Check for updates

EDITORIAL
A clinician's guide to network meta-analysis

THE EVOLUTION OF EVIDENCE SYNTHESIS

Increasing interest in promoting evidence-based clinical practice has led to methodological advancements in evidence syntheses [1, 2]. Narrative reviews have been superseded by systematic reviews, which may include meta-analysis—statistical pooling of treatment effect estimates across similar trials to improve precision [3–5]. Systematic reviews minimize the risk of selection bias by considering all evidence relevant to a clinical question; however, an important limitation of conventional meta-analyses is that they only inform treatments that have been directly compared in clinical trials. Moreover, many trials compare active interventions against placebo, usual, or standard care, whereas patients and clinicians are typically concerned with the relative effectiveness of competing interventions. Network meta-analysis (NMA) has emerged to address these limitations by allowing for calculation of the comparative effects of more than two competing interventions, even when they have not been directly compared in clinical trials [6, 7].

WHAT IS NETWORK META-ANALYSIS?

NMA requires the same steps as a conventional meta-analysis which include a systematic search of the literature, assessment of risk of bias among eligible trials, statistical pooling of reported pairwise comparisons for all outcomes of interest, and assessment of the overall certainty of evidence on an outcome-by-outcome basis. This provides the "direct" evidence for treatments that have been compared against each other, which is graphically represented by a network map. An NMA then identifies all interventions that are connected by virtue of a common comparator. For example, two different active treatments may have been compared against placebo in different trials. An NMA allows for a theoretical trial to be created that compares these active treatments against each other, based on their effect against a common comparator (placebo), which provides "indirect" evidence. Indirect comparisons provide an opportunity to fill knowledge gaps within the available evidence, providing a more comprehensive understanding of treatment options for the clinician. The network estimate is the pooled result of the direct and indirect evidence for a given comparison, or only the indirect evidence if no direct evidence is available [6, 8, 9]. Once all treatments have been compared within a network, there are different methods for ranking treatments to convey their relative net effectiveness. Limitations and advancements in the ranking methodology will be discussed in greater detail within the example provided below.

NETWORK META-ANALYSIS IN PRACTICE

An example network map on first-line medications effects on intra-ocular pressure (IOP) for primary open angle glaucoma (POAG) is shown in Fig. 1, which represents all pharmacologic treatments that have been directly evaluated in 114 clinical trials for this condition [10]. Traditional meta-analysis would be limited in comparing two of these treatments at a time, and could not inform the effectiveness of treatments that have not been directly compared; however, this NMA provides the relative effectiveness of all 15 treatments in a single investigation, even when no RCT is available to make a direct comparison between two treatments. The network map uses circles, or nodes, for each included treatment, that increase in size relative to the number of patients treated with that medication within included RCTs. The lines connecting different treatments are weighted by the number of RCTs comparing them (i.e., thicker lines convey more direct trials) [10]. In this particular study, the authors color coded their treatment nodes by drug class to improve interpretation. The network is specific to one outcome, in this case IOP, and the network assumes that the baseline characteristics of patients enrolled across trials are similar.

As Fig. 1 demonstrates, there are many RCTs assessing pharmacotherapy for POAG. Some treatments, such as Timolol or Latanoprost, have large bodies of evidence, while many others have far fewer – and smaller – trials assessing their efficacy [10]. This network enables the comparison of 14 active medications, as well as placebo, for POAG.

While the ability to summarize large bodies of evidence is also possible for traditional meta-analyses, NMAs provide comparative effectiveness data between competing treatments. It is important to note that the evidence provided by an NMA is subject to the limitations of the individual RCTs included within the network [11]. In addition, the ranking of interventions by NMAs using methods such as the Surface Under the Cumulative Ranking Curve (SUCRA) approach is problematic – despite this currently being the most common form of treatment ranking in NMAs. This approach ranks all treatments within a network from "best" to "worst" for each analyzed outcome, but only considers the effect estimate and not the associated precision or the certainty of evidence [12]. Thus, interventions supported by small, low-quality trials that report large effects are ranked highly. Minimally or partially contextualized approaches, instead, consider the magnitude of effect in the context of patient importance as well as the certainty of evidence [13, 14].

