Application of Network Meta-Analysis in The Field of Physical Activity and Health Promotion: A Case Study

A THESIS
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Xiwen Su

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

Advisor: Zan Gao, Ph.D.

August 2020

Acknowledgements

To begin, I would like to thank my main advisor, Dr. Zan Gao, without whom I wouldn't have the opportunity to be involved in the interesting field of physical activity and health research. I benefited a lot from his extraordinary mentorship over the past two years and he has provided me so many opportunities to grow my skills and expand my knowledge within this field.

Second, I would like to thank my dissertation committee members, Drs. Haitao Chu and Beth Lewis for their great help and kindly support during the completion of this thesis.

Third, I would also like to thank fellow research assistants Wenxi Liu and DJ McDonough within the Physical Activity Epidemiology Laboratory, and current PhD student Menglu Liang from Penn State University, all of whom had generously shared their precious knowledge as well as experience with me and helped me to better fit in the graduate education at the University of Minnesota. Relatedly, I thank the faculty/staff within the School of Kinesiology as well as my fellow peers for all the support they have provided me these last two years.

Finally, I want to thank my family, friends and roommates for all the support they have provided me as I have worked toward obtaining my master's degree. Their care, encouragement, and belief in my abilities have meant more to me than I can fully express. Knowing that we always share each other's laugh and tears had made me overcome every obstacle on the road and will keep supporting me going further in the future.

Dedication

This dissertation is dedicated to my mother Jing Yang, my father Shiye Xiu, and all of my family and friends who always believed in me and support me unconditionally.

Abstract

Continued advancement in the field of kinesiology and health promotion relies heavily on the synthesis of rigorous quantitative scientific evidence. As such, meta-analyses of randomized controlled trials have led to a better understanding of what intervention strategies are superior (i.e., produce the greatest effects) in physical activity-based health behavior change interventions. Indeed, standard meta-analytic approaches have allowed researchers in the field to synthesize relevant experimental evidence using pairwise procedures which produce reliable estimates of the homogeneity, magnitude, and potential biases in the observed effects. However, pairwise meta-analytic procedures are only capable to discerning differences in effects between a select intervention strategy and a select comparison condition or control condition. In order to maximize the impact of physical activity interventions on health-related outcomes, it is necessary to establish evidence concerning the comparative efficacy of *all* relevant physical activity intervention strategies. The development of network meta-analysis (NMA)—most commonly used in medical-based clinical trials—has allowed for the quantification of indirect comparisons, even in the absence of direct, head-to-head trials. Thus, it stands to reason that NMA can be applied in the physical activity and health promotion research to identify the best intervention strategies. Given this analysis technique is novel and largely unexplored in the field of kinesiology and health promotion, care must be taken in its application to ensure reliable estimates and discernment of the effect sizes between interventions.

Therefore, the purpose of this study is to first comment on the potential application and importance of NMA in the field of kinesiology and health promotion, then describe how to properly and effectively apply this technique using a specific case study evaluating the effects of different lifestyle interventions on children's body composition, and lastly suggest important considerations for its appropriate application in this field. In this paper, overviews of the foundations of NMA and commonly used approaches for conducting NMA are provided, followed by assumptions of NMA, opportunities and challenges in NMA, and a case study example of the development and conduct of an NMA, as well as the interpretation of the analysis results. The case study collect original data from

published randomized controlled studies investigating on some type of intervention on variables including body mass index (BMI), BMI z-score (BMIz), and body fat percentage, divided the used intervention into ten categories in total, from simple single intervention to multiple components mixed intervention (more than three), and used pre-processed data to carry out network meta-analysis. Results of analysis using mean difference (SD) between baseline and immediate post-intervention data showed that PA intervention ranked top two of the most effective approaches among other types of lifestyle interventions in all three variables, suggesting that promoting PA participation is crucial in children's health status and childhood obesity control. While based on the analysis using combined original pre-and-post data (SE), multiple component interventions were predicted to be the best ranked intervention approach among all ten types of intervention, indicating that taking care of more aspects in children's lifestyle may also result in an important impact for children to keep healthy and fit.

Table of Contents

List of Tables

List of Figures

Introduction on network meta-analysis

Continual advancement in the field of physical activity and health promotion depends on the accurate and timely synthesis of all available evidence from interventions based on physical activity and lifestyle (Palmer et al., 2019; Martín-García et al., 2019). In particular, meta-analyses of randomized controlled trials (RCTs) have helped us to better understand the health impact of physical activity promotion interventions and to discern the superiority, or lack thereof, of one intervention strategy over another or of no treatment at all. However, standard meta-analytic techniques only allow for pairwise comparisons (i.e., a direct comparison of one intervention strategy against another or compared to a control condition). Given that levels of physical inactivity and the prevalence of chronic diseases related to physical inactivity remain at epidemic levels (globally, 23% of men and 32% of women aged 18+ years were insufficiently physically active, according to the World Health Organization in 2016), national and global health organizations are seeking quantitative evidence synthesized from simultaneous comparisons of *multiple* intervention strategies (Molloy et al., 2018; Hutton et al., 2015).

