View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Explore Bristol Research

Molloy, GJ., Noone, C., Caldwell, D., Welton, N., & Newell, J. (2018). Network meta-analysis in health psychology and behavioural medicine: a primer. *Health Psychology Review*. https://doi.org/10.1080/17437199.2018.1457449

Peer reviewed version

Link to published version (if available):
10.1080/17437199.2018.1457449

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Taylor & Francis at https://www.tandfonline.com/doi/full/10.1080/17437199.2018.1457449. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms

For Peer Review

Network Meta-Analysis in Health Psychology and Behavioural Medicine: A Primer

Journal:	*Health Psychology Review*
Manuscript ID	RHPR-2017-0097.R3
Manuscript Type:	Conceptual Review
Keywords:	evidence synthesis, policy-making, meta-analysis, health behaviour change
Abstract:	Progress in the science and practice of health psychology depends on the systematic synthesis of quantitative psychological evidence. Meta-analyses of experimental studies have led to important advances in understanding health-related behaviour change interventions. Fundamental questions regarding such interventions have been systematically investigated through synthesising relevant experimental evidence using standard pairwise meta-analytic procedures that provide reliable estimates of the magnitude, homogeneity and potential biases in effects observed. However, these syntheses only provide information about whether particular types of interventions work better than a control condition or specific alternative approaches. To increase the impact of health psychology on health-related policy-making, evidence regarding the comparative efficacy of all relevant intervention approaches – which may include biomedical approaches - is necessary. With the development of network meta-analysis, such evidence can be synthesised, even when direct head-to-head trials do not exist. However, care must be taken in its application to ensure reliable estimates of the effect sizes between interventions are revealed. This review paper describes the potential importance of network meta-analysis to health psychology, how the technique works and important considerations for its appropriate application within health psychology.

SCHOLARONE™
Manuscripts

Abstract

Progress in the science and practice of health psychology depends on the systematic synthesis of quantitative psychological evidence. Meta-analyses of experimental studies have led to important advances in understanding health-related behaviour change interventions. Fundamental questions regarding such interventions have been systematically investigated through synthesising relevant experimental evidence using standard pairwise meta-analytic procedures that provide reliable estimates of the magnitude, homogeneity and potential biases in effects observed. However, these syntheses only provide information about whether particular types of interventions work better than a control condition or specific alternative approaches. To increase the impact of health psychology on health-related policy-making, evidence regarding the comparative efficacy of all relevant intervention approaches – which may include biomedical approaches - is necessary. With the development of network meta-analysis, such evidence can be synthesised, even when direct head-to-head trials do not exist. However, care must be taken in its application to ensure reliable estimates of the effect sizes between interventions are revealed. This review paper describes the potential importance of network meta-analysis to health psychology, how the technique works and important considerations for its appropriate application within health psychology.

Keywords: evidence synthesis, policy-making, meta-analysis, health behaviour change

For Peer Review Only

Introduction

Progressing the science and practice of health psychology depends on the systematic synthesis of evidence from health behaviour change interventions. In particular, meta-analyses of randomised controlled trials (RCTs) have led to important advances in our understanding of the health impact of health behaviour change interventions. The vast majority of these meta-analyses have involved pairwise comparisons i.e. the comparison of one intervention against another, or against a control condition. However, both national and global health policy organisations are increasingly relying on evidence synthesis involving the comparison of multiple interventions (Kanters et al., 2016).

Indirect comparisons can be made if interventions that have not been directly compared with each other, have been compared to a common alternative intervention (Bucher et al., 1997). More generally, network meta-analysis (NMA) is a tool which enables synthesis of evidence from both direct (i.e. within trial comparisons of randomised groups) and indirect (i.e. between trial) comparisons of multiple interventions that may not have been compared within the same trial (Diaz, Ades, Welton, Jansen & Sutton, 2018; Higgins & Whitehead, 1996; Lu & Ades, 2004). All that is required is that all the trial evidence being quantitatively synthesised has at least one intervention in common with another, as this allows a network of trial comparisons to be constructed. This maximises the use of available evidence, allows comparisons between any pair of interventions in the evidence network, and can increase the precision of the effect size for an intervention, compared with direct evidence alone (Caldwell, Ades, & Higgins, 2005; Ioannidis, 2006; Jansen et al., 2014). It is due to these advantages that NMA has become a key component of the development of clinical guidelines and reimbursement recommendations by national health technology assessment agencies and the World Health Organisation (Kanters et al., 2016). The utility of NMA in clinical medicine has resulted in some scholars suggesting it could constitute a higher level in the hierarchy of

For Peer Review Only

evidence than traditional systematic reviews and pairwise meta-analyses (Leucht et al., 2016; Roever & Biondi-Zoccai, 2016). However, while there has been a significant and rapid increase in the use of the method in health research more broadly over the last 10 years (Lee, 2014), uptake in the field of Health Psychology has been more limited. For example, a search of the present journal which is one of the internationally leading review journals in the discipline, as indicated by impact factor (7.24 in 2016), identified no instances use of NMA over the last 10 years. The application of NMA in health psychology has the potential to strengthen the link between evidence from behavioural trials in health and healthcare decision-making. This paper describes the potential importance of this method of evidence synthesis to health psychology, how the technique works and important considerations for its appropriate application within health psychology.

Why Network Meta-Analysis is Useful

In health psychology a considerable evidence base has been established on the effects of a wide variety of interventions for behaviour change on health. For a given patient population, there are typically several interventions available, and practitioners need to make evidence-based decisions between them. Ideally, this evidence would take the form of a well-powered RCT with as many intervention arms as there are decision options. However, it is clearly not feasible to conduct such a study, as the complexity of the study design and the resources required would be too great (Catalá-López, Aurelio, Cameron, Moher, & Hutton, 2014). For example, whereas several types of behaviour change interventions are known to be effective in reducing blood pressure, including increased physical activity, smoking cessation and dietary modifications (Mancia et al., 2013), it would be impractical to attempt implementing even one multi-arm RCT that compared the effects of changes to one of these behaviours on blood pressure, let alone an RCT that compared the different techniques used to change each of these behaviours (Grant & Calderbank-Batista, 2013). Furthermore, even if

For Peer Review Only

such complex studies could be conducted, the pairwise evidence synthesis methods normally employed in health psychology could not coherently synthesise their results.

The current evidence base for the efficacy of behavioural interventions is mostly formed from studies comparing specific types of behavioural interventions with a control condition, such as wait-list or treatment-as-usual, and occasional examples of trials evaluating competing or alternative behavioural interventions, tested against each other (Michie, Abraham, Whittington, McAteer & Gupta, 2009). There are no examples of trials comparing every possible type of behavioural intervention for a given population, illness and outcome being simultaneously evaluated against one another. Additionally, "treatment as usual" can be very different across studies, as can the behavioural interventions themselves (Oberjé, Dima, Pijnappel, Prins, & de Bruin, 2015). If ignored, this intervention-level variation can lead to high levels of heterogeneity when pooled in a meta-analysis (de Bruin, Viechtbauer, Hospers, Schaalma, & Kok, 2009). The result of working with this kind of evidence base is a tendency to rely on expert opinion in deciding what interventions to implement (Kanters et al., 2016). NMA can treat each type of control condition as a distinct intervention, and similarly for behavioural interventions with different characteristics or components, hence minimising heterogeneity.

Additionally, many health outcomes targeted by health behaviour change interventions (e.g. blood pressure reduction) are often managed, first, through medical treatment (e.g. anti-hypertensive medication). Typically, behavioural interventions are not included as comparators in clinical trials of medical interventions, as regulatory bodies only require that they be compared with placebo conditions or treatment-as-usual/standard care (Falissard et al., 2009; Sutton & Higgins, 2008; Song, Altman, Glenny, & Deeks, 2003). For example, there is very limited evidence comparing physical activity interventions to drug interventions in those with illnesses related to cardiovascular disease, as this is often not

For Peer Review Only

required for licensing (Naci & Ioannidis, 2013). Thus, to make better-informed healthcare decisions, evidence regarding the comparative efficacy of *all* available interventions, whether behavioural or medical, is required. The absence of such comparison is critical. If behavioural interventions are as effective and cost-effective as medical treatments for a given illness or if they provide clinically important amplifications to medical treatment, then the likelihood for policy change that promotes the practice of health psychology and behavioural medicine will be enhanced (Jansen et al., 2011). This can highlight future directions for confirmatory research and provide greater scientific justification for the design and implementation of RCTs (Meulemeester et al., 2018).

