

REVIEW

Characteristics and methods of incorporating randomized and nonrandomized evidence in network meta-analyses: a scoping review

Kathryn Zhang^{a,*}, Paul Arora^b, Neha Sati^c, Audrey Béliveau^d, Natalie Troke^a, Areti Angeliki Veroniki^{e,f,g}, Myanca Rodrigues^e, Patricia Rios^e, Wasifa Zarin^e, Andrea C. Tricco^{a,b,e}

^aDepartment of Public Health, Dalla Lana School of Public Health, University of Toronto, 6th floor, 155 College Street, Toronto, Ontario M5T 3M7, Canada

^bEpidemiology Division, Dalla Lana School of Public Health, University of Toronto, 6th floor, 155 College Street, Toronto, Ontario M5T 3M7, Canada

^cDepartment of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada

^dDepartment of Statistics and Actuarial Sciences, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

^eLi Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building, Toronto, Ontario M5B 1W8, Canada

^fDepartment of Primary Education, School of Education, University of Ioannina, Ioannina, Greece

^gDepartment of Surgery & Cancer, Institute of Reproductive and Developmental Biology, Faculty of Medicine, Imperial College, London W12 0NN, UK

Accepted 4 April 2019; Published online 3 May 2019

Abstract

Objectives: The objective of the study was to conduct a scoping review of the published literature on methods used to combine randomized and nonrandomized evidence (NRE) in network meta-analyses (NMAs) and their respective characteristics.

Study Design and Setting: We conducted a scoping review using a list of NMAs which incorporated NRE that were identified from a previous review. All NMAs that included NRE in the analysis of main outcomes or sensitivity analyses were eligible for inclusion. Two reviewers independently screened studies for inclusion and performed data abstraction. Data analysis involved quantitative (frequencies and percentages) and qualitative (narrative synthesis) methods.

Results: A total of 23 NMAs met the predefined inclusion criteria, of which 74% ($n = 17$) used naïve pooling, 0% used NRE as informative priors, 9% ($n = 2$) used the 3-level Bayesian hierarchical model, 9% ($n = 2$) used all methods, and 9% ($n = 2$) used other methods. Most NMAs were supplemented with additional analyses to investigate the effect estimates when only randomized evidence was included.

Conclusion: Although most studies provided justification for the inclusion of NRE, transparent reporting of the method used to combine randomized evidence and NRE was unclear in most published networks. Most NMAs used naïve pooling for combining randomized evidence and NRE. © 2019 Elsevier Inc. All rights reserved.

Keywords: Network meta-analysis; Mixed treatment comparison; Indirect treatment comparison; Evidence synthesis; Non-randomized; Statistical methods

1. Background

1.1. Introduction to NMA

The comparative effectiveness of two treatments is often evaluated through pairwise meta-analysis by synthesizing

an overall treatment effect estimate from direct head-to-head randomized controlled trials (RCTs) [1]. Network meta-analysis (NMA) extends this approach by evaluating many pairwise comparisons across a variety of interventions [2,3]. This method has become increasingly popular in health care because it is able to produce indirect estimates, for which no head-to-head evidence exists and can be used to rank the safety and effectiveness of alternative interventions. This is important because multiple treatments and interventions are becoming more commonly available for most clinical conditions.

NMAs are traditionally restricted to evidence from RCTs, as they are considered the highest quality of evidence for studying interventions and are the standard statistical basis for NMAs [4]. More recently, there has been a growing interest in including nonrandomized evidence

Conflicts of interest: Dr. Andrea C. Tricco is an Associate Editor for the Journal of Clinical Epidemiology but was not involved in the peer-review process or decision for publication. All other authors have no conflicts of interest.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics approval: Ethics approval was not necessary for this study as it is a scoping review of secondary analyses.

Availability of data and materials: All data are available from the tables, figures, and appendix of this study.

* Corresponding author.

E-mail address: Kathryn.zhang@mail.utoronto.ca (K. Zhang).

What is new?

Key findings

- Earlier network meta-analyses (NMAs) have been conducted with an evidence base comprised entirely of randomized controlled trials (RCTs). Nonrandomized evidence (NRE) is obtained from observational studies or patient registries. Recently, the use of NRE is becoming widely recognized as a potential source of evidence in NMAs. This scoping review identifies the statistical methodologies used to combine randomized evidence and NRE, and the most common method of incorporating NRE was naïve pooling. Most studies provided justification for the inclusion of NRE in NMAs; however, most studies failed to transparently report the methods used to combine RCTs and NRE.

What this adds to what was known?

- This epidemiological review empirically evaluates the practice of incorporating NRE into the published NMA literature. Our evidence base is comprised of a list of studies from a previously completed search of NMAs that incorporated NRE. This review provides an overview of the characteristics and methodologies specific to NMAs that include NRE. We conclude that the use of NRE will continue to grow in the future and careful consideration is required when incorporating NRE to produce more comprehensive and generalizable treatment effect estimates.

What is the implication and what should change now?

- This empiric data set of methodologies and characteristics of NMAs that include NRE could help further the discussion on the appropriateness of including NRE in NMAs using different methods and broaden the understanding of how often different methods are used and in what contexts.
- The incorporation of NRE in NMAs should require careful thought and consideration. Heterogeneity, consistency, and transitivity are key assumptions that should also be accounted for when including NRE. Authors should also explore and compare different methods of including NRE. Owing to growing interest of using NRE, future tutorials and training should include guidance on statistical methods to including NRE in NMAs.