As is recognized, the clinical decision-making process should be based on valid empirical evidence. While RCTs comparing the effects of two or more interventions can contribute as direct evidence, systematic reviews show advantages in their abilities to synthesize and analyze all available evidence related to the same clinical question (Li et al., 2011). Meta-analysis has been employed in clinical practice since the 1980s (Tonin et al., 2017), and has been one of the most frequently used statistical methods in systematic reviews for data synthesis. The conventional way of carrying out a meta-analysis is through pairwise comparisons between an intervention and a control (Sofia Dias & Caldwell, 2018). Pairwise meta-analysis is capable of gathering evidence from separate and relatively small studies and, through the combination of their results, may increase statistical power and detect a statistically significant difference between one intervention and another even though not all individual studies observed statistical significance in their results (Tonin et al., 2017). However, pairwise meta-analysis is only able to compare head-to-head trials using the same type of intervention, while in practice there are usually more than 2 approaches available. For clinicians, patients, and policy makers to make well-informed decisions, it is necessary to compare all intervention approaches simultaneously. Nonetheless, for various reasons, in some specific areas there may only exist limited direct evidence, making it difficult to carry out head-to-head comparisons for all types of interventions using traditional pairwise meta-analysis. Under these circumstances, the application of network meta-analysis (NMA), also termed multiple treatment comparison (MTC) or multiple treatment meta-analysis, can be very useful for synthesizing all existing evidence simultaneously (Neupane et al., 2014).

Foundations of NMA

What is NMA?

NMA is a statistical technique that combines both direct (i.e., within-trial) and indirect (i.e., between-trial) comparisons of multiple intervention strategies that may not be directly compared within the same trial (Molloy et al., 2018; Higgins & Whitehead, 1996; Lu & Ades, 2004). The most basic requisite for conducting indirect quantitative comparisons using NMA is that there should be at least one intervention strategy in common for each chain of comparison. Each type of intervention could serve as a connection to different chains of comparisons as a node. This allows for the construction of a network of trials comparing multiple interventions that can be analyzed using NMA. Therefore, compared with only direct evidence derived from standard pairwise meta-analyses, NMA maximizes the availability of evidence by allowing for the comparison of any pair of interventions linked through the constructed evidence-network, thus increasing the precision of the effect size for a given intervention strategy (Molloy et al., 2018; Caldwell et al., 2005; Ioannidis, 2006).

Advantages of NMA

NMA can be seen as an extension of conventional pairwise meta-analysis since they both share similar assumptions and have essentially the same purposes and functions (Tonin et al., 2017). In the field of physical activity and health promotion, intervention approaches mainly focus on informational, behavioral, social, environmental, and policy aspects (Lox, Martin Ginis, & Petruzzello, 2014). More detailed modifiable determinants can be listed under each aspect of approaches that different RCTs carried out by researchers from around the world might address separately. NMA is capable of combining all existing evidence under similar conditions together for analysis, as long as each piece of evidence connects to the network. Not only are direct comparisons from the RCTs taken into account, but every common comparator can contribute to making indirect comparisons as well (Li et al., 2011; Tonin et al., 2017; Dias & Caldwell, 2018). Using appropriate statistical methods, the direct and indirect evidence can be combined as a weighted average in NMAs. For example, if current clinical trials only compared A *vs.* B and B *vs.* C directly (i.e., there are no head-to-head trials comparing A *vs.* C), as shown in Fig. 1, the NMA is able to estimate the relative effect of A *vs.* C using indirect evidence through A *vs.* B and B *vs.* C under the preceding assumptions. On the other hand, if clinical trials comparing A *vs.* C are also available, the relationships of the 3 types of trials could form a closed loop, and the NMA would be able to utilize both direct and indirect sources of information and combine them with an appropriate weight. A visual comparison between pairwise meta-analysis and NMA in the use of direct and indirect evidence is provided below (Fig. 2), featuring the biggest advantage of the NMA. Longer chains of indirect comparisons may also appear under certain circumstances; for example, instead of getting indirect evidence from A *vs.* C through A *vs.* B and B *vs.* C, more common comparators may be needed on the path (i.e., A *vs.* B, B *vs.* D, and D *vs.* C).