HOW CAN YOU HAVE CERTAINTY IN THE FINDINGS OF AN NMA?

Like all study designs, there are considerations when evaluating the credibility of the findings of an NMA. These include the same issues that should be considered when evaluating a traditional pairwise meta-analysis, such as the rigor of the literature search, risk of bias among included trials, consistency of effect estimates contributing to pooled effects (heterogeneity), precision of the pooled effect estimate, publication bias, and directness of the included evidence in relation to the primary research question [8, 9, 15, 16]. However, there are two additional considerations
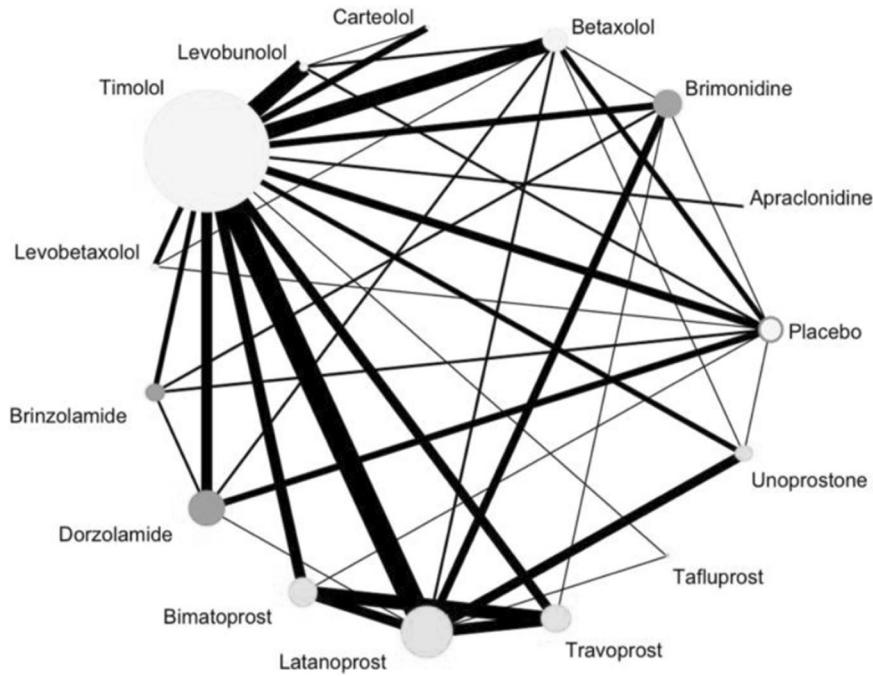
Fig. 1 Network diagram from Li et al. (2016) [10] comparing medications for POAG. Size of nodes represents the number of patients, line thickness represents number of trials.

that are specific to NMAs: incoherence and transitivity [8, 9, 15, 17].

Incoherence exists when the direct and indirect estimates for a comparison are not consistent with one another [6]. A meta-epidemiological study of 112 published NMAs found inconsistent direct and indirect treatment effects in 14% of the comparisons made [18]. This means that while in most cases it is appropriate to combine indirect and direct evidence, this is not always the case, and review authors should formally explore this issue. In the presence of incoherence, the higher certainty evidence should be presented rather than the network estimate. If the direct and indirect effects are both supported by the same certainty of evidence, then the network estimate can be used but should be downgraded one level for incoherence. The GRADE approach is increasingly used for rating the certainty in evidence for network estimates, which incorporates these aforementioned criteria [11, 15–17]. A GRADE rating can assign high, moderate, low, or very low certainty in the evidence [11, 15–17]. Clinicians should take the certainty of the evidence in consideration when determining the impact findings would have on their clinical practice, as lower certainty evidence provides less confidence in the results.

Transitivity refers to the similarity between study characteristics that allows indirect effect comparisons to be made with the assurance that there are limited factors that could modify treatment effects, aside from the intervention under investigation [6, 15]. Essentially, transitivity refers to the inclusion of studies that fundamentally address the same research questions within the same population [6]. Intransitivity can result in biased indirect estimates, which would then impact the overall findings of the network estimates [15, 17]. As previously discussed, incoherence exists when discrepancies between direct and indirect estimates are present, thus, transitivity is a common cause of incoherence [17].