(NRE) into NMAs [5]. This has largely been driven by increasing demand for the incorporation of “real-world evidence,” generated from observational studies or patient

registries, to inform medical evidence and decisions [5,6]. The inclusion of NRE can address shortcomings of RCT evidence including lower generalizability, shorter follow-up, smaller sample sizes, restrictive study populations, and impractical or unethical conduct [7].

1.2. Methods for incorporating NRE

NRE is subject to various biases such as confounding [8], and the inclusion of NRE into a network may produce biased estimates if these biases are not accounted for. The validity of including both RCTs and NRE into a network rests on the assumption of consistency and transitivity across studies. Therefore, it is important to investigate these aspects when including NRE. Definitions for these aspects are presented in [Appendix A](#).

This scoping review will focus on the three main approaches to incorporating NRE into NMAs as described in a commentary article by Cameron et al. [7] and a methodology article by Schmitz et al. [6]. Naïve pooling is the simplest method for combining both randomized evidence and NRE [6,7]. This approach incorporates evidence from all included studies ignoring differences in study designs. A second approach involves using a Bayesian framework in which NRE is incorporated into an NMA as prior information. This process involves analyzing data from NRE separately and configuring the model such that data from these studies are less influential than data from RCTs [6,7,9]. Third, a three-level Bayesian hierarchical model can be used to explicitly examine differences in study designs by modeling the two sources of evidence (RCT and NRE) separately and creating estimates at the level of specific study designs. These estimates are then combined to synthesize an overall measure of treatment effect [6,7,9].

Although other groups have reviewed statistical methods on the incorporation of NRE in NMAs [7,9], to our knowledge, no group has empirically evaluated this practice in the published NMA literature. Our objective is to provide an overview of characteristics specific to NMA's that include NRE, to describe different methodologies for incorporating NRE, how often and in what context a method is used, and to present future directions for combining both RCTs and NRE.

2. Methods

2.1. Study protocol

The methods for this scoping review were developed by Petropoulou et al. [10] and Zarin et al. [11], and the methodological framework of this review was proposed by Arksey et al. [12].

Our data set of NMAs was derived from a previously completed search of NMAs published between January 1999 and April 2015 [10]. Our evidence base comprised a total of 116 studies that were excluded from the study by Petropoulou et al. [10]. We assessed these 116 studies to determine eligibility

in our review. From these studies, a subset of 83 NMAs incorporated NRE in them, either as a primary or secondary analysis. An additional 33 studies were excluded for reasons other than the incorporation of NRE. Our database was supplemented by searching the 456 NMAs included by Petropoulou et al. [10] for sensitivity or subgroup analyses that include both RCTs and NRE. The references of the final set of included studies were also searched.

2.2. Inclusion criteria

The inclusion criteria were based on the studies by Petropoulou et al. [10] and Zarin et al. [11], with the alteration of including NMAs that combined both randomized evidence and NRE. We included NMAs that compared at least four interventions that were defined as medical, safety, or social interventions from any study design. We considered quasi-RCT, controlled before and after, cohort, case-control, and cross-sectional studies as NRE. We excluded networks that included animal studies, examined diagnostic test accuracies, and where the number of studies was smaller than the number of interventions or treatments. NMAs that included only NRE were also excluded. Owing to limited information and resources, conference abstracts for which a full-text article was unavailable or studies published in a language other than English were excluded.

2.3. Screening

Using the inclusion criteria, a pilot test was performed by two independent reviewers (K.Z. and N.T.) on a random sample of 10 studies. Subsequently, the titles and abstracts from the excluded 116 NMAs were screened independently by both reviewers, and screening results were compared. The full-text articles of potentially relevant abstract citations were then screened in the same manner. Discrepancies were discussed between reviewers and if necessary resolved by a third reviewer. Authors of conference abstracts were contacted to access the full-text article. If available, protocol proposals were replaced with full-text publications, and the full-text publications were used for inclusion, even if they were published past the date of the literature search (April 2015). If the protocol did not result in a full-text publication, the protocol study was included. The previous database [10] excluded protocol studies; we included these studies to increase comprehensiveness.

2.4. Data abstraction and data items

The data abstraction form was created in Microsoft Excel and developed as an extension to the previous review [10]. Data abstraction was performed in duplicate. Each reviewer extracted data for all included studies. A pilot test was performed on a random sample of three studies. Abstraction forms were compared and discrepancies were resolved by a third reviewer (P.A.). An extraction form template is provided in Appendix G. Data were extracted for the primary

outcome. If the primary outcome did not incorporate NRE, the network for the secondary outcome was used. In cases where both the primary outcome and secondary outcomes did not use NRE, the sensitivity or subgroup analyses for the NMA was used. Other methods used in data abstraction have been discussed elsewhere [10,11].

We abstracted data items specific to NRE, such as whether the studies provided justification for the inclusion of NRE, total number of nonrandomized studies in the network, method used for incorporating NRE (such as naïve pooling, informative priors, Bayesian 3-level hierarchical model, all methods), whether the study clearly reported the method used and if a scale or tool was used for bias assessment. We contacted the corresponding authors through email to verify the method of incorporating randomized evidence and NRE in the NMAs. A second author (A.A.V.) confirmed the statistical methodology when authors did not respond. Data on reporting characteristics (inconsistency, heterogeneity, and transitivity) grouped by method of NRE incorporation were also abstracted to increase reporting specificity.