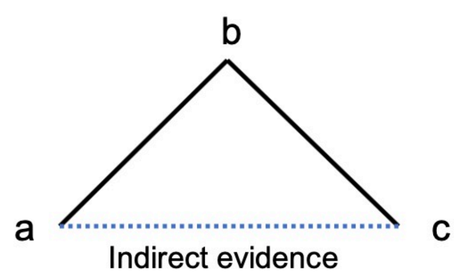


Figure 1. Example of a network of three treatments compared in two trials (solid black lines), where an indirect comparison can be made using network meta-analysis (dashed grey line).

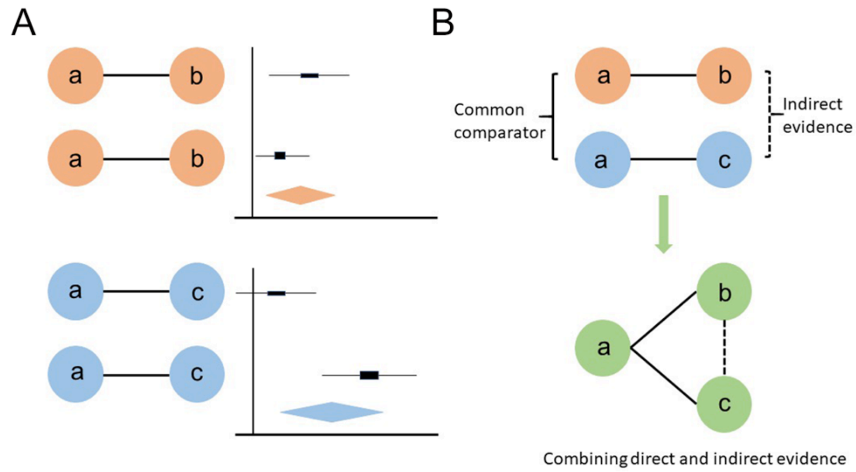


Figure 2. Comparisons between pairwise meta-analysis versus network meta-analysis in the use of direct and indirect evidence. (A) represents an example of pairwise meta-analysis. (B) shows the biggest advantage for network meta-analysis when a common comparator "a" exists.

Assumptions related to NMA and risk of bias

NMA relies on several assumptions that need to be checked prior to conducting the analysis. First, NMA shares the same assumption as pairwise meta-analysis, which is homogeneity (Molloy et al., 2018; Tonin et al., 2017). This assumption presumes no relevant heterogeneity among trial results, which means that the effect of potential modifiers should be very limited for the included RCTs (apart from sample variability), ensuring that all study-related conditions are homogenous (Molloy et al., 2018; Dias & Caldwell, 2018). However, in practice, potential modifiers may still exist and sometimes are not measured or even measurable. Examples of common potential effect modifiers include baseline characteristics of recruited participants, personnel choice of measurements, and intervention setting and dosages, to list a few. According to Dias and Caldwell (2018), an empirical way of checking this assumption is to observe in a general view and see whether the treatments and participants' characteristics are comparable among all the studies and whether it is suitable to combine them for NMA. If so, then in principle the homogeneity assumption would be satisfied. Specifically, in the field of physical activity and health promotion, one of the most likely ways to bring heterogeneity is due to the complexity of the interventions included across the studies. In practice, interventions in kinesiology studies are very likely to have degrees of flexibility or tailoring of the protocol (intensity, duration, etc.), and the outcomes may appear to be natural variabilities (Craig et al., 2013). These could all contribute as characteristics for a complex intervention and might cause the presence of statistical heterogeneity. In their 2016 paper, Caldwell and Welton (2016) addressed this issue and concluded that component-based NMA could be a good option for synthesizing complex interventions. Interested readers are referred to their work for detailed information.

Next, the consistency and transitivity assumptions are specific concerns for NMAs since they both involve the proper use of indirect evidence. The assumption of consistency states that there should not be discrepancies between direct and indirect comparisons. In other words, the results obtained from trials providing direct and indirect evidence should essentially agree with each other. Otherwise, there will be network inconsistency. The consistency assumption can only be assessed when both kinds of evidence are available (i.e., a closed loop within the network). In their previous work, Dias et al. (2013), Higgins et al. (2012), and White et al. (2012) have all provided detailed descriptions of possible strategies for consistency checking. As mentioned previously, direct comparisons may not be always available. However, the assumption of transitivity should still be assessed whether there is direct evidence or not. Transitivity refers to the assumption that for unobserved head-to-head comparisons, using indirect comparison(s) could provide valid and reliable estimates that are close to a direct comparison if one were available (Salanti, 2012). For example, in a closed loop network among A, B, and C, one should be able to conclude that A is better than C, knowing that A is better than B, and B is better than C. If at the same time there is direct evidence showing that A is better than C, then the transitivity and consistency assumptions are both met. Dias et al. (2018) introduced strategies of assessing transitivity in details in the book *Network meta-analysis for decision-making.* Broadly speaking, transitivity could be achieved by "qualitatively examining relevant clinical and methodological aspects of the relevant intervention comparators" (Molloy et al., 2018) to ensure that the potential effect modifiers are distributed evenly across all of the comparators.

