Research Techniques Made Simple: Network Meta-Analysis

Jennifer Watt[1,2,3], Andrea C. Tricco[3,4], Sharon Straus[1,2,3], Areti Angeliki Veroniki[3,5,6], Gary Naglie[1,2,7] and Aaron M. Drucker[2,8,9]

When making treatment decisions, it is often necessary to consider the relative efficacy and safety of multiple potential interventions. Unlike traditional pairwise meta-analysis, which allows for a comparison between two interventions by pooling head-to-head data, network meta-analysis (NMA) allows for the simultaneous comparison of more than two interventions and for comparisons to be made between interventions that have not been directly compared in a randomized controlled trial. Given these advantages, NMAs are being published in the medical literature with increasing frequency. However, there are important assumptions that researchers and knowledge users (e.g., patients, clinicians, and policy makers) must consider when conducting and evaluating an NMA: network connectivity, homogeneity, transitivity, and consistency. There are also multiple NMA outputs that researchers and knowledge users should familiarize themselves with in order to understand NMA results (e.g., network plots, mean ranks). Our goals in this article are to: (i) demonstrate how NMAs differ from pairwise meta-analyses, (ii) describe types of evidence in a NMA, (iii) explain NMA model assumptions, (iv) provide readers with an approach to interpreting a NMA, (v) discuss areas of ongoing methodological research, and (vi) provide a brief overview of how to conduct a systematic review and NMA.

Journal of Investigative Dermatology (2019) **139,** 4–12; doi:10.1016/j.jid.2018.10.028

CME Activity Dates: 19 December 2018
Expiration Date: 18 December 2019
Estimated Time to Complete: 1 hour

Planning Committee/Speaker Disclosure: Aaron Ducker, MD is a consultant/advisor for Sanofi-Aventis. All other authors, planning committee members, CME committee members and staff involved with this activity as content validation reviewers have no financial relationships with commercial interests to disclose relative to the content of this CME activity.

Commercial Support Acknowledgment: This CME activity is supported by an educational grant from Lilly USA, LLC.

Description: This article, designed for dermatologists, residents, fellows, and related healthcare providers, seeks to reduce the growing divide between dermatology clinical practice and the basic science/current research methodologies on which many diagnostic and therapeutic advances are built.

Objectives: At the conclusion of this activity, learners should be better able to:
- Recognize the newest techniques in biomedical research.
- Describe how these techniques can be utilized and their limitations.
- Describe the potential impact of these techniques.

CME Accreditation and Credit Designation: This activity has been planned and implemented in accordance with the accreditation requirements and policies of the Accreditation Council for Continuing Medical Education through the joint providership of Beaumont Health and the Society for Investigative Dermatology. Beaumont Health is accredited by the ACCME to provide continuing medical education for physicians. Beaumont Health designates this enduring material for a maximum of 1.0 *AMA PRA Category 1 Credit(s)*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Method of Physician Participation in Learning Process: The content can be read from the Journal of Investigative Dermatology website: http://www.jidonline.org/current. Tests for CME credits may only be submitted online at https://beaumont.cloud-cme.com/RTMS-Jan19 – click 'CME on Demand' and locate the article to complete the test. Fax or other copies will not be accepted. To receive credits, learners must review the CME accreditation information; view the entire article, complete the post-test with a minimum performance level of 60%; and complete the online evaluation form in order to claim CME credit. The CME credit code for this activity is: 21310. For questions about CME credit email cme@beaumont.edu.

[1]*Division of Geriatric Medicine, Department of Medicine, University of Toronto, Toronto, Canada;* [2]*Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Canada;* [3]*Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Canada;* [4]*Dalla Lana Faculty of Public Health, University of Toronto, Toronto, Canada;* [5]*Department of Primary Education, School of Education, University of Ioannina, Ioannina, Greece;* [6]*Institute of Reproductive and Developmental Biology, Department of Surgery and Cancer, Faculty of Medicine, Imperial College, London, England;* [7]*Baycrest Health Sciences, Toronto, Canada;* [8]*Division of Dermatology, Department of Medicine, Women's College Hospital, Toronto, Canada; and* [9]*Women's College Research Institute, Women's College Hospital, Toronto, Canada*

Correspondence: Aaron Drucker, Women's College Hospital, 76 Grenville Street, 6th Floor, Toronto, Ontario M5S1B2, Canada. E-mail: aaron.drucker@wchospital.ca

Abbreviations: NMA, network meta-analysis; RCT, randomized controlled trial

SUMMARY POINTS

Comparing pairwise meta-analysis and NMA

- Pairwise meta-analyses allow evidence comparing two interventions to be synthesized; NMAs are used to compare more than two interventions—some of which have not been directly compared in previous RCTs.
- NMAs can be used to rank interventions in terms of their relative efficacy or safety.

Limitations

- Assumptions underlying NMAs must be carefully considered, such as transitivity and consistency, because if these assumptions are not met, it may jeopardize the conclusions of NMA.
- RCTs included in a NMA are subject to the same biases as those included in pairwise meta-analyses and critical appraisal remains an important component of a well-conducted systematic review and NMA.