2.5. Statistical analyses

Data analysis included both quantitative and qualitative components. We summarized the findings of network characteristics and the methods used from the eligible studies using frequencies and percentages. We reported the qualitative results through narrative synthesis. Graphs were created in the R 3.5.0. software using the *ggplot2* package and descriptive statistics were generated in Microsoft Excel.

3. Results

3.1. Screening results

The study selection process is presented in Figure 1. After screening titles and abstracts of the 116 studies [10], 70 relevant abstracts were identified and the full-text publications were retrieved to assess eligibility. From these 70 abstracts, 14 were conference abstracts of which only 1 had an accessible full-text publication. A total of 17 full-text publications met the inclusion criteria [5,6,13–28]. Of these 17 articles, 3 protocol proposals were replaced with the published studies if they were available; one protocol study resulted in two publications [27,28], and the final article for one protocol proposal [20] had not been published yet. Five networks [29–33] were identified from screening the 456 NMAs from the study by Petropoulou et al. [10], and one network [6] was identified from scanning the references of the included studies. Overall, 23 published NMAs [5,6,13–33] met the inclusion criteria.

3.2. Study characteristics and NMA characteristics

Information related to geographic region, knowledge synthesis approach, and funding is presented in Table 1.

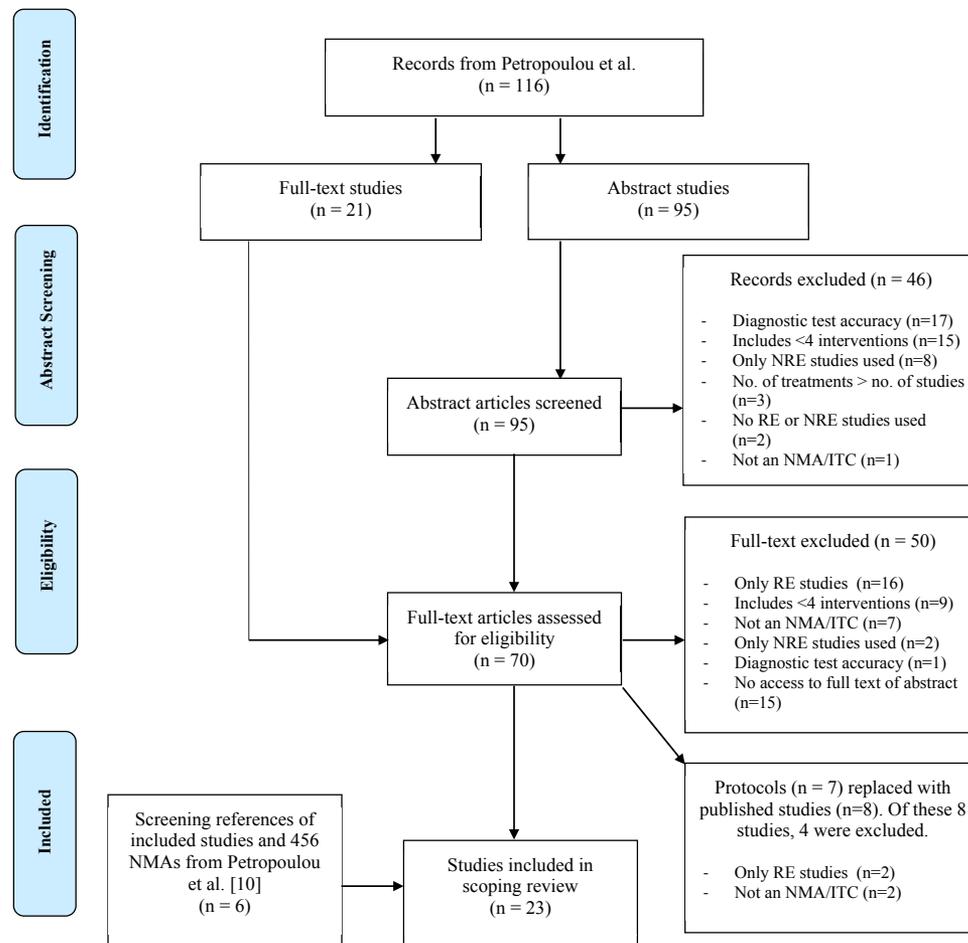


Fig. 1. PRISMA flow diagram for the study selection process of included networks. NRE, nonrandomized evidence; NMA, network meta-analyses; ITC, indirect treatment comparison; RE, randomized evidence.

Abstracted data related to closed loops (“full” networks) or absence of a closed loop (“star” network) and outcome types are reported in [Appendix C](#).

3.3. NMA model validity and reporting

The assessment of inconsistency was possible for 21 NMAs that contained at least one closed loop. Of these 21 NMAs that included NRE, just over half ($n = 13$, 57%) evaluated consistency within the network through node-splitting, random-effects metaregression, loop-specific, and/or design-by-treatment interaction models. The assessment of transitivity was reported by eight (35%) NMAs.

Nineteen (83%) included networks investigated between-study heterogeneity by performing subgroup or sensitivity analyses for the purposes of excluding the NRE to compare effect estimates with the NMA that included RCTs alone, as well as RCTs in addition to NRE. The random-effects model was the most popular method used for conducting NMAs. Thirteen (57%) networks used a random-effects model, zero networks used a

fixed-effects model, and six (26%) used both models. The most commonly used software to combine NRE and RCT data was WinBUGS/OpenBUGS ($n = 13$, 57%), followed by STATA ($n = 5$, 22%). [Appendix D](#) presents a summary of reporting characteristics.