Apart from the preceding assumptions that NMAs rely on, it is also important to account for risk of bias of individual studies included in the network. Risk of bias for NMAs shares similarities with conventional pairwise meta-analysis, although risk-of-bias assessment in NMA is far more challenging. First, publication bias or small study effects is one of the most common types of biases faced by meta-analyses since the main data source for second-hand data analyses are usually extracted from published articles. Various statistical methods have been proposed to detect or quantify the magnitude of publication bias (Lin & Chu, 2018; Lin et al., 2020). Although publication bias is hard to avoid or control, particularly from the perspective of multivariate meta-analyses (Hong et al., 2020) including NMAs, it is still necessary to at least be aware of this potential issue. In addition, risk of bias may also commonly occur when individual trials included in the analysis have potential design or execution problems, thus raising concerns regarding the validity and reliability of their results (Li et al., 2011). More importantly, if there is bias from one single trial, it is possible that the findings from this trial may affect several pooled effect estimates in NMAs, whereas only one pooled effect estimate will be affected in conventional pairwise meta-analysis.

To evaluate the certainty of the evidence from NMAs, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) was described by Puhan et al. (2014) in 2014. The guidance mostly focuses on assessing the confidence and quality of the evidence in an NMA and has been broadly used in a number of studies carrying out NMAs (Rochwerg et al., 2016; Sekercioglu et al., 2017). Later in 2018, Brignardello-Petersen et al. (2018) described recent conceptual advances of the GRADE approach and summarized them into 4 major points, mainly regarding consideration of imprecision, necessity of rating of indirect evidence, and global incoherence. In fact, given the rapidly increasing popularity of NMA in all fields, more challenges are expected to be encountered, and further development of the GRADE criteria is also anticipated through GRADE working-group meetings.

Theoretical and Statistical Approaches of NMA

Statistically, one can use either a Bayesian or a frequentist framework to conduct an NMA. When sample size is sufficient, these two frameworks should produce similar results despite the fact that the basic concepts of their statistical approaches differ. Essentially, the main difference between these two methods is whether prior information is considered when building the model. Frequentist methods do not consider information previously known (prior probability) and estimate the population parameters by infinitely repeating the present data and maximizing its likelihood function under a statistical model. The population parameters are considered as fixed unknown values and are not related to external information under this framework (Shim & Lee, 2019). On the other hand, Bayesian methods believe that the interest population parameters should have a posterior distribution that may be affected by prior information since the posterior distribution function can be obtained by multiplying the prior distribution of parameters with the likelihood function (Shim & Lee, 2019). Essentially, the observables and parameters in the model are both viewed as random quantities from the Bayesian perspective (Thomas & Spiegelhalter, 1925; Kathryn et al., 2016).

Since the Bayesian method maintains model uncertainty, its posterior distribution does not follow commonly used distributions (i.e., binomial or normal), rendering it difficult to calculate the area under the distribution curve. The Markov Chain Monte Carlo (MCMC) simulation can be used in this case for area calculation, which is essentially Monte Carlo integration using Markov Chains. MCMC emerged as an extremely popular tool for the analysis of complex statistical models for a short period during the 1990s, especially in the field of Bayesian analysis (Kathryn et al., 2016). The concept of Markov Chains involves calculating the probability of the "next state" of a random variable using an algorithm, where the next state only depends on its current state and transition probability (which is prior information). It is believed that in a Markov Chain, the value of the next state would finally reach a stable distribution with enough repetition of the calculation (Shim & Lee, 2019). Monte Carlo simulation aims to predict a targeted value using random sampling methods based on randomness. Basically, it “draws samples from the required distribution and then forms sample averages to approximate expectations” (Gilks & Richardson, 1996). Together, they form the MCMC simulation to determine posterior distribution for relatively complicated statistical models such as the Bayesian model. Fig. 3 shows a simplified schematic of the concept for the MCMC simulation used in the Bayesian approach.

