Network Meta-Analysis for Various Study Designs: Stepping Outside the Randomized Controlled Trials Comfort Zone Into the Real World

Andreas Karabis PhD, Real World and Analytic Solutions, IQVIA, Amsterdam, The Netherlands

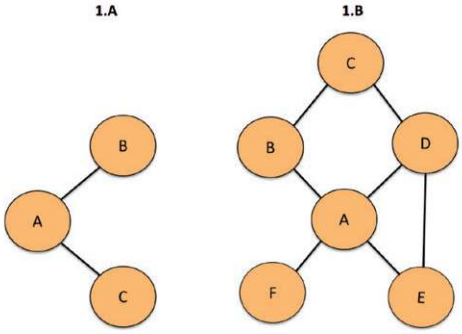
Randomized controlled trials are the dominant type of evidence in network meta-analysis. Appropriate methods for feasibility assessment, analysis, and reporting are essential for the inclusion of real-world evidence in network meta-analyses.

NETWORK META-ANALYSIS: WHAT AND WHY

Evidence from randomized controlled trials (RCTs), collected by means of systematic literature reviews, is routinely used in healthcare decision making, including clinical guidelines and reimbursement. When evidence from more than one RCT comparing the same intervention with the same comparator is available, a quantitative approach called meta-analysis is used to synthesize the results in a single outcome. Meta-analysis was introduced 40 years ago, contributing to the establishment of evidence-based clinical practice.[1]

In many cases, the competing interventions relevant to the decision are more than 2 and are not compared simultaneously in a single RCT. New drugs are often compared only with placebo or standard care, but not against all the alternative interventions of interest. Furthermore, the comparators of interest may vary by country or change over time, making the design of an RCT that includes all the alternatives impractical or not feasible.

Figure 1. Indirect treatment comparison and network meta-analysis



In the absence of RCTs comparing all interventions of interest directly, an indirect treatment comparison (ITC) can provide evidence for the difference in treatment effects.[2] For example, interventions B and C, for which we have only placebo-controlled trials, could be compared indirectly via placebo. This simple network is presented in Figure 1A, where the lines represent head-to-head RCTs.

As an extension of indirect comparison, a network of RCTs could be formed including direct evidence (by means of head-to-head RCTs) and indirect evidence (by means of ITC), in a so-called mixed-treatment comparison (Figure 1B).[3] In line with ISPOR Task Force recommendations, we use the term "network meta-analysis" (NMA) when the evidence network involves more than 2 RCTs and more than 2 interventions.

A valid NMA is based on the assumptions of transitivity (ie, indirect comparison validly estimates the unobserved head-to-head comparison) and consistency (ie, direct and indirect estimates in a network —if available—are in agreement).

THE DOMINANCE OF RCTS AND THE NEED FOR RWE

Although widely accepted and routinely used for decision making, NMAs are often limited to synthesizing evidence from RCTs. Under 4% of the NMAs published until 2014 included designs others than RCTs.[4] This is not unexpected because for decades, the gold standard for evidence generation in medical product evaluation has been the RCT. If designed appropriately and executed as planned, RCTs are expected to provide unbiased results about the effects of alternative interventions.

RCTs have well-known limitations in representing everyday clinical practice, as by definition they are conducted in selected populations and in controlled environments to ensure protocol adherence.

Beyond these fundamental limitations, there are cases where no RCTs are available to support a specific question in healthcare decision making. For example, interventions of interest may not have been studied in RCTs for ethical or other reasons. Also, certain types of effects

cannot be adequately studied in RCTs (eg, safety, long-term outcomes). There also may be cases where the evidence base consists of RCTs that are not adequately designed to address the relevant clinical question.

Even if a properly designed RCT including the interventions, population, and subpopulations of interest is feasible, it can take years from design to completion, while in many cases relevant real-world evidence (RWE) is readily available and can be used to support clinical decisions.