INTRODUCTION

A growing number of network meta-analyses (NMAs) are being published in the medical literature (Zarin et al., 2017). NMAs offer a way to make comparisons between many interventions simultaneously, helping to synthesize large amounts of data relating to clinical outcomes. NMAs can also make indirect comparisons between interventions that have not been compared in randomized controlled trials (RCTs) and rank interventions in terms of their relative efficacy or safety. While there are clear advantages to NMAs, their conduct and interpretation is more complex than that of pairwise meta-analyses. Therefore, it is important for those conducting and reading NMAs to learn how to understand and interpret the findings. In this article, we will: (i) delineate how NMAs differ from pairwise meta-analyses, (ii) describe types of evidence in a NMA, (iii) explain NMA model assumptions, (iv) provide readers with an approach to interpreting an NMA, (v) discuss areas of ongoing methodological research, and (vi) provide a brief overview of how to conduct a systematic review and NMA. Two NMAs on treatments for psoriasis will be used to illustrate these concepts (Jabbar-Lopez et al., 2017; Reich et al., 2012).

COMPARING PAIRWISE META-ANALYSIS AND NMA

Pairwise meta-analysis and NMA are compared and contrasted in Table 1. Pairwise meta-analyses are applied when the desired end point is to derive a summary effect estimate across a number of studies that compare the same two interventions (Figure 1a) (Abuabara et al., 2012). However, for many comparative effectiveness questions, the goal is to understand the relative efficacy and safety of more than two interventions. For example, therapeutic decision making for a patient with moderate to severe chronic plaque psoriasis requires comparison of all possible interventions, including adalimumab, etanercept, other biologics, traditional systemic medications, and small molecule–targeted agents.

Table 1. Comparing Pairwise and Network Meta-Analysis

Variable	Pairwise meta-analysis	Network meta-analysis
Number of comparators	2	>2
Questions answered by analysis method	What is the efficacy or risk of harm associated with one intervention compared to another?	Which interventions are efficacious and/or safe? What intervention is the most efficacious and/or safe? What is the comparative efficacy and/or safety between two interventions that hasn't been compared directly?
Systematic review question format	PICO	Modified PICO to accommodate additional treatment comparisons
Risk of bias appraisal	Cochrane Risk of Bias Tool for RCTs	Cochrane Risk of Bias Tool for RCTs
Assumptions	Homogeneity	Network connectivity Homogeneity Transitivity Consistency
Influential biases	Publication bias and small-study effects Confounding Selection bias Information bias	Publication bias and small-study effects Confounding Selection bias Information bias
Model outputs	Summary effect estimates (e.g., OR, MD, SMD) and forest plot Funnel plot	Network plot Transitivity plot or table Summary effect estimates (e.g., OR, MD, SMD) and forest plot Ranking statistic: mean rank, SUCRA value or P-score Inconsistency plot Comparison-adjusted funnel plot
Limitations	Effect modifiers create heterogeneity Biases can generate misleading results	Effect modifiers create heterogeneity and/or inconsistency Biases can generate misleading results
Reporting guidelines	PRISMA	PRISMA-NMA

Abbreviations: MD, mean difference; NMA, network meta-analysis; OR, odds ratio; PICO, population, intervention(s), comparator(s), outcome(s); PRIMSA, Preferred Reporting Guidelines for Systematic Reviews and Meta-Analyses; RCT, randomized controlled trial; SMD, standardized mean difference; SUCRA, surface under the cumulative ranking curve.

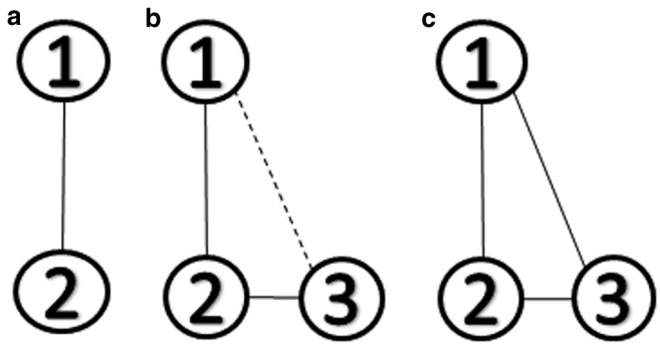

Figure 1. Illustration of intervention comparisons in pairwise and network meta-analysis. (**a**) A pairwise comparison between interventions 1 and 2. (**b**) Two direct comparisons (intervention 1 vs. 2 and intervention 2 vs. 3) and one indirect comparison (intervention 1 vs. 3) in a network meta-analysis. (**c**) Three direct comparisons (intervention 1 vs. 2, intervention 2 vs. 3, and intervention 1 vs. 3) that form a closed loop.

This can be accomplished with NMA, from which summary effect estimates can be derived across more than two interventions, some of which have never been compared directly. Like pairwise meta-analyses, NMAs can be conducted in a frequentist or Bayesian framework (Chaimani et al., 2013; Dias et al., 2018; van Valkenhoef and Kuiper, 2016).

DIRECT AND INDIRECT EVIDENCE

Estimates of relative efficacy or safety from NMA models can be derived by combining both direct and indirect evidence from intervention comparisons that form a connected network (Figure 2) (see section Assumptions of Network Meta-Analysis). Direct evidence describes data taken from at least one RCT. Indirect evidence is derived from NMA models to describe the relative efficacy or safety for intervention comparisons that have not been studied in an RCT (Figure 1b). When a comparison is informed by both direct and indirect evidence, this is referred to as a mixed effect estimate (Dias et al., 2018). For example, in the NMA conducted by Jabbar-Lopez et al. (2017), on the evaluation of biologic therapies for psoriasis, there was no RCT evidence comparing adalimumab and etanercept directly for the outcome of "clear/nearly clear"; however, there were direct comparisons between (i) adalimumab and placebo and (ii) etanercept and placebo. Authors were able to derive an indirect effect estimate comparing adalimumab and etanercept because each intervention had been compared to a common intervention (placebo) (Figure 2) (Jabbar-Lopez et al., 2017).