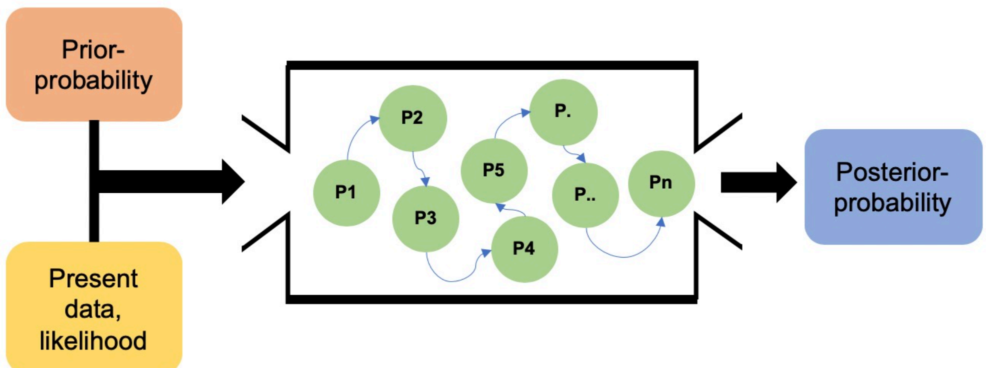


Figure 3. The overall concept of the Bayesian approach using a Markov chain Monte Carlo (MCMC) simulation.

Commonly used approaches for conducting NMA

Besides the differences in theoretical models, there are two types of approaches that can be used for NMA. The contrast-based approach is considered by many researchers to be the standard approach for meta-analysis, while arm-based models (discussed later) have more recently been purported to be an intriguing alternative to the understanding of meta-analysis (Dias & Ades, 2016). The main difference between the 2 approaches is the type of information extracted for analysis. For trials, contrast-based models pool information of relative treatment effects, while arm-based models focus on the absolute values of each outcome of interest (Lin et al., 2017). As an example, absolute measures include treatment-specific event rates, log risks, log odds, and mean outcomes for each arm included. On the other hand, the most commonly reported summary statistics for contrast measures are log-odds ratio (OR) (Zhang et al., 2014) and other statistics, such as log relative risks (RRs) and risk differences (RDs) for binary outcomes, or mean differences for continuous outcomes between treatments. There is continuing discussion regarding the strengths and weaknesses of the 2 types of models. Zhang et al. (2014) carried out a series of hypothetical NMA trials as well as reanalysis of published NMAs and compared the results of their self-processed approach using the alternative arm-based methods to the estimations of contrast-based models. They concluded that the arm-based approach outperformed the original contrast-based NMA methods in terms of bias; and for other outcomes the 2 approaches led to different treatment recommendations but shared some similarities as well (Zhang et al., 2014). In general cases, assuming the same OR or RR across different baseline risks could lead to different absolute RDs. Under these circumstances, some researchers believe that arm-based NMA might be preferred because it provides a more straightforward and accurate methodology to assess different intervention effects (Lin et al., 2017). Moreover, the arm-based methodology can use information contained in single-armed studies and therefore take into account more available treatment groups, while contrast-based studies are not capable of including such studies (Lin et al., 2017). However, the preference for the arm-based model supported by Zhang et al. (2014) and Hong et al. (2016a) was criticized by Dias and Ades (2016), who stated that "contrast-based models are to be preferred on both theoretical and practical grounds." Furthermore, Dias and Ades (2016) presented detailed comparisons and arguments supporting their conclusion that advocating for arm-based models is not helpful for the separation of absolute and relative effects, an essential problem that epidemiologists and biostatisticians have been working on for many years. Dias and Ades (2016) argued that the use of arm-based models risked biased estimates with over-inflated posterior variance and thus believed that previous studies (Zhang et al., 2014; Hong et al., 2016a) favoring the alternative arm-based method were actually mistaken. The discussion continued when Hong et al. (2016b) published a rejoinder that provided a section-by-section response to Dias and Ades (2016). Hong et al. (2016b) argued that because the assumption requirement for arm-based models is considerably higher than for contrast-based models, the payoffs were worthwhile since arm-based models allowed for significantly higher modeling flexibility, thereby making them more advantageous for model fitting and interpretation. Thus, Hong et al. (2016b) asserted their belief that arm-based models were more complete, and that a fully Bayesian approach was superior for handling missing data. These discussions allowed the researchers to fully explore every aspect of the issue and offered a chance for other analysts to better decide for themselves which model to consider for their own work. A more recent comparison by White et al. (2019) was carried out specifically targeting the arm-based model supported by Hong et al. (2016a) and the contrast-based model supported by Lu and Ades (2006). Four key differences between the 2 models were identified, but the discussion mainly focused on whether the study intercepts were random or fixed effects, which, as White et al. (2019) suggested, is the most important difference between the models. White et al. (2019) concluded that both arm-based and contrast-based models are suitable for NMA but pointed out that using random study intercepts requires a strong rationale, while models with fixed study intercepts are useful because they can be implemented with either a contrast-based or arm-based model. Wang et al. (2020) observed that a separation strategy with appropriate priors for the correlation matrix and variances performs better than strategies employing the inverse—Wishart priors used in the original arm-based NMA and can therefore reduce potential biases. Recently, Ma et al. (2018) and Lian et al. (2019) have extended arm-based NMA to simultaneously compare multiple diagnostic tests in which absolute measures, such as sensitivities and specificities, are of primary interest.