Clinicians cannot be expected to evaluate transitivity and incoherence within an NMA and authors should clearly report on these two important aspects. Indeed, the absence of reporting should lead readers to question the findings. Table 1 provides an example and overview of the core items for readers to identify for critical appraisal of published NMAs, as applied to the Li et al. (2016) POAG study [10, 19]. These criteria are based on the Users' Guides to the Medical Literature: Essentials of Evidence-Based Clinical Practice [19].

CONCLUSION

Rigorously conducted and reported NMA may provide helpful information for advancing evidence-based ophthalmology, specifically in the common scenario in which multiple treatment options exist. However, clinicians should appraise the quality of NMAs before accepting the results, and even rigorously conducted NMAs cannot provide high certainty evidence if the primary trials eligible for review are flawed.

Mark R. Phillips[ID][1], David H. Steel[2,3], Charles C. Wykoff[4,5], Jason W. Busse[1,6,7], Raveendhara R. Bannuru[ID][8], Lehana Thabane[ID][1,9], Mohit Bhandari[ID][1,10], Varun Chaudhary[ID][1,10✉] and for the Retina Evidence Trials InterNational Alliance (R.E.T.I.N.A.) Study Group*
[1]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada.* [2]*Sunderland Eye Infirmary, Sunderland, UK.* [3]*Biosciences Institute, Newcastle University, Newcastle Upon Tyne, UK.* [4]*Retina Consultants of Texas (Retina Consultants of America), Houston, TX, USA.* [5]*Blanton Eye Institute, Houston Methodist Hospital, Houston, TX, USA.* [6]*Michael G. DeGroote National Pain Center, McMaster University, Hamilton, ON, Canada.* [7]*Department of Anesthesia, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada.* [8]*Center for Treatment Comparison and Integrative Analysis, Division of Rheumatology, Tufts Medical Center, Boston, MA, USA.* [9]*Biostatistics Unit, St. Joseph's Healthcare-Hamilton, Hamilton, ON, Canada.* [10]*Department of Surgery, McMaster University, Hamilton, ON, Canada. *A list of authors and their affiliations appears at the end of the paper.*
✉*email: vchaudh@mcmaster.ca*

Table 1. Example appraisal of the Li et al. (2016) POAG NMA.