ASSUMPTIONS OF NMA

There are four key assumptions of NMAs: (i) network connectivity, (ii) homogeneity, (iii) transitivity, and (iv) consistency (Table 2). The requirement for network connectivity is unique to NMA. Interventions must be connected to the network to draw any conclusions about their direct and indirect relationships with other interventions. In Figure 2, each intervention is connected to at least one other intervention in each network. If a treatment comparison is not connected to any other treatments in the network, it cannot be a part of the NMA.

Readers are likely familiar with the concept of homogeneity: the true intervention effect should be sufficiently similar across all studies making a direct comparison between the same two intervention groups. Similar to pairwise meta-analyses, different potential sources of heterogeneity must be considered in studies included in NMAs: clinical, methodological, and statistical. If heterogeneity is anticipated between studies, then a random-effects as opposed to fixed-effects model should be implemented (Higgins and Green, 2011).

The assumptions of transitivity and consistency refer to our assessment of potential clinical and methodological effect modifiers across a network of interventions. In assessing transitivity, a judgment must be made about the distribution of effect modifiers and how they might influence direct and indirect effect estimates. For example, if all patients in one psoriasis intervention comparison have severe disease at

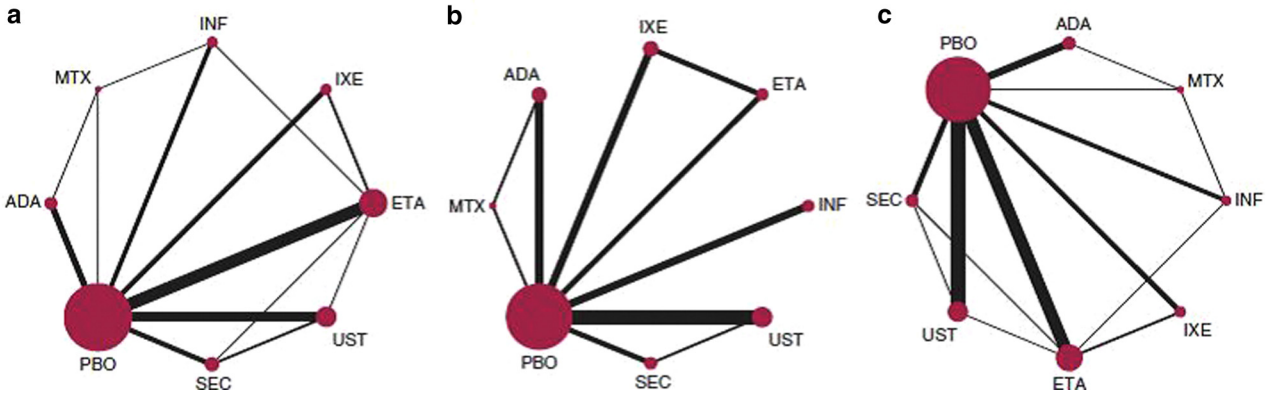

Figure 2. Examples of network plots. Connected network plots (Jabbar-Lopez et al., 2017). Nodes represent individual interventions and nodes connected by lines indicate that these two interventions have previously been compared directly in a study. In these examples, the nodes are weighted by the number of studies evaluating this treatment and the lines are weighted by the number of studies evaluating this treatment comparison. Each panel is a network plot of interventions reporting the outcome of interest: (**a**) clear/nearly clear (minimal residual activity/Psoriasis Area and Severity Index >90/0 or 1 on Physician's Global Assessment), (**b**) mean change in the Dermatology Life Quality Index, and (**c**) withdrawal due to adverse events. ADA, adalimumab; ETA, etanercept; INF, infliximab; IXE, ixekizumab; MTX, methotrexate; PBO, placebo; SEC, secukinumab; UST, ustekinumab.

Table 2. Questions to Consider When Assessing the Assumptions of a Network Meta-Analysis

Assumption	Questions to consider
Homogeneity	Is there any clinical, methodological, or statistical heterogeneity between studies that compare the same interventions? Are there effect modifiers (e.g., age, gender, illness severity) between studies making the same treatment comparison that could influence the summary effect estimate?
Network connectivity	Do all of the interventions form a connected network (as in Figure 2)?
Transitivity	Is there an imbalance in effect modifiers among studies included in the network? In theory, could any patient randomized in one study within a network have been randomized to any of the other studies in this same network?
Consistency	Where possible to assess, are the direct and indirect effect estimates from closed loops in the network in agreement?

baseline (Interventions 1 vs. 2), while all patients in the other two treatment comparisons in a loop have moderate disease at baseline (interventions 1 vs. 3 and 2 vs. 3), this violates the transitivity assumption. When there are imbalances in effect modifiers across the network, subgroup analyses or meta-regression could be used to explore their influence on NMA effect estimates, or perhaps the NMA should not be conducted.

Consistency is the statistical measure of transitivity. There may be inconsistency in a closed network loop if there is an imbalance of effect modifiers across treatment comparisons. In essence, direct and indirect effect estimates can be compared within a network to assess their level of disagreement. There are tests that assess for consistency in a network as a whole (global tests) or at certain paths (e.g., closed loops) of a network (local tests) (Dias et al., 2018). For example, the results of a loop-specific approach to the assessment of inconsistency (local test) are presented in Figure 3. There is inconsistency in the closed loop containing three comparisons: placebo-methotrexate, placebo-infliximab, and methotrexate-infliximab. This means that the direct and indirect effect estimates of one of the treatment comparisons within this closed loop are significantly different from one another (the inconsistency factor's 95% confidence interval does not cross zero). There is no inconsistency identified in the other closed loops. It is possible that statistical tests of consistency may fail to identify inconsistency; therefore, it is important to consider whether the transitivity assumption has been met prior to undertaking an NMA.