There are several statistical programs and software available that can carry out the required calculations and simulation steps for NMA. For instance, Statistical Analysis System (SAS) and statistics and data (STATA) are capable of employing NMA based on frequentist methods. Other open-access resources (e.g., OpenBUGS, WinBUGS, JAGS) can help in conducting NMA under the Bayesian framework as an MCMC sampler (Neupane et al., 2014). R, a popular open source statistical software, is frequently used among statisticians nowadays. According to the review by Neupane et al. (2014), until 2014 there were only three available R packages developed specifically for performing NMA: *gemtc*, *pcnetmeta*, and *netmeta*. The first 2 packages perform the analysis under the Bayesian framework and the last one performs it using the frequentist framework. According to the comparisons and assessments by Neupane et al. (2014), the three R packages provide different and often complementary features for performing all aspects of NMA. One or more of these packages could be used to plot the network, generate a model, detect heterogeneity and inconsistency in the network, incorporate them into the estimation, and finally generate the estimated effects sizes and rank probabilities. *Gemtc* and *netmeta* are comprehensive packages that employ Bayesian and frequentist techniques, respectively, for contrast-based NMA. In comparison, *pcnetmeta* provides Bayesian analysis for arm-based NMAs, which are generally more robust for the choice of treatments to include in the NMA (Lin et al., 2016). Table 1 summarizes the features and capabilities of the three packages.

Recently, a couple of new tools have been created to conduct NMA via the Web. Examples are MetaInsight (https://crsu.shinyapps.io/metainsightc) (Owen et al., 2019) and CINeMA (Confidence in Network Meta-Analysis, https://cinema.ispm.unibe.ch/) (Papakonstantinou et al., 2020). Both are Web-based, freely available open-source tools with no requirement for the installment of statistical software. R is used as the “backbone” for both of the tools; however, they only call the routines of R packages on the webserver rather than using the software itself. These newly developed platforms provide a more convenient and reliable source that researchers and nonspecialists can use to perform NMA and get immediate visual feedback during their research.

The software mentioned above can generate network plots, and possible configurations that imitate different situations that might occur in real studies are shown in Fig. 4. Every dot represents an arm of treatment (e.g., control, treatment 1, treatment 2, etc.); the solid lines represent edges or direct comparisons between the two treatments or comparators, which are linked by the lines; and the width of the edges has a positive association with the number of direct comparisons that occurred (i.e., the number of articles with this result reported). Generally, the more solid lines that exist, the more precise the estimation of the indirect comparisons would be because any inference should be based on the direct evidence available.

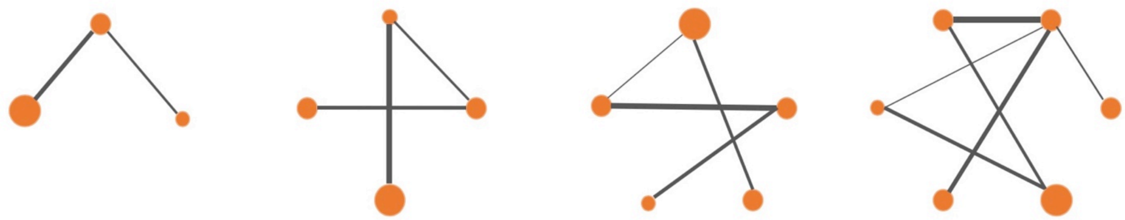


Figure 4. Possible configurations of network plots.

Table 1. Comparison of a sample of popular software packages capable of NMA.