3.4. Methods for incorporating NRE

Authors of 17 of the 23 articles provided supplementary information on the specific method of combining randomized evidence with NRE in their NMA. Used by seventeen studies (74%), naïve pooling was the most common method for incorporating both randomized evidence and NRE in an NMA. Nine percent ($n = 2$) used the Bayesian 3-level hierarchical model, none incorporated NRE using informative priors, two (9%) used all three methods, and two (9%) used “other” approaches. Most studies provided justification on why NRE was included in the network ($n = 19$, 83%). However, only four (17%) of these studies explicitly reported the method used for including both types of evidence, and three (13%) NMAs provided justification for why a method was used. The most common rationale

Table 1. Characteristics of included studies between 2007 and 2018

Author	Year of publication	Geographic region	Knowledge synthesis approach	Funding
Akbar et al. [24]	2013	North America	Systematic review	Not reported
Chen et al ^b [31]	2012	Europe	Systematic review	Publicly sponsored
Cooper et al ^b [32]	2012	Europe	Not reported	Publicly sponsored
Cota et al. [23]	2013	Central and South America	Systematic review	Publicly sponsored
Hubbard et al. [22]	2015	Europe	Not reported	Publicly sponsored
Hutton et al ^b [21]	2012	North America	Systematic review	Nonsponsored
Hutton et al. [20]	2014	North America	Systematic review	Publicly sponsored
Lewis et al ^b [18]	2011	Europe	Systematic review	Publicly sponsored
Lewis et al ^b [33]	2015	Europe	Systematic review	Publicly sponsored
Jenkins et al. [5]	2014	Europe	Systematic review	Industry and publicly sponsored
Kruidenier et al ^b [19]	2012	Europe	Not reported	Not reported
Pollock et al ^b [30]	2007	Europe	Systematic review	Publicly sponsored
Schmitz et al. [6]	2013	Europe	Systematic review	Not reported
Schwendicke et al ^b [16]	2015	Europe	Systematic review	Publicly sponsored
Schwendicke et al ^b [17]	2015	Europe	Systematic review	Publicly sponsored
Thom et al. [15]	2015	Europe	Systematic review	Industry sponsored
Tricco et al ^{a,b} [27]	2015	North America	Systematic review	Publicly sponsored
Tricco et al ^{a,b} [28]	2016	North America	Systematic review	Publicly sponsored
Tricco et al ^{a,b} [26]	2018	North America	Systematic review	Publicly sponsored
Veroniki et al ^a [25]	2017	North America	Systematic review	Publicly sponsored
Wang et al. [14]	2012	North America	Not reported	Publicly sponsored
Wilhelmus et al ^b [29]	2015	North America	Systematic review	Publicly sponsored
Zheng et al ^b [13]	2014	Australia and New Zealand	Systematic review	Publicly sponsored

Literature search was conducted between January 1999 and April 2015.

^a Protocol proposals replaced with published studies (published after April 2015). Refer to [Methods](#).

^b Studies that used naïve pooling but supplemented with additional analyses to investigate the effect of study design.

provided for including NRE in NMAs was to increase generalizability. Other reasons included the limited amount of direct evidence from RCTs, evaluation of rare outcomes, to allow for a wider range of evidence, and to evaluate the robustness of the results. Sixteen (70%) studies used a scale or tool to assess for bias in observational studies, six (26%) studies did not use any tools, and one (4%) study was unclear.

Figure 2, Table 2, and Appendix E present the methods, characteristics, and justifications of incorporating NRE. The relationship between year of publication and method used is presented Figure 3, and the proportion of non-randomized studies to all studies stratified by year of publication is presented in Appendix F. In addition, reporting of consistency, transitivity, and analyses for heterogeneity stratified by method of NRE incorporation is presented in Table 3.

4. Discussion

We conducted a scoping review using an existing empirical data set [10,11] of published NMAs to describe

characteristics and methodologies of networks that combined both randomized evidence and NRE.

Four percent (23/572) of studies included NRE in an NMA, suggesting that this is not a widely used practice. The earliest NMA that met our inclusion criteria was published in 2007 [30]. Seven NMAs (9%) were published in 2015, which suggests that the incorporation of NRE into NMAs is becoming more prevalent in the published literature. The three-level Bayesian hierarchical model was introduced in 2000 to combine different study designs [34]. However, in our database, this model was not used until 2013 to combine randomized evidence and NRE.

Most included NMAs were conducted by combining randomized evidence and NRE through naïve pooling. This method makes strong assumptions by treating all study designs the same. This model does not downweight study designs appraised as “low quality” or make bias adjustments introduced by NRE [6]. Despite its drawbacks, the naïve pooling method is attractive in practice because of its ease of implementation as a simpler and standard model [7] and can be supplemented with subgroup or sensitivity analyses to assess robustness. Naïve pooling can be advantageous by improving network connectivity and to assess consistency.