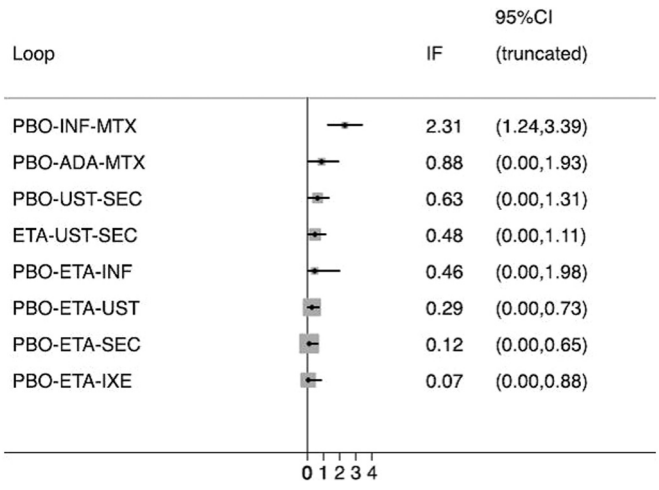

Figure 3. Example of an inconsistency plot. This is an example of an inconsistency plot with closed triangular loops of treatment comparisons evaluating the Psoriasis Area and Severity Index 75 at 12/16 weeks (Jabbar-Lopez et al., 2017). The x-axis represents the scale for the IFs. The PBO-INF-MTX loop shows evidence of inconsistency between direct and indirect evidence because the 95% CI for the IF does not include zero. There is no significant inconsistency identified in any of the other loops. ADA, adalimumab; CI, confidence interval; ETA, etanercept; IF, inconsistency factor; INF, infliximab; IXE, ixekizumab; MTX, methotrexate; PBO, placebo; SEC, secukinumab; UST, ustekinumab.

RCTs in an NMA are subject to the same biases as those included in pairwise meta-analyses. Critical appraisal of RCTS in an NMA is important because studies at high risk of bias can lead to violations of the homogeneity, transitivity, and consistency assumptions. For example, if indirect evidence from a closed network loop of studies at low risk of bias in all aspects of critical appraisal did not show a significant benefit to receiving treatment, but one study (direct evidence) at high risk of bias from lack of participant and outcome assessor blinding found a benefit to receiving treatment, this will violate the transitivity (and possibly the consistency) assumption. Similarly, between-study heterogeneity will be created if one study at high risk of bias due to lack of participant and outcome assessor blinding found a benefit to receiving a treatment, while a second study that was at low risk of bias on these aspects of critical appraisal did not find such a benefit.

INTERPRETING NMA

A number of different measures of intervention efficacy and safety can be derived from NMAs (Table 3) (Dias et al., 2018). Figures and explanations for network plots (Figure 2), surface under the cumulative ranking curves (Figure 4), an inconsistency plot (Figure 3), and a comparison-adjusted funnel plot (Figure 5) are provided (Jabbar-Lopez et al., 2017). By convention, a higher mean rank or greater surface under the cumulative ranking value indicates that an intervention is either more efficacious or safer (Dias et al., 2018). While most people are familiar with the interpretation of a frequentist effect estimate, people may be less familiar with the interpretation of a Bayesian effect estimate. Reich et al. (2012) reported the mean relative risk (and 95% credible interval) of 50%, 75%, and 90% reductions in the Psoriasis Area and Severity Index for patients with moderate to severe

Table 3. Commonly Reported Network Meta-Analysis Outputs

Network meta-analysis output	Description	Interpretation
Network plot	A diagram depicting how interventions (nodes) are connected to one another through direct comparisons (lines) (see Figure 2)	Provides an overview of the available evidence; a network estimate of an intervention's relative efficacy or safety compared to other interventions in the network can only be calculated if it is connected to the network
Transitivity plot or table	A table or plot summarizing potential effect modifiers across studies	Studies in each network should appear sufficiently similar so that the observed treatment effects are the result of receiving each treatment and not an imbalance in effect modifiers
Summary effect estimate	Estimate of the relative efficacy of interventions in the network (e.g., OR, MD, SMD, HR) compared to other network interventions, reported with a measure of uncertainty (e.g., confidence/credible intervals or predictive intervals)	Same interpretation as a summary effect estimate in a pairwise meta-analysis
Ranking statistics	Frequently presented as a mean/median rank, SUCRA value (or P-score) or probability of being the best treatment	An intervention with a higher treatment ranking, SUCRA value, or probability of being the best is more efficacious or more likely to cause harm
Inconsistency plot	A plot reporting the inconsistency factors (absolute difference between direct and indirect effect estimates) for each comparison in a closed network loop (see Figure 3)	An inconsistency factor with a confidence interval that does not include zero indicates that there is significant inconsistency between direct and indirect effect estimates
Comparison-adjusted funnel plot	Similar to a funnel plot in pairwise meta-analyses; however, the x-axis is the difference between each study-specific effect estimate and pooled effect estimate for each comparison and comparisons have been ordered in a meaningful way (e.g., chronological treatment order) (see Figure 5)	Asymmetry in the plot indicates publication bias/small-study effects

Abbreviations: HR, hazard ratio; MD, mean difference; NMA, network meta-analysis; OR, odds ratio; SMD, standardized mean difference; SUCRA, surface under the cumulative ranking curve.