Stepping outside the comfort zone of RCTs is not without challenges, starting with the definitions. An entire universe of study designs are referred to simply as "non-RCTs" or "nonrandomized studies" (NRS). Although the definition of RWE is still evolving, evidence from registry studies, claims databases and administrative data, health surveys, electronic health records, and medical chart reviews is widely accepted as RWE.[5]

However, within the frame of healthcare decision making, NMA is used to estimate the relative effects of interventions. Thus, the main interest regarding RWE is in comparative studies reporting relative treatment effects. When departing from the RCT design, the preferred evidence to be included in an NMA is in the form of well-designed, high-quality cohort studies, case-control studies, and nonrandomized comparative clinical trials, although other study designs may also be considered (Figure 2).[6] Evidence hierarchy, a system of rating the quality of evidence, could be used to navigate through the options.[7,8]

Figure 2. Beyond RCTs

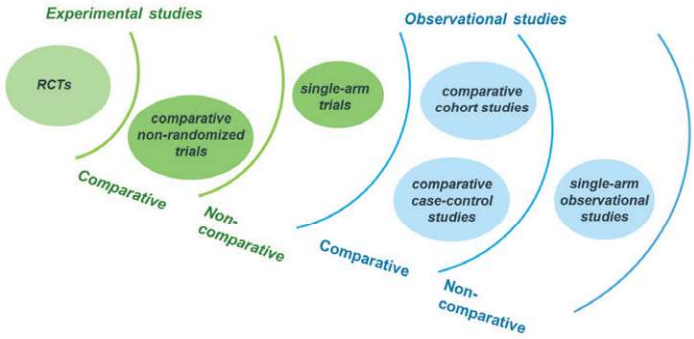
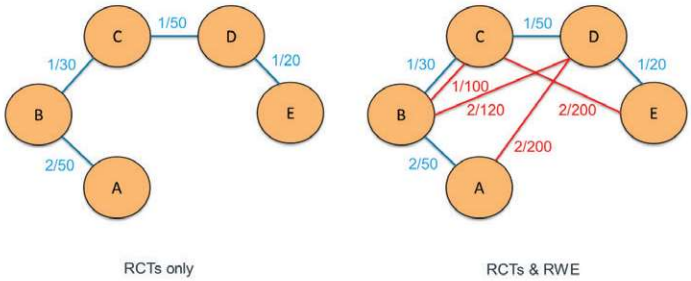


Figure 3. Reducing uncertainty with RWE in case of connected but "weak" networks



Numbers indicate "number of studies/number of participants" of the related comparisons.

The use of RWE is critical in several cases, the most important of which is probably to connect networks. The fundamental assumption of NMA methodology is the existence of a connected network of studies, while lack of RCTs for some of the interventions may lead to a disconnected network. In this case, RWE studies could be used to bridge the gaps and connect the fragmented network.[9] This can also be beneficial in reducing the underlying uncertainty also in the case of connected but "weak" networks (Figure 3) (ie, networks with a small number of RCTs connecting a relatively high number of interventions, networks of RCTs with very low number of randomized patients, or networks of RCTs of extremely poor quality), under certain conditions (eg, agreement between sources of evidence).

HOW CAN THIS WEALTH OF RWE BE USED?

Although inclusion of RWE in an NMA can complement evidence from RCTs or address some of the RCT limitations, thorough review of their quality is necessary. It is well-known that nonrandomized studies are vulnerable to biases, including confounding, thus studies that do not appropriately account for confounding factors may produce biased effect estimates. In this case, the underlying NMA assumption of transitivity may be violated, producing biased estimations for the relative treatment effects.

Before any analysis, the assessment of the quality of RWE should be conducted using appropriate tools, such as ROBINS-I.[10] It is important to ensure that the relative treatment effects in RWE are estimated using appropriate methods to minimize bias.[11] An important indicator of the RWE quality is the agreement between the evidence from RCTs and other study designs that could be easily assessed in a network of interventions connected by both types of studies.

Several methods have been proposed that account for potential bias from RWE in NMA,[6] including:

Design-adjusted analysis: In this approach, each study type can be adjusted separately using a weight w (0<w<1). Setting w=1 means that RWE has the same value as the RCTs (naïve pooling), while setting w=0 means that RWE is ignored. By changing the value of the weight, we can control the confidence we place in RWE. In addition, the point estimate of each study is shifted by a constant representing bias. The appropriate values for weight and shift are not easy to be determined and experts' opinions, as well as sensitivity analysis, are necessary.