In summary, current decision-making regarding interventions in health psychology is limited, because only evidence-based claims about *what works* can be made, rather than *what works best* (Salanti, 2012). The emergence of better comparative evidence on what interventions work best is critical for the further development of health psychology in healthcare. Network meta-analysis provides a methodology to achieve this and therefore has the potential to elevate both the science and practice of health psychology and behavioural medicine from its current status as a relatively minor component in the delivery of healthcare globally (Cheung & Hong, 2017). Despite its potential to transform the field, NMA has yet to be fully embraced by health psychology and behavioural science more broadly. As a relatively new evidence synthesis method, NMA is rarely a standard part of postgraduate training in health psychology, therefore the requisite knowledge and skills do not typically exist within this discipline.

Next, we provide a brief primer on the essential concepts which must be understood in order to conduct a NMA. See table 1 for a description of some key terms related to NMA.

For Peer Review Only

Table 1. Key terms related to NMA.

[Insert table 1 here]

How Network Meta-Analysis Works

The simplest application of NMA is the comparison of two interventions which are both viable intervention options for a given population, illness and outcome and which have been compared to similar alternative interventions (e.g. treatment-as-usual); but which have not been directly compared. Returning to the example of blood pressure reduction for people with hypertension, consider two broad types of behaviour change interventions which have been found to be effective but which, to our knowledge, have not been compared: increasing physical activity and salt-intake reduction. Interventions within these two categories are typically compared to treatment-as-usual control groups. An indirect comparison between physical activity interventions and salt reduction interventions (see Figure 1 for a *network diagram*) can then be made using the following formula (Bucher, Guyatt, Griffith, & Walter, 1997):

$$\text{Indirect Comparison}_{\text{Physical Activity VS. Salt Reduction}} = \text{Direct Comparison}_{\text{Physical Activity VS. Control Group}} - \text{Direct Comparison}_{\text{Salt Reduction VS. Control Group}}$$

Note that this assumes that the control group is similar in the Physical Activity studies to the control group employed in the Salt Reduction studies.

More generally for interventions A, B, and C, the indirect comparison can be presented as:

$$\hat{\mu}_{AB}^{Ind} = \hat{\mu}_{AC}^{Dir} - \hat{\mu}_{BC}^{Dir}$$

where $\hat{\mu}_{AB}^{Ind}$ is the indirect estimate of B vs A, $\hat{\mu}_{AC}^{Dir}$ is the direct estimate of C vs A, and $\hat{\mu}_{BC}^{Dir}$ is the direct estimate of C vs B.

For Peer Review Only

The variance of this estimate is equal to the sum of the variances of each of the direct estimates, meaning the indirect comparison alone is less precise than either of the direct estimates.

[Insert Figure 1 here]

Figure 1. An example of a network diagram.

The network represented in Figure 1 is usually referred to as a simple indirect comparison. A simple indirect comparison can be extended to include any number of interventions which have been previously tested against a single common comparator. Panel B of figure 2 provides an example of a network with four competing interventions, each of which has been compared to the common comparator intervention 'A'. This 'star' network of evidence is likely to be common in health psychology, where behavioural interventions are most often compared to treatment-as-usual (de Bruin et al. 2009; Mohr, Freedland, & Beckner, 2009). Of course, care should be taken to ensure that each treatment-as-usual intervention is similar enough across the studies to be combined into a single comparator 'node'.

[Insert Figure 2 here]

Figure 2. Some possible configurations of networks of evidence.

A 'star' network can readily be extended to include further comparisons. These can be interventions which have been compared to specific interventions present in the network i.e. they do not need to be connected via a single common comparator. There will many such situations in health psychology where more than one common comparator exists; for example, whereas many studies employ a waitlist control, some studies employ an active control group. The hypothetical evidence network depicted in panel C of figure 2 represents this situation, where A could be a waitlist control group, B to E could be competing

For Peer Review Only

interventions and F could be an active control group which has been included in trials of B and E. This network also demonstrates a *closed loop*, where there is both direct and indirect evidence available to inform the comparison conditions A and B and conditions A and E.

Panel D of figure 2 depicts another hypothetical evidence network that may arise in health psychology, where both behavioural and medical interventions are compared. How these two sources of evidence are connected will depend on the population, illness and outcome that is being investigated. In this example, we imagine a treatment-as-usual comparator, common to both behavioural and medical intervention studies, as represented by condition A. Again, the behavioural interventions are represented by conditions B to E, with condition F representing an active behavioural control group. In this example, conditions G to I represent medical interventions that have been compared to both treatment-as-usual (A) and a placebo condition (J). Still, the evidence networks which are most likely to be well connected are those where several behavioural interventions which target the same outcome are being compared. A hypothetical example can be seen in Panel E of figure 2.

It is also possible that there might be no *single* common comparator connecting all available interventions, for a given health outcome (Goring et al., 2016). For example, behavioural interventions can be compared to waitlist control groups, behavioural active control groups or treatment-as-usual, whilst medical interventions might only be compared to placebo control groups. If there is direct evidence comparing behavioural interventions directly with medical interventions, then the network "connects" and NMA can be performed. If not, then the network is disconnected (Goring et al., 2016). Standard NMA techniques cannot be applied to disconnected networks unless the different types of control can be considered similar enough to "lump" together and connect the network. A recent example of this is a health technology assessment of smoking cessation interventions (Health Information and Quality Authority, 2017). Behavioural interventions and pharmacological interventions

For Peer Review Only

were analysed separately because there were systematic differences in the nature and effects of the control groups used in trials of these two types of intervention. Some extensions of network meta-analysis have been proposed which can analyse disconnected networks but these rely on extra assumptions (Goring et al., 2016).

To estimate the indirect comparisons in the more complex networks that may emerge in the synthesis of evidence from behavioural interventions that are typically studied in the health psychology literature, additional modern statistical models such as NMA are required (Dias, Ades, Welton, Jansen, & Sutton, 2018). Such methods produce more precise effect sizes, than using direct evidence alone (Caldwell et al., 2005; Ioannidis, 2006; Jansen et al., 2014). However, for all NMA models there are some key assumptions that must be met to ensure the resulting effect size estimates are meaningful.

Assumptions of Network Meta-Analysis

In NMA, as in pairwise meta-analysis, care must be taken to estimate and account for heterogeneity. Heterogeneity across a set of studies implies the presence of effect modifiers, examples of which may include: participant characteristics at baseline; intervention dosages; intervention setting; type and timing of measurements, among others. However, these effect modifiers may or may not be measured or even measurable. If measurable and measured, a trial-level variable is shown to be an effect modifier when it interacts significantly with the intervention effect (Dias, Welton, Sutton, & Ades, 2013). Critically, estimates of the effect sizes from NMA can be confounded by the uneven distribution of effect modifiers across the network of evidence (Kovic et al., 2017). This is an example of the violation of the key assumption underpinning NMA, which can be considered in two parts *(i)* transitivity and *(ii)* consistency.

For Peer Review Only

According to Salanti (2012, p.83), *transitivity* refers to the assumption that the "indirect comparison validly estimates the unobserved head-to-head comparison". It should be possible, in principle, that participants could be randomised to any of the interventions included in the evidence network in a hypothetical RCT (Salanti, 2012). For example, receiving one kind of intervention technique should not mean that another one is contraindicated. *Consistency* is the term used for the statistical manifestation of transitivity, and can only be assessed when both direct and indirect evidence is available. Estimates in a NMA are said to be consistent when the indirect evidence and the direct evidence agrees. Checking that the conditions for both transitivity and statistical consistency are met is an essential step in running a NMA, where evidence is available from both direct and indirect sources (see the following for a detailed description of strategies for checking consistency; Dias et al, 2013; Higgins et al., 2012; White, Barrett, Jackson, & Higgins, 2012). However, when direct evidence is absent, and a statistical check of consistency is therefore not possible, transitivity must still be assessed. It is always possible to check for transitivity, regardless of whether direct evidence is available or not. This can be achieved by qualitatively examining relevant clinical and methodological aspects of the relevant intervention comparators to ascertain whether there is an even distribution of clinical and methodological effect modifiers across the intervention comparators (Diaz, Ades, Welton, Jansen & Sutton, 2018) .

The assumption of transitivity is crucial to the validity of the results of any NMA as the violation of this assumption leads to biased indirect comparison estimates, which leads to biased NMA estimates (i.e. the estimates which integrate both direct and indirect evidence; Jansen & Naci, 2013). The next section discusses specific challenges which may arise in applying NMA in health psychology. These challenges may affect the validity with which health behaviour change intervention studies can be synthesised by NMA.

For Peer Review Only

Challenges in Applying Network Meta-analysis in Health Psychology

Although there are many potential benefits of using NMA in health psychology, particular care must be taken in comparing multiple behavioural interventions, as there may be important differences in the reasons why a particular behaviour is being targeted or why a particular set of behaviour change techniques (BCTs) is being used, or additionally why a specific comparator is chosen.