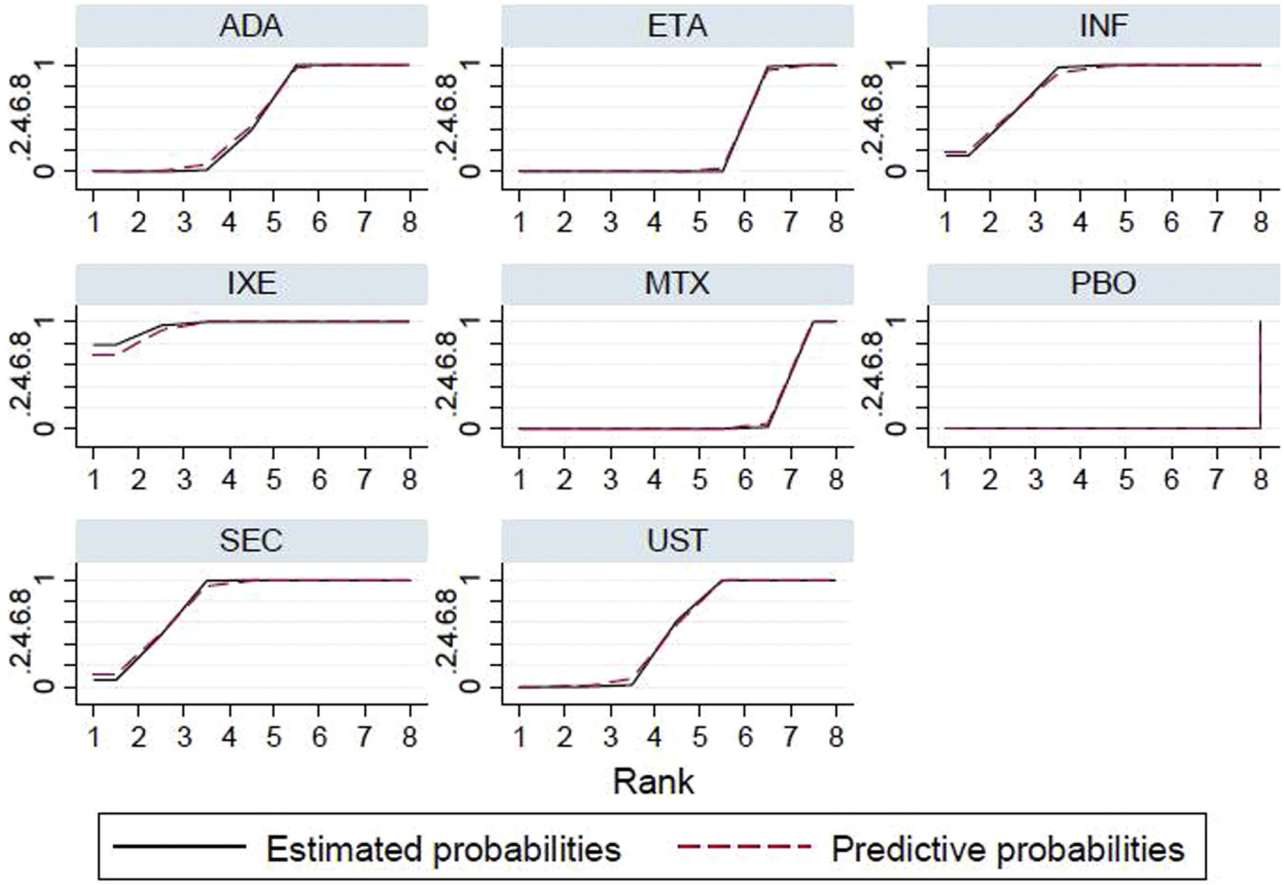

Figure 4. Examples of SUCRA curves. SUCRA curves of treatments evaluating the Psoriasis Area and Severity Index 75 at 12/16 weeks (Jabbar-Lopez et al., 2017). The cumulative probability that each treatment is ranked among the top n (e.g., 1, 2, …, 8) treatments (y-axis) is plotted against each possible rank (x-axis) for treatments in the network. Predictive probabilities incorporate the uncertainty in our network estimates from heterogeneity. IXE has the highest SUCRA value (96.4%) and PBO has the lowest SUCRA value (0%). ADA, adalimumab; ETA, etanercept; INF, infliximab; IXE, ixekizumab; MTX, methotrexate; PBO, placebo; SEC, secukinumab; SUCRA, Surface Under the Cumulative Ranking; UST, ustekinumab.

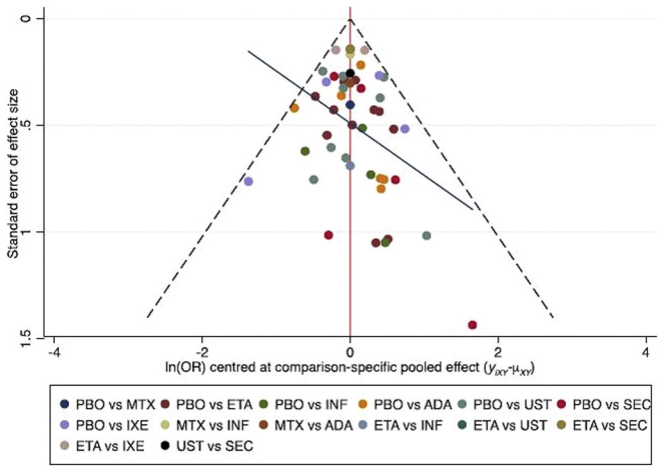

Figure 5. Example of a comparison-adjusted funnel plot. This is an example of a comparison-adjusted funnel plot of treatment comparisons evaluating the Psoriasis Area and Severity Index 75 at 12/16 weeks (Jabbar-Lopez et al., 2017). Comparisons are color-coded as per the legend at the bottom of the figure. The y-axis represents the standard error of each study-specific effect estimate. The x-axis represents the difference between the ln(OR) for each study-specific effect estimate and the pooled effect estimate for each comparison (e.g., all of the study-specific estimates reporting on the PBO vs. ADA comparison). The blue diagonal line represents a linear regression of the x-axis variable on the y-axis variable. The paucity of studies in the bottom left of the plot indicates there may be small studies missing that would have favored established treatments. ADA, adalimumab; ETA, etanercept; INF, infliximab; IXE, ixekizumab; MTX, methotrexate; OR, odds ratio; PBO, placebo; SEC, secukinumab; UST, ustekinumab.

