Care Excellence; NMA = Network meta-analyses; RCT = Randomised controlled trial; US = United States

Similarities across agencies were greater than differences, with no obvious issue that would prevent submission of one NMA to multiple jurisdictions.[26]

Given that HTA guidance on NMAs methodology is relatively limited, it would be informative to know the feedback received during HTA assessment of different technologies. Two recent abstracts[29,30] have investigated feedback from the NICE Evidence Review Group (ERG) following Single Technology Appraisal (STA), and a summary is provided in **Table 4**.

In addition to differing expectations amongst some ERGs, the main recurring findings in the manufacturers' submissions were poor justification of key decisions (e.g. trial selection), lack of transparency and insufficient or inadequate detail (e.g. assessment of heterogeneity, inconsistency, risk and impact of bias, justification for methodology); deficiencies broadly consistent with the findings of those who have reviewed the quality of published NMAs.[31]

NMA methodology is a dynamic field, and one emerging areas is whether the network of studies can be extended to include "real-world" evidence (RWE). Up to the end of 2012, less than 4% of published NMAs included a non-

*Table 4:* **NICE Evidence Review Group (ERG) Feedback Following Single Technology Appraisal (STA)[29,30,32]**

| Author | STA Dates; No. of STAs Reviewed | No. of STAs with ICs | Author's Findings | ERG Feedback | Future Recommendations |
|---|---|---|---|---|---|
| Fleetwood et al 2016 | STAs (May 2015 – end Apr 2016); n=39<br><br>NB: Rapid reviews and terminated appraisals excluded | 25 (64%) | • NMAs widely used but do not always conform to NICE guidelines<br>• Manufacturers used standard as well as more complex methods<br>• Manufacturers did not consistently report all the details suggested by NICE guidelines<br>• 28% of STAs did not assess source of heterogeneity<br>• <50% of NMAs assessed risk of bias and only one explored its *impact*<br>• Bayesian approaches included in majority of relevant STAs<br>• ERGs differed in their NMA expectations (some accepted more complex methods) | • Poor justification of key decisions (e.g. why specific trials were included/excluded)<br>• Lack of transparency (e.g. study selection process, input datasets for each network; NMAs should be reproducible)<br>• Insufficient detail (e.g. about studies included in the NMA; methodological details for Bayesian NMAs) | • Use the NICE DSU TSD 7 checklist to review NMA<br>• Ensure key assumptions and decisions clearly justified<br>• Ensure NMA methods are transparent and reproducible |
| Sarri et al 2016 | STAs (Dec 2015/Jan 2016 - Jun 2016); n=22 | Not clear – possibly 16 (73%): | Heterogeneity:<br>• Heterogeneity is a key issue in NMAs of mixed-populations<br>• Only partially explicit NICE guidance; no clear guidance on a suitable proportion of target patients in mixed-population studies, nor on the acceptable level of heterogeneity in analyses<br>Selection of NMA type:<br>• Selection of different types of NMA not clearly justified<br>Analyses:<br>• Submissions did not always follow the NICE guidance on non-proportional hazards (the proportional-hazard assumption is that hazard ratios for a given outcome remain constant over time) | Heterogeneity:<br>• A common ERG criticism was that heterogeneity (e.g. due to differences in disease stages/severity) was either inadequately assessed or not explored in subgroup analyses in NMAs of mixed-population trials<br>• In a few cases, ERG agreed that studies were too heterogeneous to be synthesised in an NMA<br>Selection of NMA type:<br>• ERGs noted that distributional assumptions not always tested before the selection of NMA type to be used in submission<br>Analyses:<br>• Manufacturers were criticised by the ERGs for not taking steps to control for non-proportional hazards | • Apply rigorous thresholds for the proportion of **target** patients in mixed-population studies (e.g. those in NICE Clinical Guidelines)<br>• Clearly specify sources of heterogeneity and, if present, use appropriate analysis models<br>• Use subgroup analysis<br>• Use sensitivity analysis to test validity of NMA<br>• Explore direct and indirect evidence inconsistencies. If present, explain and resolve |

DSU = Decision Support Unit; ERG = Evidence Review Group; NICE = National Institute for Health and Care Excellence; NMA = Network Meta-Analysis; RCT = randomised controlled trial; STA = Single Technology Appraisals; TSD = Technical Support Document

randomised study.[19,33] NICE, however, is reported to be moving further towards the use of RWE in its decision making and accepts it is important to "assemble all relevant evidence" in a way that "minimises the risk of biased selection".[20,33–35] NICE guidance further states that inferences from non-randomised studies (or those without controls) will be treated in a "more circumspect" manner than those from RCTs and that biases should be identified, quantified and adjusted for, and should be subject to sensitivity analyses.[33] EUNetHTA guidelines also highlight the importance of using "all relevant studies, including observational and unpublished data".[8]