Item	Considerations	Example from Li et al. (2016) POAG NMA [10]
Systematic review processes		
Study eligibility	Systematic review with clearly defined, explicit eligibility criteria should be included. A pre-published protocol should summarize the planned conduct of the NMA (ex. PROSPERO registration).	The methods have the section "Eligibility criteria for considering studies for this review", which provides explanation of their eligibility criteria.
Literature search	Reproduceable, systematic search that has confidently retrieved all relevant literature on the topic.	Within the methods, "Search methods for identifying studies" summarizes the systematic approach used for the literature search.
Study selection and assessment	Systematic screening process is used to identify and select all relevant literature from the search that was conducted. Additionally, a risk of bias should be conducted for each included study.	The "Study selection" subheading summarizes the study screening process.
Between study differences	Did the study plan for subgroup, sensitivity, and/or meta-regression analyses to address hypotheses on between study differences?	Sensitivity analyses that evaluated specific concentrations of Bimatoprost and Timolol were conducted. Justification for this is provided in the section "Measures of association".
NMA Analysis and Results		
Amount of evidence	For each comparison, consider the number of RCTs informing the effect estimate. Estimates based on a small amount of evidence may be less reliable than those informed by a large number of RCTs.	Table 1 provides an overview of the number of studies included within each comparison.
Certainty in each comparison	For each comparison in the network, authors should provide an evaluation of the certainty of evidence; which is most commonly the GRADE assessment. The GRADE assessment includes assessment of: Risk of bias, inconsistency, indirectness, imprecision, and publication bias – as well as incoherence between direct and indirect estimates (see below).	They did not provide an evaluation of the certainty of evidence. Certainty of evidence ratings is an important aspect of ensuring proper interpretation of future NMA results [11, 15].
Incoherence	An NMA should provide exploration of the direct and indirect estimates to identify any instances of incoherence. Explanations of incoherence, such as transitivity, should be discussed in the context of the results.	They provided multiple versions of statistical assessment of incoherence under the "Evaluation of the assumption for network meta-analysis" section in the methods, and inconsistency portion of the results.
Treatment rankings	.As previously discussed, older methods such as SUCRA rankings have major concerns and limitations, more modern methods in minimally or partially contextualized approaches can give more informative rankings of the included treatments [13, 14].	A key limitation is the utilization of SUCRA scores to draw key conclusions within their abstract and manuscript. While it may be warranted to include SUCRA scores within an NMA, it is problematic to use these scores for drawing conclusions over the effect estimates observed.
Robust results across sensitivity assumptions	Often, NMA authors will conduct sensitivity analyses to determine the robustness of the results across different assumptions. If results remain consistent across these sensitivity analyses, greater confidence may be put on the overall findings.	The sensitivity analyses conducted demonstrated comparable results to the main analysis – although minor changes in SUCRA rankings were observed. As previously stated, this should not drive clinical decision making.
Application of Results to Clinical Practice		
Included outcomes	Whenever possible, patient-important outcomes should be directly assessed, opposed to a surrogate outcome.	For POAG, intra-ocular pressure (IOP) is a core outcome, as it is the primary modifiable risk factor for this condition. It is important to also consider patient-reported outcome measures, when applicable. The Li et al. (2016) NMA only provides insight into IOP, although an attempt was made to also evaluate visual field.
Included treatment options	When an NMA does not include all potential treatment options, it leaves uncertainty for clinicians in the potential effects of those omitted treatments. This does not give a full picture of the possible treatment options to consider in a clinical scenario.	They compared a comprehensive list of the available treatment options for POAG, increasing the applicability of the findings.
Subgroup effects	When subgroup effects have been explored, these results can inform the specific clinical scenarios in which particular effects may be observed.	NA – as they did not conduct subgroup analyses.
Overall quality and limitations	Once all of these considerations have been made, clinicians may have a greater sense of the quality of the NMA, and thus can make an informed decision on the implementation of the NMA findings in their clinical practice.	Taking all of these aspects into account, Li et al. conducted a thoughtful NMA on an important topic area, although it was not without some limitations. The primary concerns for interpretation are the omission of a certainty of evidence assessment, as well as the heavy reliance on SUCRA rankings to draw conclusions. Clinicians should come to their own conclusions based on these limitations as to its applicability to practice

REFERENCES

1. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet Lond Engl. 2017;390:415–23. https://doi.org/10.1016/S0140-6736(16)31592-6.
2. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med. 2009;6: e1000097. https://doi.org/10.1371/journal.pmed.1000097.
3. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ. 1997;315:1533–7. https://doi.org/10.1136/bmj.315.7121.1533.
4. Crowther M, Lim W, Crowther MA. Systematic review and meta-analysis methodology. Blood. 2010;116:3140–6. https://doi.org/10.1182/blood-2010-05-280883.
5. Leucht S, Kissling W, Davis JM. How to read and understand and use systematic reviews and meta-analyses. Acta Psychiatr Scand. 2009;119:443–50. https://doi.org/10.1111/j.1600-0447.2009.01388.x.
6. Mills E, Thorlund K, Ioannidis J. Demystifying trial networks and network meta-analysis. BMJ. 2013;346:f2914 https://doi.org/10.1136/bmj.f2914.
7. Neupane B, Richer D, Bonner AJ, Kibret T, Beyene J. Network meta-analysis using r: a review of currently available automated packages. PLOS ONE. 2014;9: e115065. https://doi.org/10.1371/journal.pone.0115065.
8. Foote CJ, Chaudhry H, Bhandari M, Thabane L, Furukawa TA, Petrisor B, et al. Network meta-analysis: users' guide for surgeons: Part I – credibility. Clin Orthop Relat Res. 2015;473:2166–71. https://doi.org/10.1007/s11999-015-4286.

1526

9. Chaudhry H, Foote CJ, Guyatt G, Thabane L, Furukawa TA, Petrisor B, et al. Network meta-analysis: users' guide for surgeons: Part II - certainty. Clin Orthop. 2015;473:2172–8. https://doi.org/10.1007/s11999-015-4287-9.