Choosing to change a specific health behaviour and applying specific BCTs to achieve this involves careful development work that considers patient characteristics, available resources and contextual factors (Bartholomew, Parcel, & Kok, 1998; Michie, van Stralen, & West, 2011). Each decision in the intervention development process has the potential to modify the intervention effect. Therefore, in applying NMA in health psychology, researchers must examine how each intervention in the evidence network was developed, in order to ensure transitivity and consistency. Combining behavioural interventions that apply multiple interacting BCTs in different ways, across different settings, and with different patient groups, has the potential to violate transitivity if there is an uneven distribution of clinical and methodological characteristics across the set of interventions being analysed. Therefore, we strongly recommend that this methodology is only used when there is appropriate statistical and clinical expertise within the review team, as this is essential to apply this method appropriately.

Researchers should also consider the type of control groups used in testing different interventions, which may include the application of some BCTs, and which may be unevenly distributed across control conditions (de Bruin et ak., 2009). This is a considerable threat to the assumption of transitivity and one that is difficult to identify due to the poor reporting of the contents of control conditions (Oberjé et al., 2015). However, if the contents of control

For Peer Review Only

conditions are coded carefully rather than lumped together, NMA can be usefully applied to identify how intervention effects differed according to the type of control group employed. Notably, the use of NMA identified different intervention effects for cognitive-behavioural therapy in depression depending on the nature of the control group employed and revealed a possible nocebo effect attributable to waiting-list control groups (Furukawa et al., 2014). Note however, that by creating distinct control group effects, the precision in the summary intervention effect estimates will be reduced.

It is likely that network meta-analyses in health psychology will rely on indirect evidence. This is due to the common practice of comparing interventions to treatment-as-usual rather than suitable alternative, competing interventions (Ayling et al., 2015; Bruin & Viechtbauer, 2014; Freedland, Mohr, Davidson, & Schwartz, 2011; Oberjé et al., 2015). As discussed above, this precludes statistical assessment of consistency. Care must be taken in the design of any NMA in health psychology as a clear definition of the population, interventions, comparators and outcomes (PICO) will enhance the validity of the analysis.

Another characteristic of NMA that may limit its usefulness in health psychology, as in other areas of psychology, is the predominance of small studies (Crutzen & Peters, 2017). These may suffer from methodological limitations usually associated with small sample sizes which can lead to biased estimates (Roever & Biondi-Zoccai, 2016). This issue applies equally to pairwise meta-analysis, but bias can propagate through a network and affect different parts of the network in different ways (Li et al., 2011). NMA would not be recommended in cases where evidence is only available from very small, underpowered trials.

Finally, the suitability of NMA for synthesising evidence in health psychology is expected to improve as existing calls for increased rigour and reproducibility are heeded.

For Peer Review Only

Health psychologists should continue to respond to calls for: better measurement (Beauchamp & McEwan, 2017); increased use of standard outcome sets (Williamson et al., 2012); more transparent reporting of intervention methodology and results (Boutron, Moher, Altman, Schulz, & Ravaud, 2008; Hoffmann et al., 2014); and the compulsory sharing of individual-level data (Peters, Abraham, & Crutzen, 2015).

Opportunities for Network Meta-Analysis in Health Psychology

There are many opportunities to apply NMA and synthesise evidence regarding behavioural intervention for some of the most pressing health problems. Foremost among these include the main behavioural contributors to mortality such as smoking, sedentary behaviour, dietary behaviour, sleep and alcohol consumption. Indeed, there are several recent and ongoing NMAs that aim to elucidate the comparative efficacy of behavioural and medical interventions for addressing health outcomes and related behaviours (Suissa, et al., 2017; Ifikhar et al., 2017; Schwingshackl et al., 2017; Cheng et al., 2017). The increased application of NMA in addressing health relevant behaviours, in recent times, demonstrates that researchers, in a variety of fields, have identified NMA as a potential means of providing both richer syntheses of existing evidence and new insights into whether and which behavioural interventions should be prioritised in healthcare.

Another important area for future development involves linking NMA to other recent developments in meta-analysis, such as spatiotemporal, multivariate, and automated meta-analyses (Card, 2017). The integration of these methods would increase the amount of valuable information contributing to decision-making regarding the comparative effectiveness of health interventions. Specifically, spatiotemporal meta-analysis is a technique designed to account for heterogeneity in research findings due to variability in study environments (Johnson et al., 2017). This approach expands the traditional process of

For Peer Review Only

conducting meta-analysis to include methods for the coding and modelling of geographical and temporal information. Factors related to the timing and location of interventions can be significant effect modifiers. Integrating the spatiotemporal meta-analysis and NMA will therefore allow for more accurate and systematic examination of the assumption of transitivity.

Multivariate meta-analysis is an extension of meta-analysis which allows for the examination of intervention effects for multiple outcomes (Jackson, Riley, & White, 2011). In addition to the primary outcome, studies in health research usually involve several secondary outcomes, which are correlated to some extent e.g. healthy eating and participation in regular physical activity. Like multivariate meta-analysis, methods have been developed for including multiple outcomes in NMA (Jackson, Bujkiewicz, Law, Riley, & White, 2017). For example, Taieb et al. (2015) analysed the effects of two classes of anti-diabetic drugs (i.e. dipeptidyl peptidase-4 inhibitors and sulphonylureas) and placebo pills on three outcomes related to glycaemic control in Type-2 diabetes patients, including change in HbA1c from baseline, the change in fasting plasma glucose (FPG) from baseline and the proportion of patients reaching HbA1c < 7%. The advantage of multivariate network meta-analysis is that it allows for the estimation of intervention effects across all comparators for all outcomes of interest - even those for which there is currently no direct evidence available. In this case, no evidence was available regarding the proportion of patients reaching HbA1c < 7% for the comparison of sulfonylureas and placebo pills. Multivariate NMA not only revealed that these drugs had a significant benefit, but also produced more precise estimates of the intervention effects of the other drugs included in the analysis (Taieb, Belhadi, Gauthier, & Pacou, 2017). Examining multiple outcomes is vital to ensuring that all relevant outcomes, including benefits and harms, contribute to the estimation of the intervention effect and also

For Peer Review Only

avoids problems related to overestimation of the variance of effects sizes, biased effect sizes and type-2 error due to multiple comparisons (Mavridis & Salanti, 2011).

With respect to automated meta-analyses, one particularly ambitious project focuses on developing advanced techniques for synthesising health research is the Human Behaviour Change Project (Michie et al., 2017). This project aims to identify the extent to which health behaviour change interventions work and the contribution of effect modifiers, such as participant characteristics, setting and target behaviour. This project will apply artificial intelligence and machine learning technology to code studies based on an ontology of behaviour change and then extract data in order to perform automated meta-analyses (Larsen et al., 2016). While the prospect of evidence synthesis being facilitated in this way is exciting, the decision-making value of the outputs of this project will be limited if a purely pairwise approach to meta-analysis is taken. For the Human Behaviour Change Project to fulfil its aims, it must integrate network and multivariate analytic approaches into its design. Such an approach, known as live cumulative NMA, has already been developed in clinical medicine, though further development of the methodology and of the reporting of systematic reviews in health research is needed before it is commonly applied (Créquit, Trinquart, & Ravaud, 2016; Vandvik, Brignardello-petersen, & Guyatt, 2016).

Not only are there interesting opportunities for application of NMA in health psychology, there are also exciting opportunities for health psychology to contribute to the development of NMA, particularly in the area of evidence synthesis for complex interventions. It has been proposed that NMA would provide a useful framework for analysing the contribution of specific components (i.e. elements of an intervention which actively influence the intervention effect; Kühne, Ba, Härter, & Kriston, 2015) within complex interventions (Caldwell & Welton, 2016; Madan et al., 2014; Welton et al., 2009). A high degree of heterogeneity is introduced by attempting to synthesise evidence from

For Peer Review Only

complex interventions in pairwise meta-analysis (Kühne et al., 2015). This is because complex interventions, by definition, involve multiple components which may interact and these components can vary between studies (Craig, Dieppe, Macintyre, & Michie, 2008). Applying NMA allows for components (e.g. which are common across interventions in an evidence network to be represented as nodes in the network (Caldwell & Welton, 2016). Welton and colleagues (2009) have demonstrated three analytic models which make different assumptions regarding the relationships between intervention components. The additive main effects model assumes that the effects of each intervention component sum together. In this model, the components are assumed not to interact or cancel each other out in any way. The two-way interaction model allows pairs of components to have a larger or smaller effect when found together in an intervention than that would be expected of an intervention involving one of those components alone. The full-interaction model treats each specific combination of intervention components as a unique intervention with an associated intervention effect (Caldwell & Welton, 2016).