The strengths and limitations of different methodologies for the inclusion of RWE data from non-randomised studies in NMA have very recently been subject to thorough evaluation. The authors concluded that RWE inclusion has the potential to corroborate findings from RCTs and increase the precision of the estimate, thereby enhancing the decision-making process.[36] However, where there is a very high risk of bias or the study is incompatible with the specific aims of the NMA, exclusion is recommended.[19]

In summary, the use of NMA is likely to increase alongside the need to estimate the relative effectiveness of a plethora of treatment options within a single analysis using direct and/or indirect evidence. The approach is critical to decision makers such as HTA agencies as well as to the development of clinical guidelines. HTA guidance is currently not comprehensive and, as a consequence, can differ both across and within jurisdictions. Although NICE ERGs can differ in their NMA expectations, commonly recurring issues and a number of mitigation approaches are provided which may help minimise these issues for future submissions. Finally, the field of NMA methodology is dynamic and the incorporation of RWE into NMA is just one area that is likely to develop in the medium term.

Further information on this subject can be found in the references provided.

**Martin Jones and Paul Gandhi;**
Corresponding author:
martinjones@bridgemedical.org

## 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 Jun;50(6):683-91. https://www.ncbi.nlm.nih.gov/pubmed/9250266
2. Lumley T. Network meta-analysis for indirect treatment comparisons. Stat Med. 2002 Aug 30;21(16):2313-24. https://www.ncbi.nlm.nih.gov/pubmed/12210616
3. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105-24. https://www.ncbi.nlm.nih.gov/pubmed?cmd=historysearch&querykey=19
4. 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 Jun;14(4):417-28. https://www.ncbi.nlm.nih.gov/pubmed/21669366
5. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ. 2005 Oct 15;331(7521):897-900. https://www.ncbi.nlm.nih.gov/pubmed?cmd=historysearch&querykey=5
6. Salanti G, Higgins JP, Ades AE, et al. Evaluation of networks of randomized trials. Stat Methods Med Res. 2008 Jun;17(3):279-301. Epub 2007 Oct 9. https://www.ncbi.nlm.nih.gov/pubmed/17925316
7. 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 Jun;3(2):80-97. https://www.ncbi.nlm.nih.gov/pubmed/26062083
8. European Network for Health Technology Assessment (EUNetHTA). Guideline: Comparators & Comparisons – Direct and Indirect Comparisons. Feb 2013. http://www.eunethta.eu/sites/5026.fedimbo.belgium.be/files/Direct%20and%20indirect%20comparisons.pdf
9. Dias S, Welton NJ, Sutton A, et al. NICE Decision Support Unit technical support document 1: Introduction to Evidence Synthesis for Decision Making. April 2012 http://www.nicedsu.org.uk/TSD1%20Introduction.final.08.05.12.pdf
10. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009 Feb 28;373(9665):746-58. https://www.ncbi.nlm.nih.gov/pubmed/19185342
11. 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 Sep 14;382(9896):951-62. https://www.ncbi.nlm.nih.gov/pubmed/23810019
12. Cipriani A, Barbui C, Salanti G, et al. Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. Lancet. 2011 Oct 8;378(9799):1306-15. https://www.ncbi.nlm.nih.gov/pubmed/21851976
13. Petropoulou M, Nikolakopoulou A, Veroniki AA, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. J Clin Epidemiol. 2017 Feb;82:20-28. https://www.ncbi.nlm.nih.gov/pubmed/27864068
14. Kanters S, Ford N, Druyts E, et al. Use of network meta-analysis in clinical guidelines. Bull World Health Organ. 2016 Oct 1;94(10):782-784. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5043215/pdf/BLT.16.174326.pdf
15. Leucht S, Chaimani A, Cipriani AS, et al. Network meta-analyses should be the highest level of evidence in treatment guidelines. Eur Arch Psychiatry Clin Neurosci. 2016 Sep;266(6):477-80. https://www.ncbi.nlm.nih.gov/pubmed/27435721
16. Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. BMJ. 2013 May 14;346:f2914. https://www.ncbi.nlm.nih.gov/pubmed/23674332
17. Efthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods. 2016 Sep;7(3):236-63. https://www.ncbi.nlm.nih.gov/pubmed/26754852
18. Cipriani A, Higgins JP, Geddes JR, et al. Conceptual and technical challenges in network meta-analysis. Ann Intern Med. 2013 Jul 16;159(2):130-7. https://www.ncbi.nlm.nih.gov/pubmed/23856683; http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.689.7412&rep=rep1&type=pdf
19. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. Stat Med. 2017 Jan 12. doi: 10.1002/sim.7223. [Epub ahead of print] https://www.ncbi.nlm.nih.gov/pubmed/28083901
20. Faria R, Hernandez AM, Manca, A, et al. NICE DSU Technical Support Document 17: The use of observational data to inform estimates of treatment effectiveness for Technology Appraisal: Methods for comparative individual patient data. 2015. Available from http://www.nicedsu.org.uk; http://www.nicedsu.org.uk/TSD17%20-%20DSU%20Observational%20data%20FINAL.pdf
21. Mavridis D, Giannatsi M, Cipriani A, et al. A primer on network meta-analysis with emphasis on mental health. Evid Based Ment Health. 2015 May;18(2):40-6. https://www.ncbi.nlm.nih.gov/pubmed/25908686
22. 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 Med. 2013 Jul 4;11:159. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3707819/pdf/1741-7015-11-159.pdf
23. Hoaglin DC, Hawkins N, Jansen JP, 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 Health. 2011 Jun;14(4):429-37. https://www.ncbi.nlm.nih.gov/pubmed/21669367
24. Coleman CI, Phung OJ, Cappelleri JC, et al. Use of mixed treatment comparisons in systematic reviews. AHRQ Publication No. 12-EHC119-EF. Agency for Healthcare Research and Quality, 2012 https://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0076753/pdf/PubMedHealth_PMH0076753.pdf
25. Dequen P. Clinical evidential and methodological challenges of early assessments of new health technologies. PhD Thesis. 2016. https://lra.le.ac.uk/bitstream/2381/37930/1/2016DEQUENPPhD.pdf
26. Laws A, Kendall R, Hawkins N. A comparison of national guidelines for network meta-analysis. Value Health. 2014 Jul;17(5):642-54. http://www.valueinhealthjournal.com/article/S1098-3015(14)01884-1/pdf
27. Kleijnen S, George E, Goulden S, et al. Relative effectiveness assessment of pharmaceuticals: similarities and differences in 29 jurisdictions. Value Health. 2012 Sep-Oct;15(6):954-60. https://www.ncbi.nlm.nih.gov/pubmed?term=2012%5Bpdat%5D+AND+Kleijnen+S%5Bfirst+author%5D&cmd=detailssearch
28. Wells GA, Sultan SA, Chen L, et al. Indirect Evidence: Indirect Treatment Comparisons in Meta-Analysis. Ottawa: Canadian Agency for Drugs and Technologies in Health; 2009. https://www.cadth.ca/media/pdf/H0462_itc_tr_e.pdf
29. Fleetwood K, Glanville J, McCool R, et al. A Review of the Use of Network Meta-Analysis In Nice Single Technology Appraisals. Value in Health. 2016 Nov 1;19(7):A348. http://www.valueinhealthjournal.com/article/S1098-3015(16)31376-6/abstract; http://www.yhec.co.uk/yhec-content/uploads/2016/11/ISPOR_Review-of-NMAs-in-NICE-STAs.pdf
30. Sarri G, Rizzo M. Systematic Review and Network Meta-Analysis (NMA) in Reimbursement Submission; What Nice Says Versus what Nice Wants. Value in Health. 2016 Nov 1;19(7):A493. http://www.valueinhealthjournal.com/article/S1098-3015(16)32211-2/fulltext?rss=yes; https://www.ispor.org/research_pdfs/54/pdffiles/PHP302.pdf
31. Hutton B, Salanti G, Chaimani A, et al. The quality of reporting methods and results in network meta-analyses: an overview of reviews and suggestions for improvement. PLoS One. 2014 Mar 26;9(3):e92508. http://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0092508&type=printable
32. Dias s, Welton NJ, Sutton A, et al. NICE Decision Support Unit technical support document 1: Introduction to Evidence Synthesis for Decision Making. April 2012 http://www.nicedsu.org.uk/TSD1%20Introduction.final.08.05.12.pdf
33. Bell H, Wailoo AJ, Herandez M et al. NICE Decision Support Unit. The use of real world data for the estimation of treatment effects in NICE decision making. 2016 http://www.nicedsu.org.uk/RWD%20DSU%20REPORT%20Updated%20DECEMBER%202016.pdf
34. Sutton A, Ades AE, Cooper N, et al. Use of indirect and mixed treatment comparisons for technology assessment. Pharmacoeconomics. 2008;26(9):753-67. https://www.ncbi.nlm.nih.gov/pubmed/18767896
35. NICE. Guide to the methods of technology appraisal. 2013. www.nice.org.uk/process/pmg9; https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781
36. Cameron C, Fireman B, Hutton B, et al. Network meta-analysis incorporating randomized controlled trials and non-randomized comparative cohort studies for assessing the safety and effectiveness of medical treatments: challenges and opportunities. Syst Rev. 2015 Nov 5;4:147. https://www.ncbi.nlm.nih.gov/pubmed/26537988