Tasks	Features	*netmeta*	*gemtc*	*pcnetmeta*
Estimation framework	Bayesian		√	√
	Frequentist	√		
Forms of input data	Arm-level data		√	√
	Contrast-level data	√	√	
	Accepts multi-arm (≥ 3) trials	√	√	√
Types of outcome data that can be analyzed	Binary	√	√	√
	Count	√	√	
	Continuous	√	√	√
	Survival	√	√	
Extracts descriptive measures	Total # of studies	√	√	
	Total # of multi-arm studies	√	√	
	Total # of participants		√	
	Total # of treatments	√	√	
Network plot and options	Network plot	√	√	√
	Add node labels	√	√	√
	Node size reflects network characteristic			√
	Edge thickness reflects network characteristic	√		√
Assessing heterogeneity	Visual inspection - forest plot	√	√	
	Pairwise statistics	√	√	
	Global statistics	√	√	
Assessing inconsistency	Visual inspection—forest plot of direct vs. indirect		√	
	Visual inspection—heat map	√		
	Consistency statistics	√	√	
	Back-calculation		√	
	Node-split/decomposition	√	√	
MCMC sampler (when under Bayesian modeling)	WinBUGS	N/A	√	
	OpenBUGS	N/A	√	
	JAGS	N/A	√	√

Opportunities and challenges related to NMA

NMA has most commonly been used in clinical fields where researchers test the effectiveness of different drug interventions. The clinical conditions involving drug interventions most often evaluated by NMA include cardiovascular diseases, oncological disorders, mental health disorders, and infectious diseases (Tonin et al., 2017). However, other NMA applications have developed rapidly in the past decade. According to Tonin et al. (2017), very few systematic reviews containing NMAs were published prior to 2008, yet now there are more than 400 published. NMA is gaining popularity in comparing clinical treatments because there are typically a variety of treatments or drugs targeting similar disease categories, rendering it difficult for clinicians and patients to compare them thoroughly in a pairwise fashion before making informed treatment decisions. Properly conducted NMA studies have the potential to overcome such issues. In fact, using NMA to make indirect comparisons among studies has become a critical component of evidence synthesis and decision making in healthcare (Molloy et al., 2018). In the field of physical activity and health promotion, NMA is gaining attention as a valuable analysis tool and research method. The main trend involves the use of exercise as one of the treatment intervention approaches and comparing this treatment arm with other treatment strategies, or comparing the health benefits of different types of exercise (e.g., aerobic exercise, high-intensity training, resistance training, etc.) in populations with chronic disease or other clinical populations. Specifically, NMA allows for comparisons of the efficacy of behavioral (i.e., physical activity) and biomedical (i.e., pharmacological treatments) intervention strategies on a common health outcome (e.g., weight, body mass index, blood pressure) that would otherwise not have been compared previously in head-to-head trials when they share common comparators, such as a control or placebo group.

In the field of physical activity and health promotion, we identified 11 articles that utilized a physical activity or exercise program as one of the treatment arms within the RCTs. Among the 11 studies, 5 were carried out in Europe (Austria and UK) (Schwingshackl et al., 2013; Schwingshackl et al., 2014); 4 were carried out in Asia (China and Japan) (Zou et al., 2018; Pan et al., 2018) and the other 2 studies were carried out in the US (George et al., 2018) and Brazil (Andreato et al., 2019). As for software choices for the analysis, it appears that STATA and WinBUGS were most popular. Five of the studies used STATA to either generate network plots as a first step or to perform the full NMA analysis (Xia et al., 2018; Yamaoka et al., 2019). The open-source software WinBUGS was used in 5 of the studies to carry out the analysis (Schwingshackl et al., 2013; Naci & John, 2013; Naci et al., 2019), and 1 study used the R package "netmeta" (Pan et al., 2018). Most of the studies included in the analysis used a Bayesian approach (Schwingshackl et al., 2013; Naci & John, 2013; Uthman et al., 2013), while 2 employed the frequentist approach (Pan et al., 2018; George et al., 2018). Six of the studies examined the effects of exercise or other physical-activity-based interventions (e.g., lifestyle interventions) on patients with diseases such as coronary heart disease, type 2 diabetes mellitus, and nonalcoholic fatty liver disease, as well as the effects on mortality outcomes (Schwingshackl et al., 2014; Pan et al., 2018; Xia et al., 2018; Naci & John, 2013). Three studies focused on the effect of exercise training on the participants' change in body weight, adiposity level, or other anthropometric characteristics; all the interventions were moderately effective (Schwingshackl et al., 2013; George et al., 2018; Andreato et al., 2019). Furthermore, 2 studies focused on less severe chronic diseases such as lower limb osteoarthritis and hypertension (Naci et al., 2019; Uthman et al., 2013). Taken together, it is apparent that exercise and physical activity intervention programs are moderately to highly effective for attenuating chronic diseases and obesity-related health problems.

However, kinesiology is a far broader field than exercise science, where most of the focus of NMA currently lies. The promotion of physical activity is needed to improve the health of most populations, given that the prevalence of physical inactivity has become alarmingly higher in the past decade. Numerous studies have reported the use of different approaches for promoting physical activity, and researchers from all over the world are employing various interventions in an effort to find more effective and suitable ways to prevent the incidence of chronic diseases and promote health. Thus, it is important to pool and compare their work in an effort to identify the most effective physical activity and health promotion methods.