However, there is debate regarding the best way to identify and model the components within complex interventions. Many methods of coding intervention components can be employed. These have been described as falling into two categories: clinically meaningful unit methods and component dismantling methods. Focusing on the clinically meaningful unit means addressing which broad approach to intervention is most effective. Dismantling methods involve the examination of how specific components (or their combinations) affect intervention efficacy (Melendez-Torres, Bonell, & Thomas, 2015). This debate represents an opportunity for health psychology to contribute a considerable amount of accumulated knowledge regarding the coding of intervention components in terms of modes of delivery, settings, behaviour change techniques, theoretical constructs and

For Peer Review Only

mechanisms of action (van Genugten, Dusseldorp, Webb, & Empelen, 2016; Kok et al., 2016; Michie et al., 2013).

Conducting a Network Meta-Analysis

Once the assumptions of NMA are met, there are models available for conducting an NMA on many different types of effect size estimates including those most commonly used in health psychology, mean differences and odds ratios. NMA can be carried out within a frequentist or Bayesian framework. Comparisons of the two approaches appear to show similar outcomes (Hong et al., 2013). However, Bayesian methods for conducting NMA are more flexible, as they can make use of prior information regarding model estimates; account for uncertainty and inconsistency; and yield easily interpretable results (Hong et al., 2013; Neupane, Richer, Bonner, Kibret, & Beyene, 2014).

Bayesian NMA is most commonly conducted using *Bayesian inference Using Gibbs Sampling* (BUGS) software, including WinBUGS and OpenBUGS (Lunn, Thomas, Best, & Speigelhalter, 2000). These programs were developed to allow for the use of Markov Chain Monte Carlo methods for analysing Bayesian statistical models. Dias and colleagues provide WINBUGS/OpenBUGS code for a wide range of commonly encountered evidence/outcome types (Dias et al., 2011). Similar programs include JAGS and Stan (Stephenson, Fleetwood, & Yellowlees, 2015). While the BUGS environment may be difficult to adapt to, Brown et al. (2014) have developed an accessible tool called NetMetaXL, which runs within Microsoft Excel and interfaces with WinBUGS to better facilitate Bayesian NMA. The `gemtc` (van Valkenhoef & Kuiper, 2016), LaplacesDemon (Hall et al., 2016) and pcnetmeta (Lin, Zhang, & Chu, 2016) packages for the R environment can also be used for the same purpose. There are packages available in Stata for conducting NMA within the frequentist framework, including `mvmeta, network` (White, 2009) and `network graphs` (Chaimani, Higgins,

For Peer Review Only

Mavridis, Spyridonos, & Salanti, 2013). The 'netmeta' package for the R environment is also based in a frequentist framework (Rücker, Scharzer, Krahn, & König, 2017). Most of these software packages are available free and many come with accessible guides on how to use them. See table 1 for a comparison of some of the most popular packages available. Next, we present a step-by-step example of the application of NMA to a set of trials of behavioural interventions.

Table 2. Comparison of a sample of popular software packages capable of NMA. Adapted from Neupane, Richer, Bonner, Kibret, & Beyene (2014).

[Insert Table 2 here]

A Step-by-step Example of the Development and Conduct of a Network Meta-analysis

Background: Kanters and colleagues (2017) provide a useful illustration of how NMA has been applied in synthesising the evidence on these behaviour change interventions which are not often compared directly to each other. The main steps involved in conducting this NMA are described below.

Step 1: The research question for this study was generated in the context of a need to update the WHO global consolidated guidelines on HIV. This required the examination of the comparative effectiveness of medication adherence interventions on adherence to ART and HIV viral load.

Step 2: A detailed protocol was developed using the PRISMA extension to NMA (Hutton et al., 2015) to guide the study design, analyses and reporting. This set out a clear focus on the population (people living with HIV), interventions (those targeting enhanced adherence to

For Peer Review Only

ART), comparators (standard care) and outcomes (treatment adherence and viral suppression; PICO) and described the key search terms.

Step 3: The database search was conducted and supplemented by additional standardised strategies to identify grey literature. .

Step 4: Two investigators independently reviewed any identified abstracts and subsequently relevant full text articles to identify the relevant RCTs. The quality of the included studies were assessed using the Cochrane tool for assessing risk of bias (Higgins et al., 2016) and the GRADE criteria for assessing the strength of evidence in NMAs (Caldwell et al., 2016).

Step 5: Two investigators independently extracted the pre-specified data.

Step 6: They categorised intervention and control arms in the identified RCTs using the following categories: standard of care, enhanced standard of care, telephone, SMS, behavioural skills training or medication adherence training, multimedia, cognitive behavioural therapy, supporter, incentives, and device reminder interventions. Due to the considerable heterogeneity across the term standard of care, they defined enhanced standard of care as interventions that provided more support than the usual standard of care. Standard of care was defined as instructions by the health-care provider at treatment initiation regarding how to take ART medication and the importance of adhering to it. Included studies were also classified according to whether they were based in high income and low-income and middle income (LMIC) settings.

Step 7: NMAs were conducted to compare the effect of intervention categories on adherence and viral suppression for all study settings (i.e. the global network) and for studies in the LMIC network only. These NMAs were conducted using logistic regression models which included dichotomised variables indicating medication adherence success and viral load suppression as outcome variables. Both fixed-effects and random-effects models were considered – the model with the lowest deviance information criterion was selected. Potential

For Peer Review Only

effect modifiers were identified (e.g. sample characteristics and time of measurement), and meta-regression was used to evaluate their influence. Sensitivity analyses were conducted to assess the influence of different follow-up periods and the use of either the intention-to-treat or per-protocol results. All analyses were carried out with R (version 3.1.2) and OpenBugs (version 3.23). The authors do not report any analysis of the consistency between direct evidence and indirect evidence for the comparisons in the evidence network. Ideally, models for checking for the presence of consistency are applied – for example, the design-by-treatment interaction model by Higgins et al. (2011). This is an informative approach as it provides information on the appropriateness of the categorisation of the nodes and the reliability of the effect size estimates. Tabular ranking strategies and visual depictions of intervention rank are also sometimes employed to identify the best intervention approaches (Salanti, Ades, & Ioannidis, 2011). In Kanters et al., (2016) forest plots were employed to compare effect sizes for intervention approaches on ART adherence and HIV viral load.

Step 8: The results of these NMAs demonstrated, using the direct and indirect evidence available, an estimate for the effect size between each pair of interventions for both ART adherence and viral suppression. These are presented as a table of odds ratios with each effect size representing the comparisons between the interventions.

Considering these estimates, the authors concluded that supportive strategies and behavioural strategies are more effective than standard adherence support. Medication adherence interventions which involved both in-person and telephone support were more effective than most other interventions.

For a summary of the steps usually taken in conducting a study involving NMA, see table 3.

Table 3. Generic steps in the planning and execution of a NMA.

[Insert table 3 here]

For Peer Review Only

Exemplar Applications of Network Meta-analysis of Relevance to Health Psychology

NMAs that have a particular resonance for health psychology and behavioural medicine are increasingly being reported over the last 5 years. We briefly illustrate three such studies here. One such example examined the comparative efficacy of exercise and drug interventions on mortality outcomes (Naci & Ioannidis, 2013). This analysis incorporated data from 305 RCTs and found that exercise and many drug interventions are often similarly effective with respect to their impact on survival, in the context of secondary prevention of coronary heart disease, rehabilitation after stroke, treatment of heart failure and prevention of diabetes. The study also found that diuretics were more effective than exercise in reducing mortality in those with heart failure. The findings from this analysis highlighted the need to perform RCTs on the comparative effectiveness of exercise and drug interventions. These findings are important for health psychology as they demonstrate that behavioural intervention, in the form of physical activity promotion, may be as effective as medical intervention (i.e. secondary prevention medications) in some contexts.

Mayo-Wilson et al. (2014), examined the comparative efficacy of psychological and pharmacological interventions for social anxiety disorder in adults. They used a "class-effect" model, where each type of intervention is considered to be distinct, but that effects are similar within classes. This provides a balance between avoiding heterogeneity due to "lumping", and avoiding imprecision due to "splitting". The analysis used data from 101 RCTs and found that the efficacy of some psychological interventions for social anxiety disorder (e.g. individual cognitive behavioural therapy), were comparable to some classes of pharmacological interventions (e.g. selective serotonin-reuptake inhibitors and serotonin–norepinephrine reuptake inhibitors). As cognitive behavioural therapy has been shown to have lower risk of side-effects than some pharmacological interventions, this review

For Peer Review Only

recommended that it should be regarded as the best intervention for the initial treatment of social anxiety disorder. Once again, such findings are important for health psychology, as they demonstrate that psychological intervention may be as effective as medical treatment, but with the added benefit of a reduced risk of adverse side-effects. This evidence, an integration of direct and indirect comparisons derived through NMA, supports the prioritisation of psychological intervention for this significant health problem.