Use of comparative RWE as prior information: In a Bayesian framework, prior knowledge can be combined with the current data to derive the posterior distribution, representing the updated state of science on the parameters of interest (eg, treatment effects). Following this approach, the results of a Bayesian NMA including only RWE is considered as prior knowledge for the NMA of the RCTs, in a 2-step approach. The posterior of Step 1 is then adjusted for bias and used as priors for the RCT NMA. There are different approaches for the construction of the RCT NMA prior to the RWE NMA results.[6]

Three-level Bayesian hierarchical model: Three-level hierarchical models can be used to synthesize data from studies >

with different designs (eg, RCTs, cohort studies, case–control). At the first level, each study is analyzed separately to obtain estimates of the relative effects of the interventions that are compared in the study. At the second level, studies of the same design are grouped and synthesized by means of NMA. At the third level, a NMA will synthesize the results of the design-specific NMAs to a single estimate. Furthermore, the estimates from each study (first level) or from each design (second level) can be down-weighted by inflating the variance of the estimates obtained.

Even if a properly designed RCT including the interventions, population, and subpopulations of interest is feasible, it can take years from design to completion, while in many cases relevant real-world evidence is readily available and can be used to support clinical decisions.

Beyond comparative studies: In some clinical areas, the available evidence consists mainly of single-arm studies. In this case, the family of population-adjusted indirect comparison methods has been proposed, including the matching adjusted indirect comparison and the simulated treatment comparison.[12] If patient-level data are available for the single-arm study, these can be connected to the network, following one of these methods.

BABY STEPS OUTSIDE THE COMFORT ZONE

Conducting a valid NMA with RCTs does require the use of proper methods and thorough review of the evidence base. The key is to ensure that patient and study characteristics that can act as potential treatment effect modifiers are balanced and the NMA assumptions are valid. Adding RWE is even more challenging, as it is difficult to predict the magnitude or direction of possible biases, especially when patient-level data are not available. Advancing the statistical methods, understanding the strengths and weaknesses of various data sources, and providing guidance for transparent analysis and reporting are all critical and complementary for valid NMAs, including RWE.

Despite the challenges, the wealth of RWE that is being produced at an increasing rate will become more and more difficult to be ignored or simply excluded from the decision process. The increasing granularity and complexity of RWE, together with the recent advances in data science analytics, offer opportunities to leverage these data for decision support through NMA.

REFERENCES

1. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. Nature, 2018;555(7695):175.

2. 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(6):683-691.

3. Dias S, Ades AE, Welton, NJ, Jansen JP, Sutton AJ. Network meta-analysis for decision-making. Hoboken, NJ: John Wiley & Sons; 2018.

4. Nikolakopoulou A, Chaimani A, Veroniki AA, Vasiliadis HS, Schmid CH, Salanti G. Characteristics of networks of interventions: a description of a database of 186 published networks. PLoS ONE; 2014; 9(1), e86754.

5. GetReal -Project No. 115546, WP1: Deliverable D1.3, Glossary of Definitions of Common Terms (http://www.imi-getreal.eu/Portals/1/Documents/01 deliverables/D1.3. Revised GetReal glossary - FINAL updated version_25Oct16_webversion.pdf.

6. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. Stat Med. 2017;36.8:1210-1226.

7. Howick J, Chalmers I, Glasziou P, et al. Explanation of the 2011 Oxford Centre for Evidence-Based Medicine (OCEBM) Levels of Evidence (Background Document). Oxford Center for Evidence-Based Medicine. https://www.cebm.net/index.aspx?o=5653. Accessed November 3, 2018.

8. Guyatt G.H. et al. GRADE Guidelines: 1. Introduction--GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64:383-394.

9. Jenkins DA, Martina R, Dequen P, Bujkiewicz S, Abrams K. The added value of real-world evidence to connect disconnected networks for network meta-analysis: A case study in rheumatoid arthritis. Value in Health. 2016;19(7):A393.

10. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. BMJ. 2016;355:i4919.

11. Faria R, Hernandez Alava M, Manca A, Wailoo AJ. 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. NICE DSU technical support document. Available from http://www.nicedsu.org.uk. Accessed November 3, 2018.

12. Phillippo D, Ades AE, Dias S, Palmer S, Abrams KR, Welton NJ. NICE DSU Technical Support Document 18: Methods for population-adjusted indirect comparisons in submissions to NICE. 2016. Available from http://www.nicedsu.org.uk. Accessed November 3, 2018.