psoriasis receiving biologics. In this case, the relative risk value represents the mean of the relative risk posterior distribution for each relative treatment effect, and the 95% credible interval represents the range of values within which there is a 95% probability that the true value of the relative risk is found, given the observed data. In contrast, Jabbar-Lopez et al. (2017) used a frequentist NMA approach. In a frequentist framework, the 95% confidence interval means that there is a 95% chance of the true relative risk value being found within the intervals, given repeated randomized sampling. Frequentist modeling treats data as random and parameters as fixed unknown constants, whereas, Bayesian modeling treats data as fixed and parameters as random (Kadane, 1995).

Knowledge users can use the International Society for Pharmacoeconomics and Outcomes Research tool for interpreting NMAs in health care decision making or the *Journal of the American Medical Association Users' Guide to the Medical Literature* on NMAs for interpreting and critically appraising a systematic review and NMA (Jansen et al., 2014; Mills et al., 2012). The GRADE (Grades of Recommendation, Assessment, Development and Evaluation) approach has also been extended to assess the certainty of NMA results. It provides a framework for determining the quality of evidence in NMA-derived effect estimates for each outcome (Brignardello-Petersen et al., 2018; Salanti et al., 2014).

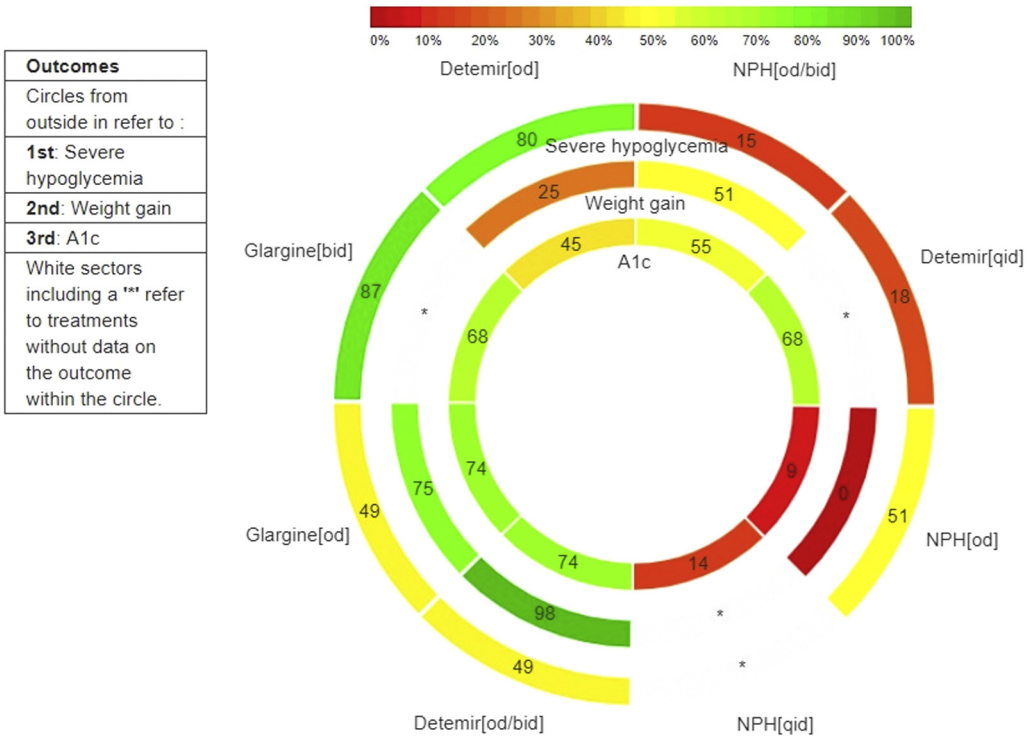

Figure 6. Example of a rank-heat plot. This is an example of a rank-heat plot of outcomes associated with insulin use in patients with type 1 diabetes mellitus. Each ring represents a different outcome. Outcomes are also specified in the legend. Each "slice" represents a different treatment. Treatments are ranked according to their surface under the cumulative ranking curve values. Higher surface under the cumulative ranking curve values (in green) indicate more efficacious and safer treatments. Uncolored areas indicate that the treatment was not included in the network meta-analysis of that outcome. A1c, hemoglobin A1c; bid, twice daily, OD, once daily; qid, four times per day.