A final example of NMA that may potentially change health psychology intervention for Type-2 diabetes treatment was reported by Pillay et al. (2015). Pillay and colleagues' review aimed to identify factors moderating the effectiveness of behavioural programmes for adults with Type-2 diabetes. This synthesis included 132 RCTs and found that several aspects of the content and delivery of these programmes were associated with outcomes. For example, self-management education, offering 10 or fewer hours of contact with delivery personnel, provided little benefit and that these programs seem to benefit persons with suboptimal or poor glycaemic control more than those with good control.

These findings have resonance for health psychology as they provide indirect comparative effectiveness data that can be used to optimise the delivery of health psychology intervention in the context of a specific chronic illness. When considering these and any other applications of NMA, it is vital to scrutinise how the evidence network was determined, whether transitivity and consistency were established. Useful tools for evaluating the quality of studies which have applied NMA can be found in the work of Salanti and colleagues (Salanti, Giovane, Chaimani, Caldwell, & Higgins, 2014), Chaimani and colleagues (Chaimani, Salanti, Leucht, Geddes, & Cipriani, 2017) and Jansen and colleagues (Jansen et al., 2014).

For Peer Review Only

Conclusion

The primacy of direct evidence will, and should, continue to determine the most effective and cost-effective means of health psychology intervention to improve health outcomes. However, the appropriate and judicious use of indirect comparisons can provide insights that can shed light on the potential value of health psychology interventions that may influence the role of the discipline in the delivery of healthcare. Network meta-analysis and its variants provide a useful evidence synthesis methodology that is currently underused in health psychology. This methodology is expected to make a significant contribution to the evolution of both the science and practice of health psychology in the years to come.

Acknowledgements: This work was supported by the Irish Research Council under the New Horizons grant REPRO/2016/31. Thanks to Dr Chris Dwyer for proofing this manuscript and providing advice on style.

For Peer Review Only

References

Ayling, K., Brierley, S., Johnson, B., Heller, S., Eiser, C., Brierley, S., … How, C. E. (2015). How standard is standard care ? Exploring control group outcomes in behaviour change interventions for young people with type 1 diabetes. *Psychology & Health*, *30*(1), 85–103. doi:10.1080/08870446.2014.953528

Bartholomew, L. K., Parcel, G. S., & Kok, G. (1998). Intervention mapping: a process for developing theory- and evidence-based health education programs. *Health Education & Behavior: The Official Publication of the Society for Public Health Education*, *25*(5), 545–563. doi:10.1177/109019819802500502

Beauchamp, M. R., & McEwan, D. (2017). Response Processes and Measurement Validity in Health Psychology BT - Understanding and Investigating Response Processes in Validation Research. In B. D. Zumbo & A. M. Hubley (Eds.) (pp. 13–30). Cham: Springer International Publishing. doi:10.1007/978-3-319-56129-5_2

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* Wiley: Chichester.

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT statement to trials reporting nonpharmacological treatments: extension and elaboration. *Ann Intern Med*, *148*(4), 295–309. doi:10.7326/0003-4819-148-4-200802190-00008

Brown, S., Hutton, B., Clifford, T., Coyle, D., Grima, D., Wells, G., & Cameron, C. (2014). A Microsoft-Excel-based tool for running and critically appraising network meta-analyses---an overview and application of NetMetaXL. *Systematic Reviews*, *3*(1), 110. doi:10.1186/2046-4053-3-110

For Peer Review Only

Bruin, M. De, & Viechtbauer, W. (2014). Standard Care Quality Determines Treatment Outcomes in Control Groups of HAART- Adherence Intervention ... Standard Care Quality Determines Treatment Outcomes in Control Groups of HAART-Adherence Intervention Studies : Implications for the Interpretation and Comparison of Intervention Effects, (November 2009). doi:10.1037/a0015989

Bucher, H. C., Guyatt, G. H., Griffith, L. E., & Walter, S. D. (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology*, *50*(6), 683–691. doi:10.1016/S0895-4356(97)00049-8

Caldwell, D. M., & Welton, N. J. (2016). Approaches for synthesising complex mental health interventions in meta-analysis. *Evidence-based mental health, 19*(1), 16-21. doi:eb-2015-102275

Caldwell, D. M., Ades, A. E., & Higgins, J. P. T. (2005). Simultaneous comparison of multiple treatments: Combining direct and indirect evidence. *British Medical Journal*, *331*(7521), 897–900.

Caldwell, D. M., Ades, A. E., Dias, S., Watkins, S., Li, T., Taske, N., … Welton, N. J. (2016). A threshold analysis assessed the credibility of conclusions from network meta-analysis. *Journal of Clinical Epidemiology*, *80*, 68–76. doi:10.1016/j.jclinepi.2016.07.003

Caldwell, D. M., Dias, S., & Welton, N. J. (2015). Extending treatment networks in health technology assessment: how far should we go? *Value in Health, 18*(5), 673-681. doi: 10.1016/j.jval.2015.03.1792

Card, N. A. (2017). Advances in meta-analysis methodologies contribute to advances in research accumulation: Comments on Cheung & Hong and Johnson et al. *Health Psychology Review*. doi: 10.1080/17437199.2017.1345646

For Peer Review Only

Catalá-López, F., Aurelio, T., Cameron, C., Moher, D., & Hutton, B. (2014). Network meta‐analysis for comparing treatment effects of multiple interventions : an introduction. *Rheumatology International*, *34*, 1489–1496. doi:10.1007/s00296-014-2994-2

Chaimani, A., Higgins, J. P. T., Mavridis, D., Spyridonos, P., & Salanti, G. (2013). Graphical Tools for Network Meta-Analysis in STATA. *PLoS ONE*, *8*(10). doi:10.1371/journal.pone.0076654

Chaimani, A., Salanti, G., Leucht, S., Geddes, J. R., & Cipriani, A. (2017). Common pitfalls and mistakes in the set-up, analysis and interpretation of results in network meta-analysis: what clinicians should look for in a published article. *Evidence-based mental health*, *20*(3), 88-94. doi:10.1136/eb-2017-102753

Cheng, H. Y., Elbers, R. G., Higgins, J. P., Taylor, A., MacArthur, G. J., McGuinness, L., ... & Kessler, D. (2017). Therapeutic interventions for alcohol dependence in non-inpatient settings: a systematic review and network meta-analysis (Protocol). *Systematic reviews, 6*(1), 77. doi:10.1186/s13643-017-0462-2

Cheung, M. W. L., & Hong, R. Y. (2017). Applications of meta-analytic structural equation modelling in health psychology: examples, issues, and recommendations. *Health Psychology Review*, *11*(3), 265-279. doi: 10.1080/17437199.2017.1343678

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *Bmj*, *337*, a1655. doi:10.1136/bmj.a1655

Créquit, P., Trinquart, L., & Ravaud, P. (2016). Live cumulative network meta-analysis: protocol for second-line treatments in advanced non-small-cell lung cancer with wild-type or unknown status for epidermal growth factor receptor. *BMJ open, 6*(8), e011841.

For Peer Review Only

doi:10.1136/bmjopen-2016-011841

Crutzen, R., & Peters, G.-J. Y. (2017). Targeting Next Generations to Change the Common Practice of Underpowered Research. *Frontiers in Psychology, 8,* 1184 doi: 10.3389/fpsyg.2017.01184 .

de Bruin, M., Viechtbauer, W., Hospers, H. J., Schaalma, H. P., & Kok, G. (2009). Standard Care Quality Determines Treatment Outcomes in Control Groups of HAART-Adherence Intervention ... Standard Care Quality Determines Treatment Outcomes in Control Groups of HAART-Adherence Intervention Studies : Implications for the Interpretation and Comparison of Intervention Effects, *Health Psychology, 28*(6)*,* 668-674 doi:10.1037/a0015989

Dias, S., Ades, A.E., Welton, N.J., Jansen, J. & Sutton, A.J. (2018). *Network Meta-Analysis for Decision-Making.* Oxford, UK: John Wiley & Sons Ltd.