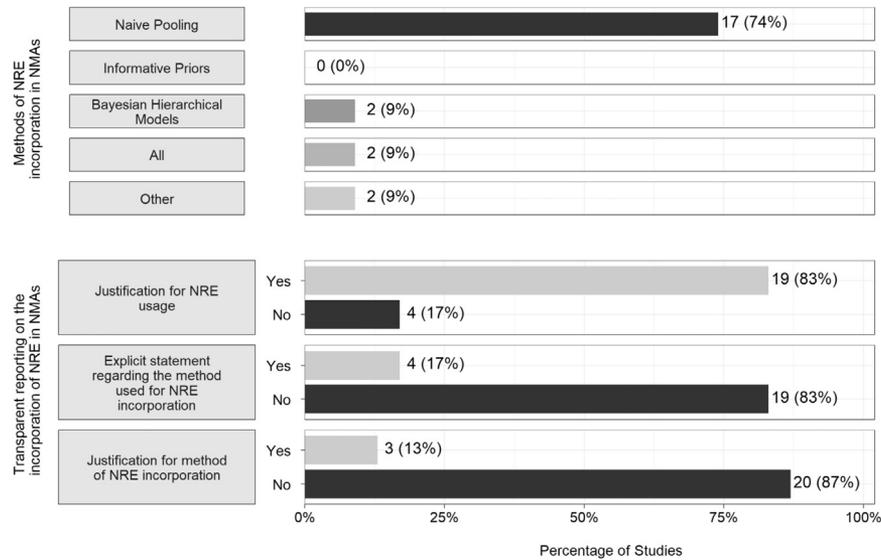


Fig. 2. Frequency of various methods used to incorporate NRE into NMAs. NRE, nonrandomized evidence; NMA, network meta-analysis.

In the study by Thom et al. [15], a closed loop within the network is present because of the inclusion of evidence from observational studies. The network would be disconnected in the absence of NRE, that is, if it had been restricted to solely RCTs.

NRE was not included in NMAs as prior information, with the exception of studies ($n = 2$, 9%) that used all methods to combine randomized evidence and NRE [5,6]. This method is advantageous over naïve pooling, as it allows for the adjustment of potential biases because of the study design by downweighting observational study data [6,9]. However, between-study heterogeneity including all potential study designs is not accounted for. A 3-level Bayesian hierarchical model was used in two (9%) NMAs. This method is able to include different forms of evidence by accounting for between-study heterogeneity, both within each study design and across varied study designs [5,6,35]. The 3-level Bayesian hierarchical model has been previously applied in clinical settings; more weight was assigned to the randomized evidence compared with other studies, based on the premise that RCTs might be less biased than estimates from observational studies [35–37]. These NMAs were also adjusted for bias introduced by NRE. In the study by Veroniki et al. [25], this method was used to combine both randomized evidence and NRE. However, the number of nonrandomized studies included in the network was 77, whereas the number of RCTs included was one. Each type of study represents a level of “clustering” in the model and because this NMA included only one RCT in the cluster, this method would produce very similar results as naïve pooling. Therefore, this approach is useful in cases where the number of studies per study type is more balanced. A common problem for all three approaches is the ability to maintain the internal and external validities of RCT data and also the ability to adjust for

Table 2. Characteristics of NMAs that included NRE

Characteristics of NMAs ($n = 23$)	Count (%)
Method of incorporating NRE	
Naïve pooling	17 (74%)
Informative priors	0 (0%)
3-Level hierarchical model	2 (9%)
All methods	2 (9%)
Other ^a	2 (9%)
Explicit reporting of method for incorporating NRE	
Yes	4 (17%)
No	19 (83%)
Reporting of justification for the chosen method of incorporating NRE	
Yes	3 (13%)
No	20 (87%)
Description of the chosen method used to incorporate NRE	
Yes	3 (13%)
No	20 (86%)
Reporting of justification for including NRE into the NMA	
Yes	19 (83%)
No	4 (17%)
Using a risk of bias scale/tool to assess bias	
Yes	16 (70%)
No	6 (26%)
Unclear	1 (4%)

Abbreviations: NMA, network meta-analyses; NRE, nonrandomized evidence; IPD, individual patient data.

^a Other methods include metaregression analysis including IPD (available IPD for 10/12 studies) as presented in the study by Hubbard et al. [22] and propensity score matching based on baseline characteristics as presented in the study by Thom et al. [15].

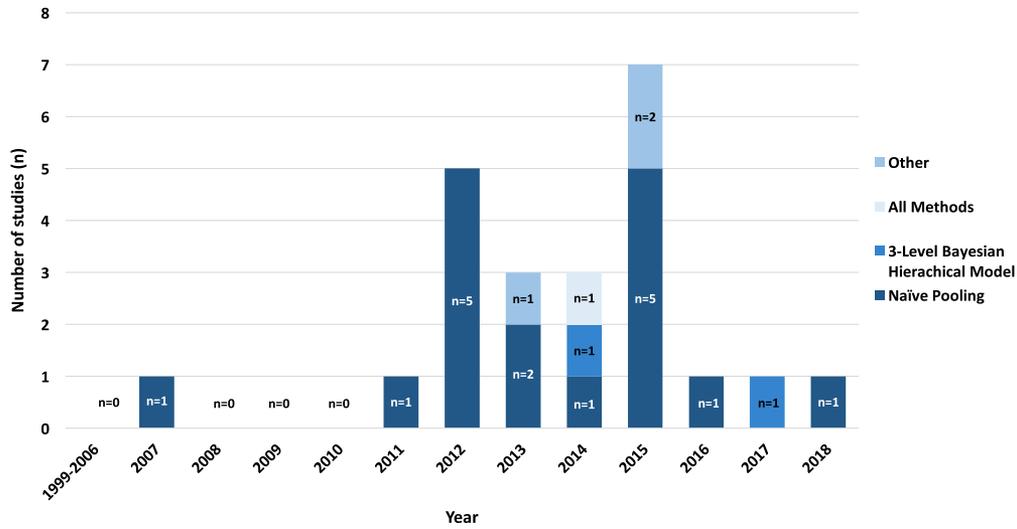


Fig. 3. Method of NRE incorporation stratified by year of publication. Studies published after April 2015 are protocol proposals replaced with published studies. Refer to [Methods](#). One study published in 2015 was a protocol proposal replaced with the published study. NRE, nonrandomized evidence.