Table 4. Conducting a Systematic Review and Network Meta-Analysis

Steps to follow when conducting a systematic review and network meta-analysis
1. Follow a modified PICO format when developing clinical questions for systematic reviews and NMAs because you are considering multiple intervention and comparator groups.
2. Register your systematic review and NMA protocol with PROSPERO and consider publishing the protocol in a peer-reviewed journal.
3. Develop a comprehensive literature search strategy that will encompass all of the interventions and outcomes of interest.
4. Complete all steps relating to article screening, data abstraction, and risk of bias appraisal independently in duplicate.
5. Inspect network plots to ensure all interventions form a connected network.
6. Make judgments concerning the homogeneity and transitivity assumptions prior to conducting NMA. Be explicit about how you model heterogeneity in your NMA if you implement a random-effects model.
7. Describe any assessments of global and local inconsistency. If there is inconsistency in your NMA, state how this is addressed.
8. Assess for small-study effects and publication bias by using a plot such as the comparison-adjusted funnel plot.
9. Present summary effect estimates for interventions and an estimate of heterogeneity. You can also present ranking statistics such as a mean rank and a SUCRA value for each intervention.
10. Follow the recommendations of the PRISMA extension statement for the reporting of NMAs when submitting your systematic review and NMA for publication (Hutton et al., 2015).

Abbreviations: NMA, network meta-analysis; PICO, population, interventions, comparators, outcome(s); PRISMA, Preferred Reporting Guidelines for Systematic Reviews and Meta-Analyses; PROSPERO, International Prospective Register of Systematic Reviews; SUCRA, surface under the cumulative ranking curve.

AREAS OF ONGOING METHODOLOGICAL RESEARCH IN NMA

There remain a number of questions about how to apply NMA methods in clinical and policy decision making. For example, what is the best way to present NMA results to knowledge users? In addition to reporting summary effect estimates, is it best to report all of the surface under the cumulative ranking curve values individually or should a method like a rank-heat plot be utilized (Veroniki et al., 2016a)? A rank-heat plot is a collated graphical representation of ranking statistics demonstrating the comparative effect of interventions on a number of outcomes (Figure 6). How can data from non-randomized studies be incorporated into NMAs? For adverse event data, in particular, this is an important topic because many RCTs are underpowered to detect the potential for harm. Several models have been proposed to include non-randomized studies in NMAs: (i) naïve pooling, (ii) data from non-randomized studies as prior information, and (iii) a three-level hierarchical model with an additional level of uncertainty to account for the inclusion of different study designs (Schmitz et al., 2013). Lastly, how can individual patient-level data best be included in NMAs to account for potential effect modifiers? Meta-analysts are using several methods to incorporate individual patient-level data, including one- and two-stage Bayesian hierarchical NMA models (Veroniki et al., 2016b).

CONDUCTING A SYSTEMATIC REVIEW AND NMA

We provide an overview of the steps necessary to conduct a systematic review and NMA in Table 4. There are statistical packages available to conduct frequentist and Bayesian NMAs (Chaimani et al., 2013; van Valkenhoef and Kuiper, 2016). In conducting a Bayesian NMA, special consideration needs to be given to the choice of prior information for stochastic model parameters (Dias et al., 2018). Reich et al. (2012) implemented vague prior distributions for study-specific baselines in their NMA of biologic treatments for moderate to severe psoriasis, but minimally informative and informative priors are also used in Bayesian NMAs (Dias et al., 2018; Reich et al., 2012).

SUMMARY

Researchers may wish to undertake a systematic review and NMA because they can make indirect comparisons between interventions that have not been previously compared in RCTs, compare the relative efficacy or safety of more than two interventions simultaneously, and rank interventions in terms of their relative efficacy or safety. Much work has been done to improve the reporting and interpretability of NMA results; however, researchers and knowledge users must be cautious when reading NMA results and carefully consider many of the same limitations that face pairwise meta-analyses, including potential threats to the validity of meta-analytic findings from systematic biases.

CONFLICT OF INTEREST

Aaron M. Drucker served as an investigator and has received research funding from Sanofi and Regeneron and has been a consultant for Sanofi, RTI Health Solutions, Eczema Society of Canada and Canadian Agency for Drugs and Technology in Health. He has received honoraria from Astellas Canada, Prime Inc, Spire Learning, CME Outfitters, and Eczema Society of Canada. Jennifer Watt is funded by a doctoral research award from the Canadian Institutes of Health Research and the University of Toronto Department of Medicine Eliot Phillipson Clinician Scientist Training Program. Andrea C. Tricco is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. Andrea C. Tricco receives funding from the Government of Canada through a Canada Research Chair in Knowledge Synthesis. Sharon Straus is funded by a Tier 1 Canada Research Chair in Knowledge Translation. The remaining authors state no conflict of interest.

AUTHOR CONTRIBUTIONS

Jennifer Watt drafted the manuscript. Jennifer Watt, Andrea C. Tricco, Sharon Straus, Areti Angeliki Veroniki, Gary Naglie, and Aaron M. Drucker contributed to the conception, design, and critical revision of the manuscript, and approved the final manuscript.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to this paper. Teaching slides are available as supplementary material.

MULTIPLE CHOICE QUESTIONS

1. Which of the following are advantages of conducting a network meta-analysis as compared to a pairwise meta-analysis?
   A. Make indirect comparisons between interventions that have not been previously compared in randomized controlled trials.
   B. Rank interventions in terms of their relative efficacy or safety.
   C. Increase the precision of our summary effect estimates by including both direct and indirect evidence.
   D. All of the above

2. You read an article reporting the results of a systematic review and network meta-analysis. The authors report there was no inconsistency detected in their network meta-analysis models. You should:
   A. Accept the network meta-.analysis results as robust because there was no inconsistency identified
   B. Read further in the study methods and results section to see if the authors evaluated the transitivity assumption prior to conducting the network meta-analysis.
   C. Consider the similarities and differences between the studies included in the network meta-analysis to evaluate the transitivity assumption.
   D. B and C