Dias S, Welton NJ, Sutton AJ, & Ades AE. (2011). *A generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials.* NICE Decision Support Unit Evidence Synthesis Technical Support Document 2. Retrieved from http://scharr.dept.shef.ac.uk/nicedsu/technical-support-documents/evidence-synthesis-tsd-series/

Dias, S., Sutton, A. J., Ades, A. E., & Welton, N. J. (2013). Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Medical Decision Making,* 33(5), 607-617. doi/abs/10.1177/0272989X12458724

Dias, S., Welton, N. J., Sutton, A. J., & Ades, A. E. (2013). Evidence Synthesis for Decision Making 5: The Baseline Natural History Model. *Medical Decision Making*, *33*(5), 597–606. doi:10.1177/0272989X13487604

For Peer Review Only

Dias, S., Welton, N. J., Sutton, A. J., Caldwell, D. M., Lu, G., & Ades, A. E. (2013). Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Medical Decision Making*, *33*(5), 641-656. doi: 10.1177/0272989X12455847

Falissard, B., Zylberman, M., Cucherat, M., Izard, V., & Meyer, F. (2009). Real medical benefit assessed by indirect comparison. *Thérapie, 64*, 225–232. DOI:10.2515/therapie/2009032.

Forouzanfar, M. H., Liu, P., Roth, G. A., Ng, M., Biryukov, S., Marczak, L., ... & Ali, R. (2017). Global burden of hypertension and systolic blood pressure of at least 110 to 115 mm Hg, 1990-2015. *JAMA, 317*(2), 165-182. doi:10.1001/jama.2016.19043

Freedland, K. E., Mohr, D. C., Davidson, K. W., & Schwartz, J. E. (2011). Usual and Unusual Care: Existing Practice Control Groups In Randomized Controlled Trials of Behavioral Interventions. *Psychosomatic Medicine*, *73*(4), 323–335. doi:10.1097/PSY.0b013e318218e1fb

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., ... & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatrica Scandinavica*, *130*(3), 181-192. doi: 10.1111/acps.12275

Goring, S. M., Gustafson, P., Liu, Y., Saab, S., Cline, S. K., & Platt, R. W. (2016). Disconnected by design: analytic approach in treatment networks having no common comparator. *Research synthesis methods*, *7*(4), 420-432. doi: 10.1002/jrsm.1204

Grant, E. S. & Calderbank-Batista. (2013). Network Meta-Analysis for Complex Social Interventions: Problems and Potential. *Journal of the Society for Social Work and Research, 4*(4), 406 – 420. doi:10.5243/jsswr.2013.25

For Peer Review Only

Hall, B., Hall, M., Statisticat, L., Brown, E., Hermnson, R., Charpentier, E., & Singmann, H. (2016). *LaplacesDemon*. R package version 3.0. Available : https://cran.r-project.org/web/packages/LaplacesDemon/

Higgins, J. P. T., & Whitehead, A. (1996). Borrowing strength from external trials in a meta□analysis. *Statistics in medicine*, *15*(24), 2733-2749. doi: 10.1002/(SICI)1097-0258(19961230)15:24<2733::AID-SIM562>3.0.CO;2-0

Higgins, J. P. T., Sterne JAC, Savović J, Page MJ, Hróbjartsson A, Boutron I, Reeves B, Eldridge S. A revised tool for assessing risk of bias in randomized trials In: Chandler J, McKenzie J, Boutron I, Welch V (editors). Cochrane Methods. *Cochrane Database of Systematic Reviews* 2016, Issue 10 (Suppl 1). DOI:10.1002/14651858.CD201601.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. doi:10.1002/sim.1186

Higgins, J. P. T., Jackson, D., Barrett, J. K., Lu, G., & Ades, A. E., & White, I. R. (2012). Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods, 3,* 98–110. doi: 10.1002/jrsm.1044

Health Information and Quality Authority (2017). Health technology assessment (HTA)of smoking cessation interventions. Available: https://www.hiqa.ie/sites/default/files/2017-04/Smoking%20Cessation%20HTA.pdf

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., … Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ (Clinical Research Ed.)*, *348*(March), g1687. doi:10.1136/bmj.g1687

Hong, H., Carlin, B. P., Shamliyan, T. A., Wyman, J. F., Ramakrishnan, R., Sainfort, F., & Kane, R. L. (2013). Comparing Bayesian and frequentist approaches for multiple

For Peer Review Only

outcome mixed treatment comparisons. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, *33*(5), 702–14. doi:10.1177/0272989X13481110

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., … Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. *Annals of Internal Medicine*, *162*, 777–784. doi:10.7326/M14-2385

Iftikhar, I. H., Bittencourt, L., Youngstedt, S. D., Ayas, N., Cistulli, P., Schwab, R., ... & Magalang, U. J. (2017). Comparative efficacy of CPAP, MADs, exercise-training, and dietary weight loss for sleep apnea: a network meta-analysis. *Sleep Medicine, 30*, 7-14. doi : 10.1016/j.sleep.2016.06.001

Ioannidis, J. P. (2006). Indirect comparisons: the mesh and mess of clinical trials. *Lancet*, *368*(9546), 1470–1472. doi:10.1016/S0140-6736(06)69615-3

Jackson, D., Bujkiewicz, S., Law, M., Riley, R., & White, I. (2017). A matrix-based method of moments for fitting multivariate network meta-analysis models with multiple outcomes and random inconsistency effects. *ArXiv E-Prints*.

Jackson, D., Riley, R., & White, I. R. (2011). Multivariate meta-analysis: Potential and promise. *Statistics in Medicine*, *30*(20), 2481–2498. doi:10.1002/sim.4172

Jansen, J. P., & Naci, H. (2013). Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Medicine*, *11*. doi:10.1186/1741-7015-11-159

Jansen, J. P., Fleurence, R., Devine, B., Itzler, R., Barrett, A., Hawkins, N., ... & Cappelleri, J. C. (2011). Interpreting indirect treatment comparisons and network meta-analysis

For Peer Review Only

for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value in Health, 14*(4), 417-428. doi: 10.1016/j.jval.2011.04.002

Jansen, J. P., Trikalinos, T., Cappelleri, J. C., Daw, J., Andes, S., Eldessouki, R., & Salanti, G. (2014). 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 in Health*, *17*, 157–173. doi:10.1016/j.jval.2014.01.004

Johnson, B. T., Cromley, E. K., & Marrouch, N. (2017). Spatiotemporal Meta-Analysis: Reviewing Health Psychology Phenomena over Space and Time. *Health Psychology Review, 11,* 280-291. doi: 10.1080/17437199.2017.1343679

Kanters, S., Park, J. J., Chan, K., Socias, M. E., Ford, N., Forrest, J. I., ... & Mills, E. J. (2017). Interventions to improve adherence to antiretroviral therapy: a systematic review and network meta-analysis. *The Lancet HIV*, *4*(1), e31-e40. doi: 10.1016/S2352-3018(16)30206-5

Kanters, S., Ford, N., Druyts, E., Thorlund, K., Mills, E. J., & Bansback, N. (2016). Use of network meta-analysis in clinical guidelines. *Bulletin of the World Health Organization*, *94*(10), 782. doi: 10.2471/BLT.16.174326

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., … Bartholomew, L. K. (2016). A taxonomy of behaviour change methods: an Intervention Mapping approach. *Health Psychology Review*, *10*(3), 297–312. doi:10.1080/17437199.2015.1077155

Kovic, B., Zoratti, M. J., Michalopoulos, S., Silvestre, C., Thorlund, K., & Thabane, L. (2017). Deficiencies in addressing effect modification in network meta-analyses : a

For Peer Review Only

meta-epidemiological survey. *Journal of Clinical Epidemiology*. doi:10.1016/j.jclinepi.2017.06.004

Kühne, F., Ehmcke, R. E., Härter, M., & Kriston, L. (2015). Conceptual decomposition of complex health care interventions for evidence synthesis: a literature review. *Journal of Evaluation in Clinical Practice, 21*, 817–823. doi:10.1111/jep.12384

Larsen, K. R., Michie, S., Hekler, E. B., Gibson, B., Ahern, D., Rebecca, H. C., … Hesse, B. (2016). Behavior change interventions : the potential of ontologies for advancing science and practice. *Journal of Behavioral Medicine*. doi:10.1007/s10865-016-9768-0

Lee, A.W. (2014). Review of mixed treatment comparisons in published systematic reviews shows marked increase since 2009. *Journal of Clinical Epidemiology, 67,* 138-143. doi: 10.1016/j.jclinepi.2013.07.014

Leucht, S., Chaimani, A., Cipriani, A. S., Davis, J. M., Furukawa, T. A., & Salanti, G. (2016). Network meta-analyses should be the highest level of evidence in treatment guidelines. *European Archives of Psychiatry and Clinical Neuroscience, 266*(6), 477 - 480. doi: 10.1007/s00406-016-0715-4

Li, T., Puhan, M. A., Vedula, S. S., Singh, S., & Dickersin, K. (2011). Network meta-analysis-highly attractive but more methodological research is needed. *BMC medicine*, *9*(1), 79. doi: 10.1186/1741-7015-9-79

Lin L, Zhang J, Chu H (2014) *pcnetmeta: Methods for patient-centered network meta-analysis.* R package version 1.2. Available: http://CRAN.R-project.org/package5pcnetmeta.