10. Li T, Lindsley K, Rouse B, Hong H, Shi Q, Friedman DS, et al. Comparative effectiveness of first-line medications for primary open-angle glaucoma: a systematic review and network meta-analysis. Ophthalmology. 2016;123:129–40. https://doi.org/10.1016/j.ophtha.2015.09.005.

11. Puhan MA, Schünemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ. 2014;349:g5630. https://doi.org/10.1136/bmj.g5630.

12. Mbuagbaw L, Rochwerg B, Jaeschke R, Heels-Andsell D, Alhazzani W, Thabane L, et al. Approaches to interpreting and choosing the best treatments in network meta-analyses. Syst Rev. 2017;6:1–5. https://doi.org/10.1186/s13643-017-0473-z.

13. Brignardello-Petersen R, Florez ID, Izcovich A, Santesso N, Hazlewood G, Alhazzani W, et al. GRADE approach to drawing conclusions from a network meta-analysis using a minimally contextualised framework. BMJ. 2020;371:m3900. https://doi.org/10.1136/bmj.m3900

14. Brignardello-Petersen R, Izcovich A, Rochwerg B, Florez ID, Hazlewood G, Alhazzani W, et al. GRADE approach to drawing conclusions from a network meta-analysis using a partially contextualised framework. BMJ. 2020;371:m3907. https://doi.org/10.1136/bmj.m3907

15. Brignardello-Petersen R, Bonner A, Alexander PE, Siemieniuk RA, Furukawa TA, Rochwerg B, et al. Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. J Clin Epidemiol. 2018;93:36–44. https://doi.org/10.1016/j.jclinepi.2017.10.005.

16. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64:383–94. https://doi.org/10.1016/j.jclinepi.2010.04.026.

17. Brignardello-Petersen R, Mustafa RA, Siemieniuk RAC, Murad MH, Agoritsas T, Izcovich A, et al. GRADE approach to rate the certainty from a network meta-analysis: addressing incoherence. J Clin Epidemiol. 2019;108:77–85. https://doi.org/10.1016/j.jclinepi.2018.11.025.

18. Song F, Xiong T, Parekh-Bhurke S, Loke YK, Sutton AJ, Eastwood AJ, et al. Inconsistency between direct and indirect comparisons of competing interventions: meta-epidemiological study. BMJ. 2011;343:d4909. https://doi.org/10.1136/bmj.d4909.

19. Mills E, Ioannidis JPA, Thorlund K, Schünemann HJ, Puhan MA, Guyatt GH. Chapter 24: network meta-analysis. In: Guyatt G, Rennie D, Meade MO, Cook DJ (eds). Users' guides to the medical literature: a manual for evidence-based clinical practice. 3rd ed. McGraw Hill Education; 2015. 327–56.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Varun Chaudhary.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

FOR THE RETINA EVIDENCE TRIALS INTERNATIONAL ALLIANCE (R.E.T.I.N.A.) STUDY GROUP

Varun Chaudhary [1,10 ✉], Mohit Bhandari [1,10], Charles C. Wykoff[4,5], Sobha Sivaprasad[11], Lehana Thabane [1,9], Peter Kaiser[12], David Sarraf[13], Sophie J. Bakri[14], Sunir J. Garg[15], Rishi P. Singh[16,17], Frank G. Holz[18], Tien Y. Wong[19,20] and Robyn H. Guymer[21,22]

[11]NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital, London, UK. [12]Cole Eye Institute, Cleveland Clinic, Cleveland, OH, USA. [13]Retinal Disorders and Ophthalmic Genetics, Stein Eye Institute, University of California, Los Angeles, CA, USA. [14]Department of Ophthalmology, Mayo Clinic, Rochester, MN, USA. [15]The Retina Service at Wills Eye Hospital, Philadelphia, PA, USA. [16]Center for Ophthalmic Bioinformatics, Cole Eye Institute, Cleveland Clinic, Cleveland, OH, USA. [17]Cleveland Clinic Lerner College of Medicine, Cleveland, OH, USA. [18]Department of Ophthalmology, University of Bonn, Bonn, Germany. [19]Singapore Eye Research Institute, Singapore, Singapore. [20]Singapore National Eye Centre, Duke-NUD Medical School, Singapore, Singapore. [21]Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia. [22]Department of Surgery (Ophthalmology), The University of Melbourne, Melbourne, Australia.