Table 3. Reporting of the method used to incorporate NRE grouped by reporting of consistency, transitivity, analysis of heterogeneity, and use of IPD

Naïve pooling	Reporting characteristic	Yes (N)	(%)	No (N)	(%)	Unclear (N)	(%)	Total (N)	Total (%)
	Consistency	8	53%	5	33%	2	13%	15	100%
	Transitivity	6	35%	11	65%	0	0%	17	100%
	Heterogeneity	14	82%	3	18%	0	0%	17	100%
	Use of IPD	0	0%	17	100%	0	0%	17	100%
Information priors									
	Consistency	0	0%	0	0%	0	0%	0	0%
	Transitivity	0	0%	0	0%	0	0%	0	0%
	Heterogeneity	0	0%	0	0%	0	0%	0	0%
	Use of IPD	0	0%	0	0%	0	0%	0	0%
3-Level Bayesian									
	Consistency	2	100%	0	0%	0	0%	2	100%
	Transitivity	1	50%	1	50%	0	0%	2	100%
	Heterogeneity	2	100%	0	0%	0	0%	2	100%
	Use of IPD	0	0%	2	100%	0	0%	2	100%
All methods									
	Consistency	2	100%	0	0%	0	0%	2	100%
	Transitivity	0	0%	2	100%	0	0%	2	100%
	Heterogeneity	1	50%	1	50%	0	0%	2	100%
	Use of IPD	0	0%	2	100%	0	0%	2	100%
Other ^a									
	Consistency	1	50%	1	50%	0	0%	2	100%
	Transitivity	1	50%	1	50%	0	0%	2	100%
	Heterogeneity	2	100%	0	0%	0	0%	2	100%
	Use of IPD	2	100%	0	0%	0	0%	2	100%

Abbreviations: IPD, individual patient data; NRE, nonrandomized evidence.

^a Other methods include metaregression analysis including IPD (available IPD for 10/12 studies) as presented in the study by Hubbard et al. [22] and propensity score matching based on baseline characteristics as presented in the study by Thom et al. [15].

observational bias. Thus, the results of NMAs that include NRE should be interpreted with caution. More than half ($n = 16$, 70%) of the studies used a scale or tool to assess for bias in NRE. The scale that was most commonly applied was the Newcastle Ottawa Scale [38]. Most studies handled the multiplicity of bias by conducting sensitivity or subgroup analyses using only RCTs. However, most studies do not explicitly report the specific method of bias adjustment. In addition, there is no widely accepted statistical software in the literature for the specific combination of randomized evidence and NRE.

Most NMAs ($n = 20$, 87%) provided a rationale for the inclusion of NRE but did not transparently report the method used to combine randomized evidence and NRE. The incorporation of NRE into a network of evidence may violate assumptions about transitivity and exchangeability, and thus, studies should provide reasonable justification and report clearly on the methods used.

The combination of NRE with RCTs in an NMA provides conceptual advantages but poses methodological challenges. A reasonable approach to combining both randomized evidence and NRE into a network includes performing sensitivity or subgroup analyses to investigate sources of heterogeneity, for example, performing an NMA using only randomized evidence and then comparing these results to an NMA with both randomized evidence and NRE included. Naïve pooling can be used as a first-step analysis to synthesize useful insight about the effect of including NRE into the NMA. Exploring the effect of all three methods as demonstrated in the studies by Jenkins et al. [5] and Schmitz et al. [6] would provide useful information on the treatment effect estimates produced by each approach. The evaluation of consistency and transitivity is imperative for an NMA. These two components should be evaluated to ensure that different sources of evidence are compatible and that the underlying source population is similar. Most NMAs ($n = 21$, 91%) in our evidence base did not use individual patient data (IPD) to adjust for patient-level covariates, a common situation as IPD is often not available [39–41]. IPD may be used to explore heterogeneity and inconsistency within a network [42–44], and access to IPD can improve parameter estimation of an NMA [42]. A critical advantage of IPD over study-level data is the ability to estimate separate effects for a patient-level effect modifier and a study-level effect modifier [45].

Our review has some limitations to note. We did not extend this review to unpublished studies, which we expect to be more voluminous than published studies. Thus, it is difficult to achieve an overall view of the most common methods used to combine randomized evidence and NRE in both published and unpublished NMAs. We also did not extend the search to 2016 and 2017, and although this would have broadened our data set, we relied on the existing database from the study by Petropoulou et al. [10], with the expectation that the general pattern would be captured. We were also limited to studies in English as we did not

have the capacity for article translation. However, 1 (< 1%) NMA from screening titles and abstracts from the study by Petropoulou et al. [10] was non-English.

5. Conclusion

The combination of both randomized evidence and NRE will continue to grow in health care because of the absence of direct evidence from RCTs and their limitations. In our evidence base, naïve pooling was most commonly used to incorporate NRE, and although justification for incorporating NRE was commonly provided, explicit reporting on the specific methods used or their justification was rare. The integration of randomized evidence and NRE requires careful consideration because of potential biases that may be introduced from NRE. The incorporation of NRE into NMAs will most likely increase in the future and will require more research to further evaluate the validity of the three approaches discussed in this scoping review.