Lu, G. & Ades, A. E. (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine, 23*, 3105–3124. doi:10.1002/sim.1875.

For Peer Review Only

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine, 21*, 2313–2324. DOI:10.1002/sim.1201.

Lunn, D.J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337. doi: 10.1023/A:1008929526011

Madan, J., Chen, Y. F., Aveyard, P., Wang, D., Yahaya, I., Munafo, M., ... & Welton, N. (2014). Synthesis of evidence on heterogeneous interventions with multiple outcomes recorded over multiple follow□up times reported inconsistently: a smoking cessation case□study. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *177*(1), 295-314. doi: 10.1111/rssa.12018

Mancia, G., Fagard, R., Narkiewicz, K., Redon, J., Zanchetti, A., Böhm, M., ... & Galderisi, M. (2013). 2013 ESH/ESC guidelines for the management of arterial hypertension: the Task Force for the Management of Arterial Hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). *Blood pressure, 22*(4), 193-278. doi: 10.3109/08037051.2013.812549

Mavridis, D., & Salanti, G. (2013). A practical introduction to multivariate meta-analysis. *Statistical Methods in Medical Research, 22*(2), 133-158. doi:10.1177/0962280211432219

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *The Lancet Psychiatry*, *1*(5), 368–376. doi:10.1016/S2215-0366(14)70329-3

For Peer Review Only

Melendez-Torres, G. J., Bonell, C., & Thomas, J. (2015). Emergent approaches to the meta-analysis of multiple heterogeneous complex interventions. *BMC Medical Research Methodology*, *15*, 47. http://doi.org/10.1186/s12874-015-0040-z

Meulemeester, J. De, Fedyk, M., Jurkovic, L., Reaume, M., Stotts, G., & Shamy, M. (2018). Many RCTs May Not Be Justified: A Cross-Sectional Analysis of the Ethics and Science of Randomized Clinical Trials. *Journal of Clinical Epidemiology*. doi:10.1016/j.jclinepi.2017.12.027

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology, 28*(6), 690-701. doi: 10.1037/a0016136.

Michie, S., Thomas, J., Johnston, M., Mac Aonghusa, P., Shawe-Taylor, J., Kelly, M. P., ... & O'Mara-Eves, A. (2017). The Human Behaviour-Change Project: harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implementation Science, 12*(1), 121. doi.org/10.1186/s13012-017-0641-5

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., … Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, *46*(1), 81–95. doi:10.1007/s12160-013-9486-6

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science : IS*, *6*, 42. doi:10.1186/1748-5908-6-42

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., ... & Kaplan,

For Peer Review Only

R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and psychosomatics, 78*(5), 275-284. doi:10.1159/000228248

Naci, H., & Ioannidis, J. P. (2013). Comparative effectiveness of exercise and drug interventions on mortality outcomes: meta-epidemiological study. *BMJ,* 347, f5577. doi:10.1136/bmj.f5577

Neupane, B., Richer, D., Bonner, A. J., Kibret, T., & Beyene, J. (2014). Network meta-analysis using R: A review of currently available automated packages. *PLoS ONE*, *9*(12), 1–17. doi:10.1371/journal.pone.0115065

Oberjé, E. J., Dima, A. L., Pijnappel, F. J., Prins, J. M., & de Bruin, M. (2015). Assessing treatment-as-usual provided to control groups in adherence trials: exploring the use of an open-ended questionnaire for identifying behaviour change techniques. *Psychology & Health, 30*(8), 897-910. doi:10.1080/08870446.2014.1001392

Peters, G. J., Abraham, C., & Crutzen, R. (2015). Full disclosure: doing behavioural science necessitates sharing. *European Health Psychologist*, *14*(4), 77-84.

Pillay, J., Armstrong, M. J., Butalia, S., Donovan, L. E., Sigal, R. J., Vandermeer, B., … Dryden, D. M. (2015). Behavioral programs for type 2 diabetes mellitus: A systematic review and network meta-Analysis. *Annals of Internal Medicine*, *163*(11), 848–860. doi:10.7326/M15-1400

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of random effects meta-analyses. *BMJ*, *342*. Retrieved from http://www.bmj.com/content/342/bmj.d549.abstract

For Peer Review Only

Roever, L., & Biondi-Zoccai, G. (2016). Network Meta-analysis to Synthesize Evidence for Decision Making in Cardiovascular Research. *Arquivos Brasileiros de Cardiologia*, *106*(4), 333–337. doi: 10.5935/abc.20160052

Rucker G, Schwarzer G, Krahn U, Konig J (2014) *netmeta: Network meta-analysis with R*. R package version 0.5-0. Availiable: http://CRAN.R-project.org/package5netmeta.

Salanti, G. (2012). Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods*, *3*(2), 80–97. doi:10.1002/jrsm.1037

Salanti, G., Ades, A. E., & Ioannidis, J. P. (2011). Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology, 64*(2), 163-171. doi: 10.1016/j.jclinepi.2010.03.016

Salanti, G., Del Giovane, C., Chaimani, A., Caldwell, D. M., & Higgins, J. P. (2014). Evaluating the quality of evidence from a network meta-analysis. *PloS one, 9*(7), e99682. doi:10.1371/journal.pone.0099682

Salanti, G., Marinho, V., & Higgins, J. P. (2009). A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *Journal of Clinical Epidemiology*, *62*, 857–864. doi:10.1016/j.jclinepi.2008.10.001

Schwingshackl, L., Chaimani, A., Hoffmann, G., Schwedhelm, C., & Boeing, H. (2017). Impact of different dietary approaches on blood pressure in hypertensive and prehypertensive patients: protocol for a systematic review and network meta-analysis. *BMJ Open, 7*(4), e014736. doi: 10.1136/bmjopen-2016-014736

For Peer Review Only

Song, F., Altman, D. G., Glenny, A. M., & Deeks, J. J. (2003). Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. *BMJ*, *326*(7387), 472. doi: 10.1136/bmj.326.7387.472

Stephenson, M., Fleetwood, K., & Yellowlees, A. (2015). Alternatives to WinBUGS for Network Meta-Analysis. *Value in Health, 18*, A720. doi: 10.1016/j.jval.2015.09.2730

Suissa, K., Larivière, J., Eisenberg, M. J., Eberg, M., Gore, G. C., Grad, R., ... & Filion, K. B. (2017). Efficacy and Safety of Smoking Cessation Interventions in Patients With Cardiovascular Disease. *Circulation: Cardiovascular Quality and Outcomes, 10*(1), e002458. doi: 10.1161/CIRCOUTCOMES.115.002458

Sutton, A. J., & Higgins, J. (2008). Recent developments in meta-analysis. *Statistics in Medicine, 27*(5), 625-650. doi:10.1002/sim.2934

Taieb, V., Belhadi, D., Gauthier, A., & Pacou, M. (2017). Multivariate Network Meta-Analysis: An Example In Type 2 Diabetes For The Analysis Of Glycaemic Control. *Value in Health*, *18*(7), A687. doi:10.1016/j.jval.2015.09.2542

van Genugten, L., Dusseldorp, E., Webb, T. L., & van Empelen, P. (2016). Which combinations of techniques and modes of delivery in internet-based interventions effectively change health behavior? A meta-analysis. *Journal of medical Internet research, 18*(6). doi:10.2196/jmir.4218.

van Valkenhoef, G. & Kuiper, J. (2014) GeMTC network meta-analysis. R package version 0.6. Available: http://CRAN.R-project.org/package5gemtc

Vandvik, P. O., Brignardello-Petersen, R., & Guyatt, G. H. (2016). Living cumulative network meta-analysis to reduce waste in research: A paradigmatic shift for systematic reviews?. *BMC medicine, 14*(1), 59. doi:10.1186/s12916-016-0596-4

For Peer Review Only

Welton, N. J., Caldwell, D. M., Adamopoulos, E., & Vedhara, K. (2009). Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *American Journal of Epidemiology, 169*(9), 1158-1165. doi:10.1093/aje/kwp014

Welton, N. J., Caldwell, D. M., Adamopoulos, E., & Vedhara, K. (2009). Practice of Epidemiology Mixed Treatment Comparison Meta-Analysis of Complex Interventions : Psychological Interventions in Coronary Heart Disease, *169*(9), 1158–1165. doi:10.1093/aje/kwp014

White, I. R. (2009). Multivariate random-effects meta-analysis. *Stata Journal*, *9*(1), 40–56. Retrieved from http://www.stata-journal.com/article.html?article=st0156