3. Which of the following model outputs are common to both pairwise and network meta-analysis?
   A. Summary effect estimate (e.g., odds ratio, mean difference)
   B. Mean rank
   C. Surface under the cumulative ranking curve value
   D. Inconsistency plot

4. Which of the following scenarios best describes a homogeneous comparison?
   A. The mean age of patients enrolled in studies evaluating comparison AB is 65 years; whereas, the mean age of patients enrolled in studies evaluating comparison AC is 70 years.
   B. Among three studies evaluating comparison AB, the mean age of patients enrolled in study #1 is 65 years, the mean age of patients enrolled in study #2 is 66 years, and the mean age of patients enrolled in study #3 is 63 years.
   C. The mean age of patients enrolled in studies evaluating comparison AB is 65 years; whereas, the mean age of patients enrolled in studies evaluating comparison AC is 66 years.
   D. Among three studies evaluating comparison AB, the mean age of patients enrolled in study #1 is 65 years, the mean age of patients enrolled in study #2 is 45 years, and the mean age of patients enrolled in study #3 is 80 years.

5. You conduct a network meta-analysis on the comparative risk of death from new drugs used to treat atopic dermatitis. The mean ranks for four of the new drugs are as follows:

   Drug A 6.2

   Drug B 3.4

   Drug C 8.1

   Drug D 1.5

   Which of the following is true?
   A. Drug A is associated with a greater risk of death compared to Drug B.
   B. Drug D is associated with a lower risk of death compared to Drug C.
   C. Drug A is associated with a lower risk of death compared to Drug B.
   D. Drug D is associated with a lower risk of death compared to Drug A.

REFERENCES

Abuabara K, Freeman EE, Dellavalle R. The role of systematic reviews and meta-analysis in dermatology. J Invest Dermatol 2012;132(11):e2.

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.

Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. PLoS One 2013;8(10):e76654.

Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. Network Meta-Analysis for Decision-Making. Hoboken, NJ: Wiley; 2018.

Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. Updated March 2011.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med 2015;162:777–84.

Jabbar-Lopez ZK, Yiu ZZN, Exton LS, Firouz Mohd Mustapa M, Samarasekera E, David Burden A, et al. Quantative evaluation of biologic therapy options for psoriasis: a systematic review and network meta-analysis. J Invest Dermatol 2017;137:1646–54.

Jansen JP, Trikalinos T, Cappelleri JC, Daw J, Andes S, Eldessouki R, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force Report. Value Health 2014;17:157–73.

Kadane JB. Prime Time for Bayes. Control Clin Trials 1995;16:313–8.

Mills EJ, Ioannidis JPA, Thorlund K, Schunemann HJ, Puhan MA, Guyatt GH. How to use an article reporting a multiple treatment comparison meta-analysis. JAMA 2012;308:1246–53.

Reich K, Burden AD, Eaton JN, Hawkins NS. Efficacy of biologics in the treatment of moderate to severe psoriasis: a network meta-analysis of randomized controlled trials. Br J Dermatol 2012;166:179–88.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. PLoS One 2014;9(7):e99682.

Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. Stat Med 2013;32:2935–49.

Van Valkenhoef G, Kuiper J. Package 'gemtc'. December 23, 2016. https://cran.r-project.org/web/packages/gemtc/gemtc.pdf. Accessed: November 19, 2018.

Veroniki AA, Straus SE, Fyraridis A, Tricco AC. The rank-heat plot is a novel way to present the results from a network meta-analysis including multiple outcomes. J Clin Epidemiol 2016a;76:193–9.

Veroniki AA, Straus S, Soobiah C, Elliot MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. BMC Med Res Methodol 2016b;16(47).

Zarin W, Veroniki AA, Nincic V, Varaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. BMC Med 2017;15(3):61.

DETAILED ANSWERS

1. Which of the following are advantages of conducting a network meta-analysis as compared to a pairwise meta-analysis?

Correct answer: D. All of the above

You should never rely solely on tests of inconsistency to detect inconsistency in a network meta-analysis. You must first conduct an assessment of transitivity across comparisons in the network to ensure effect modifiers are balanced. Authors should conduct an assessment of transitivity and they should provide a way for readers of their study to assess the transitivity assumption as well.

2. You read an article reporting the results of a systematic review and network meta-analysis. The authors report there was no inconsistency detected in their network meta-analysis models. You should:

Correct answer: D. B and C

Summary effect estimates are reported in both pairwise and network meta-analyses.

3. Which of the following model outputs are common to both pairwise and network meta-analysis?

Correct answer: A. Summary effect estimate (e.g., odds ratio, mean difference)

Summary effect estimates are reported in both pairwise and network meta-analyses.

4. Which of the following scenarios best describes a homogeneous comparison?

Correct answer: B. Among three studies evaluating comparison AB, the mean age of patients enrolled in study #1 is 65 years, the mean age of patients enrolled in study #2 is 66 years, and the mean age of patients enrolled in study #3 is 63 years.

The three studies that have compared treatments A and B have a similar distribution of patient ages, which indicates there is homogeneity with regards to patient age within this treatment comparison. Choice c is an example of the transitivity assumption. Patients enrolled in studies comparing treatments A and B and treatments A and C are similar in age, which confirms the transitivity assumption to be valid with regards to the potential effect modifier of patient age.

5. You conduct a network meta-analysis on the comparative risk of death from new drugs used to treat atopic dermatitis. The mean ranks for four of the new drugs are as follows:
Drug A 6.2
Drug B 3.4
Drug C 8.1
Drug D 1.5
Which of the following is true?

Correct answer: A. Drug A is associated with a greater risk of death compared to Drug B.

Drugs with a lower mean rank are associated with a higher risk of death.