CRedit authorship contribution statement

Kathryn Zhang: Methodology, Validation, Formal analysis, Writing - original draft, Visualization, Supervision, Project administration. **Paul Arora:** Conceptualization, Methodology, Validation, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Neha Sati:** Writing - review & editing, Visualization. **Audrey Béliveau:** Writing - review & editing. **Natalie Troke:** Methodology. **Areti Angeliki Veroniki:** Writing - review & editing. **Myanca Rodrigues:** Writing - review & editing. **Patricia Rios:** Writing - review & editing. **Wasifa Zarin:** Writing - review & editing. **Andrea C. Tricco:** Writing - review & editing, Funding acquisition.

Acknowledgments

The authors would like to thank Justin J. Slater for input on confirming statistical methods for incorporating nonrandomized evidence into a network meta-analysis and Bruno Kovic for valuable input during the conceptualization of the study.

All authors contributed to all aspects of the manuscript.

Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jclinepi.2019.04.023>.

References

- [1] Sutton A, Abrams K. Bayesian methods in meta-analysis and evidence synthesis. *Stat Methods Med Res* 2001;10(4):277–303.

- [2] Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. *Value Health* 2008;11(5):956–64.
- [3] Caldwell DM. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;331:897–900.
- [4] Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, 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;14(4):417–28.
- [5] Jenkins D, Bujkiewicz S, Martina R, Dequen P, Abrams KR. Methods for the inclusion of real world evidence in network meta-analysis. *Value Health* 2014;17(7):A576.
- [6] Schmitz S, Adams R, Walsh C. Incorporating data from various trial designs into a mixed treatment comparison model. *Stat Med* 2013;32:2935–49.
- [7] Cameron C, Fireman B, Hutton B, Clifford T, Coyle D, Wells G, 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;4(1):1–8.
- [8] Grimes DA, Schulz KF. Bias and causal associations in observational research. *Lancet* 2002;359:248–52.
- [9] Efthimiou O, Mavridis D, Debray TPA, Samara M, Belger M, Siontis GCM, et al. Combining randomized and non-randomized evidence in network meta-analysis. *Stat Med* 2017;36:1210–26.
- [10] Petropoulou M, Nikolakopoulou A, Veroniki AA, Rios P, Vafaei A, Zarin W, et al. Bibliographic study showed improving statistical methodology of network meta-analyses published between 1999 and 2015. *J Clin Epidemiol* 2017;82:20–8.
- [11] Zarin W, Veroniki AA, Nincic V, Vafaei A, Reynen E, Motiwala SS, et al. Characteristics and knowledge synthesis approach for 456 network meta-analyses: a scoping review. *BMC Med* 2017;15(1):1–11.
- [12] Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol Theor Pract* 2005;8(1):19–32.
- [13] Zheng H, Barnett AG, Merollini K, Sutton A, Cooper N, Berend T, et al. Control strategies to prevent total hip replacement-related infections: a systematic review and mixed treatment comparison. *BMJ Open* 2014;4(3):e003978.
- [14] Wang SY, Chu H, Shamliyan T, Jalal H, Kuntz KM, Kane RL, et al. Network meta-analysis of margin threshold for women with ductal carcinoma in situ. *J Natl Cancer Inst* 2012;104:507–16.
- [15] Thom HHZ, Capkun G, Cerulli A, Nixon RM, Howard LS. Network meta-analysis combining individual patient and aggregate data from a mixture of study designs with an application to pulmonary arterial hypertension. *BMC Med Res Methodol* 2015;15:34.
- [16] Schwendicke F, Jäger AM, Paris S, Hsu LY, Tu YK. Treating pit-and-fissure caries: a systematic review and network meta-analysis. *J Dent Res* 2015;94(4):522–33.
- [17] Schwendicke F, Paris S, Tu YK. Effects of using different criteria for caries removal: a systematic review and network meta-analysis. *J Dent* 2015;43(1):1–15.
- [18] Lewis R, Williams N, Matar H, Din N, Fitzsimmons D, Phillips C, et al. The clinical effectiveness and cost-effectiveness of management strategies for sciatica: systematic review and economic model. *Health Technol Assess* 2011;15:1–578.
- [19] Kruidenier LM, Viechtbauer W, Nicolai SP, Büller H, Prins MH, Tejjink JAW. Treatment for intermittent claudication and the effects on walking distance and quality of life. *Vascular* 2012;20(1):20–35.
- [20] Hutton B, Joseph L, Yazdi F, Tetzlaff J, Hersi M, Kokolo M, et al. Checking whether there is an increased risk of post-transplant lymphoproliferative disorder and other cancers with specific modern immunosuppression regimens in renal transplantation: protocol for a network meta-analysis of randomized and observational studies. *Syst Rev* 2014;3(1):1–7.
- [21] Hutton B, Joseph L, Fergusson D, Mazer CD, Shapiro S, Tinmouth A. Risks of harms using antifibrinolytics in cardiac surgery: systematic review and network meta-analysis of randomised and observational studies. *BMJ* 2012;345:e5798.
- [22] Hubbard S, Cooper N, Kendrick D, Young B, Wynn PM, He Z, et al. Network meta-analysis to evaluate the effectiveness of interventions to prevent falls in children under age 5 years. *Inj Prev* 2015;21(2):98–108.
- [23] Cota GF, de Sousa MR, Fereguetti TO, Rabello A. Efficacy of anti-leishmania therapy in visceral leishmaniasis among HIV infected patients: a systematic review with indirect comparison. *PLoS Negl Trop Dis* 2013;7(5):e2195.
- [24] Akbar A, Abu Dayyeh BK, Baron TH, Wang Z, Altayar O, Murad MH. Rectal nonsteroidal anti-inflammatory drugs are superior to pancreatic duct stents in preventing pancreatitis after endoscopic retrograde cholangiopancreatography: a network meta-analysis. *Clin Gastroenterol Hepatol* 2013;11(7):778–83.
- [25] Veroniki AA, Cogo E, Rios P, Straus SE, Finkelstein Y, Kealey R, et al. Comparative safety of anti-epileptic drugs during pregnancy: a systematic review and network meta-analysis of congenital malformations and prenatal outcomes. *BMC Med* 2017;15(1):1–20.
- [26] Tricco AC, Ashoor HM, Soobiah C, Rios P, Veroniki AA, Hamid JS, et al. Comparative effectiveness and safety of cognitive enhancers for treating Alzheimer's disease: systematic review and network meta-analysis. *J Am Geriatr Soc* 2018;66(1):170–8.
- [27] Tricco AC, Soobiah C, Blondal E, Veroniki AA, Khan PA, Vafaei A, et al. Comparative efficacy of serotonin (5-HT₃) receptor antagonists in patients undergoing surgery: a systematic review and network meta-analysis. *BMC Med* 2015;13(1):1–13.
- [28] Tricco AC, Blondal E, Veroniki AA, Soobiah C, Vafaei A, Ivory J, et al. Comparative safety and effectiveness of serotonin receptor antagonists in patients undergoing chemotherapy: a systematic review and network meta-analysis. *BMC Med* 2016;14(1):1–16.
- [29] Wilhelmus KR. Antiviral treatment and other therapeutic interventions for herpes simplex virus epithelial keratitis (Review). *Cochrane Database Syst Rev* 2011:1–305.
- [30] Pollock A, Baer G, Pomeroy V, Langhorne P. Physiotherapy treatment approaches for the recovery of postural control and lower limb function following stroke. *Cochrane Database Syst Rev* 2003:1–94.
- [31] Chen Y-FF, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. *Health Technol Assess* 2012;16:1–205.
- [32] Cooper NJ, Kendrick D, Achana F, Dhiman P, He Z, Wynn P, et al. Network meta-analysis to evaluate the effectiveness of interventions to increase the uptake of smoke alarms. *Epidemiol Rev* 2012;34(1):32–45.
- [33] Lewis RA, Williams NH, Sutton AJ, Burton K, Din NU, Matar HE, et al. Comparative clinical effectiveness of management strategies for sciatica: systematic review and network meta-analyses. *Spine J* 2015;15(6):1461–77.
- [34] Prevost TC, Abrams KR, Jones DR. Hierarchical models in generalized synthesis of evidence: an example based on studies of breast cancer screening. *Stat Med* 2000;19:3359–76.
- [35] Verde PE, Ohmann C. Combining randomized and nonrandomized evidence in clinical research: a review of methods and applications. *Res Synth Methods* 2015;6(1):45–62.
- [36] Mesgarpour B, Heidinger BH, Schwameis M, Kienbacher C, Walsh C, Schmitz S, et al. Safety of off-label erythropoiesis stimulating agents in critically ill patients: a meta-analysis. *Intensive Care Med* 2013;39(11):1896–908.
- [37] Bittl JA, He Y, Jacobs AK, Yancy CW, Normand S-LT. Bayesian methods affirm the use of percutaneous coronary intervention to improve survival in patients with unprotected left main coronary artery disease. *Circulation* 2013;127:2177–85.
- [38] Wells GA, O'Connell D. The Newcastle-Scale for assessing the quality of nonrandomised studies in meta-analyses. Ottawa: Ottawa Hospital