White, I. R., Barrett, J. K., Jackson, D., & Higgins, J. (2012). Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research synthesis methods*, *3*(2), 111-125. doi: 10.1002/jrsm.1045

Whittaker, R., McRobbie, H., Bullen, C., Rodgers, A., & Gu, Y. (2016). Mobile phone-based interventions for smoking cessation. *The Cochrane Database of Systematic Reviews*, *4*, CD006611. doi:10.1002/14651858.CD006611.pub4

Williamson, P. R., Altman, D. G., Blazeby, J. M., Clarke, M., Devane, D., Gargon, E., & Tugwell, P. (2012). Developing core outcome sets for clinical trials: issues to consider. *Trials*, *13*(1), 132. doi:10.1186/1745-6215-13-132

Wu, P., Wilson, K., Dimoulas, P., & Mills, E. J. (2006). Effectiveness of smoking cessation therapies: a systematic review and meta-analysis. *BMC Public Health*, *6*, 300. doi:10.1186/1471-2458-6-300

For Peer Review Only

Term	Definition
Pairwise meta-analysis	A statistical analysis method for synthesising evidence from a set of individual trials which involved similar populations and which all compare the same (or very similar) two intervention conditions with a focus on the same (or a very similar) outcome
Network meta-analysis *Also known as "mixed treatment comparison meta-analysis" and "multiple treatment meta-analysis"*	A statistical method for synthesising both direct and indirect evidence from a set of individual trials which involved similar populations, and which may include multiple different intervention conditions with a focus on the same (or a very similar) outcome
Indirect treatment comparison *Also known as "adjusted indirect comparison" and "simple indirect comparison"*	A statistical analysis method for synthesising evidence from individual trials of two interventions which have not been directly compared in head-to-head trials, but which have been compared to a common intervention in head-to-head trials
Evidence network	A body of evidence from trials which compared multiple interventions in a homogenous population with a focus on the same (or a very similar) outcome
Network diagram	A graphical representation of an evidence network which usually uses nodes, the size of which represent the number of participants which took part in a specific intervention across multiple trials, and edges – lines connecting the nodes which indicate what interventions have been compared. The thickness of the edges represents the number of trials which have compared the two interventions represented by the connected nodes.
Closed loop	A closed loop can be seen in a network diagram whenever there is both direct and indirect evidence connecting a set of three or more interventions
Disconnected network	Disconnection occurs when there is neither direct nor indirect comparisons between certain interventions in the network
Effect modifiers	Clinical or methodological characteristics of studies which affect the relative effect between interventions
Transitivity	Transitivity implies that interventions, methods and populations in an evidence network are comparable in terms of the distribution of effect modifiers

For Peer Review Only

Consistency	Consistency is the statistical demonstration of agreement between the direct evidence and the indirect evidence for all pairwise comparisons in a network for which both direct and indirect evidence are present

Please see Diaz, Ades, Welton, Jansen & Sutton (2018) for further information and a comprehensive definitive resource on NMA.

For Peer Review Only

Statistical Framework	Features	gemtc - R	pcnetmeta - R	netmeta - R	laplacesDemon - R	WinBUGS/ OpenBUGS/ JAGS/ Stan/ NetMetaXL	mvmeta/ network graphs - Stata
Bayesian				■			■
Frequentist		■	■		■	■	
Tasks							
Forms of Input Data	Arm-level data			■			
	Contrast-level data		■				■
	Accepts multi-arm trials						
Types of Outcome Data that Can be Analysed	Binary						
	Count		■				
	Continuous		■				
	Survival		■				
Extracts descriptive measures	Total number of studies		■				
	Total number of multi-arm studies		■				
	Total number of participants		■	■			
	Total number of treatments		■				
Network plot and options	Network plot				■		
	Add node labels				■		
	Node size reflects network characteristics	■		■	■		
	Edge thickness reflects network characteristic	■			■		

Note: White blocks indicates presence of the feature; black blocks indicates that the feature is not present in the software package.

Step	Aim	Considerations
1	Generate Research Question	• The research question should be constructed with consideration of both the clinical and methodological characteristics of the studies of interest
2	Plan Systematic Review	• This should be guided by PRISMA extension for NMA • A clear definition of the PICO must be presented and the associated inclusion and exclusion criteria should allow the inclusion of as many relevant interventions and comparators as possible • Potential effect modifiers should be identified • The plan for the systematic review, should be registered in PROSPERO and detailed in a study protocol
3	Conduct Search	• In situations where a large body of literature exists and high-quality systematic reviews have been carried out, the search may focus on identifying these, as identifying individual studies through a primary search may not be feasible.
4	Select Studies	• Studies should be selected according to the inclusion and exclusion criteria ideally by two independent reviewers • Studies which involve interventions which are not central to the research question may be included if they are compared to interventions which are central to the research question and this provides more useful evidence to the network
5	Extract Data	• This stage will generally focus on extracting the relevant data regarding outcomes and potential effect modifiers • Risk of bias and evidence quality should be assessed using the tools provided by Cochrane and GRADE as these characteristics also affect transitivity
6	Build Network	• Decisions regarding splitting and lumping are made at this stage and planned approaches may have to be modified according to the nature of the collected data (e.g. if there is a lack of data, some lumping may have to be done)

For Peer Review Only

		• A network diagram should be constructed and its geometry should be evaluated e.g. Figure 2
7	Analyse Data	• For all comparisons for which there is both direct and indirect evidence, consistency checks should be carried out to ensure that the direct and indirect evidence agrees. • Generally, pairwise analyses are conducted first and then NMA models are conducted • The data should be analysed as set out in the study protocol
8	Interpret and Report Results	• The PRISMA extension for NMA provides guidance on reporting the results in a clear and comprehensive manner. • Data from individual studies should be summarized in tables • The estimates of comparative effectiveness are usually presented in tables and sometimes in a rankogram

For Peer Review Only

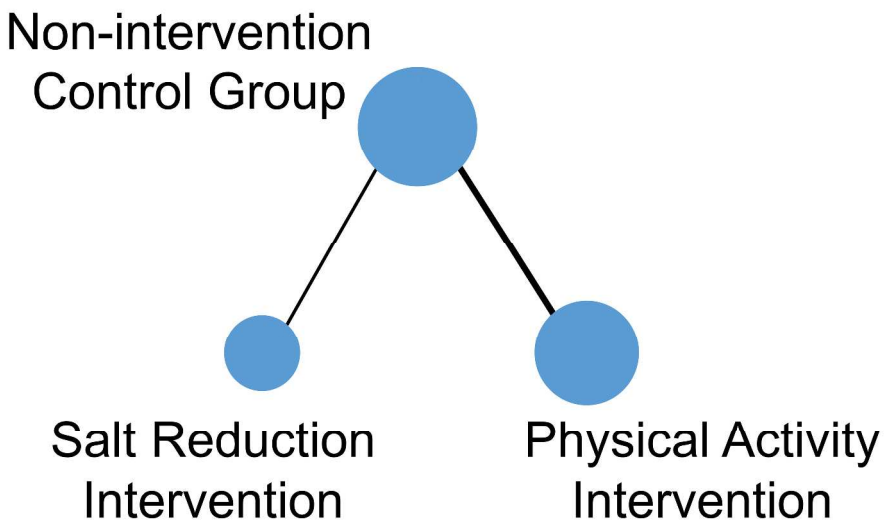

Figure 1. An example of a network diagram

For Review Only

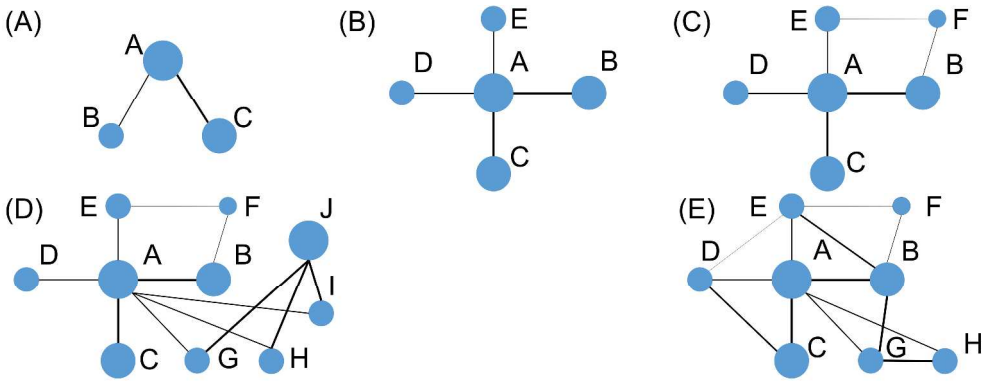

Figure 2. Some possible configurations of networks of evidence

Peer Review Only