- Research Institute; 2011. Available at http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
- [39] Ravva P, Karlsson MO, French JL. A linearization approach for the model-based analysis of combined aggregate and individual patient data. *Stat Med* 2014;33:1460–76.
- [40] Donegan S, Williamson P, D'Alessandro U, Garner P, Smith CT. Combining individual patient data and aggregate data in mixed treatment comparison meta-analysis: individual patient data may be beneficial if only for a subset of trials. *Stat Med* 2013;32:914–30.
- [41] Riley RD, Simmonds MC, Look MP. Evidence synthesis combining individual patient data and aggregate data: a systematic review identified current practice and possible methods. *J Clin Epidemiol* 2007; 60:431–9.
- [42] Saramago P, Sutton AJ, Cooper NJ, Manca A. Mixed treatment comparisons using aggregate and individual participant level data. *Stat Med* 2012;31:3516–36.
- [43] Veroniki AA, Straus SE, Soobiah C, Elliott MJ, Tricco AC. A scoping review of indirect comparison methods and applications using individual patient data. *BMC Med Res Methodol* 2016;16:47.
- [44] Donegan S, Williamson P, D'Alessandro U, Smith CT. Assessing the consistency assumption by exploring treatment by covariate interactions in mixed treatment comparison meta-analysis: individual patient-level covariates versus aggregate trial-level covariates. *Stat Med* 2012;31:3840–57.
- [45] Jansen JP. Network meta-analysis of individual and aggregate level data. *Res Synth Methods* 2012;3(2):177–90.