Although the use of NMA is spreading rather quickly among various research fields, more methodological research is needed because certain interpretational aspects of the approach are poorly understood. Researchers must consider many different aspects of NMA in order to obtain valid simulations and estimations. These aspects include (1) the strength of evidence and risk of bias for each of the comparisons, (2) the analytical challenges, tools, and opportunities in detecting and exploring heterogeneity within and between comparisons, and (3) the interpretation of widely used statistical models and effect measures. Consideration of these points will help ensure high-quality synthesis of evidence and reasonable analysis when conducting an NMA.

An NMA methodology meeting was held at the Johns Hopkins Bloomberg School of Public Health in May 2010. According to Li et al. (2011), the attendees discussed the methodological challenges and research opportunities for NMA relevant to each aspect of the systematic review process. The main points addressed included (1) clearly defining the review question and eligibility criteria, (2) searching for and selecting valid and high-quality studies for data analysis, (3) accurately assessing risk of bias and quality of evidence, (4) conducting a quantitative evidence synthesis, and (5) properly interpreting the results and reporting findings. Although the commentary from the meeting is relatively old, most parts of the discussion are still meaningful for guiding the NMA process. However, new software, such as R, has been developed since the 2010 meeting, but much of the meeting discussion indicates that "most network meta-analysis to date use WinBUGs software", which is limited in functionality and accessibility to the non-statistician. R and other Web-based tools have improved greatly since then. With the rapid growth of NMA studies in the past few years, we have reason to believe that NMA will continue to be a promising analysis technique and will play a significant role in other health-related fields, including physical activity and health promotion.

Introduction in children's body composition and physical activity

With rapid economic development, people's life quality improves along greatly, and more and more people are paying attention on forming a healthy lifestyle not only for themselves but also for their next generations. Engaging in appropriate amount of regular physical activity has been proven to bring lots of health benefits and researchers have also developed interest in carrying out studies focusing on the relationship between physical activity and health as well as effective ways to promote physical activity to broader populations. An abundance of scientific evidence has shown the efficacy of receiving improvement on participants health status after physical activity interventions, with lots of studies reporting exercise treatment for certain kinds of disease to be successful. In fact, the concept of "Exercise Is Medicine" and similar voices is far from new in a worldwide range. Evidence suggests a dose–response relationship such that being active, even to a modest level, is preferable to being inactive or sedentary (Matthews et al., 2016).

Physical activity is especially important for children and adolescents as it affects their physical, psychological and cognitive health greatly (Matthews et al., 2016). The second edition of Physical Activity Guidelines for Americans (2018) made special suggestions for children and adolescents for them to engage more daily physical activities than adults. According to the guidelines, reschool-aged children (ages 3 – 5 years) are encouraged to set their target for moving and playing actively for three hours per day of all intensity activities, while school-aged youth (ages 6 through 17 years) are recommended to have moderate- and vigorous-intensity physical activity for 60 minutes or more each day to achieve substantial health benefits such as muscle- and bone-strengthening. The engagement of physical activity during childhood has been considered vital because general consensus agrees that there should have close relationship among childhood activity and health with adult activity and health. Figure 5 shows hypothetical relationships between them, which was reproduced by Boreham & Riddoch (2001). It appears that childhood activity implies reflection of childhood health and vice versa. It also has a potential causal effect for adult activity and health since children's past experience and lifestyle could have a big impact on their adulthood as well.

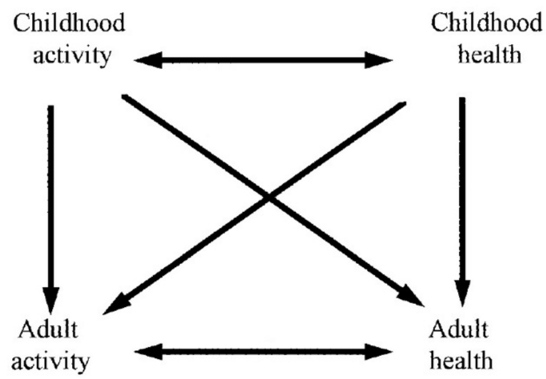


Figure 5. Hypothetical relationships between physical activity and health in children and adults. Adapted from Boreham & Riddoch (2001).

There are plenty of health indicators that could been used to investigate the effectiveness of physical activity or other lifestyle interventions on children and adolescents. For feasibility reasons most studies have chosen to examine body composition as an important indication of children's health status. Body composition usually mainly describes the percentages of fat, bone, water and muscle in human bodies. Body Mass Index (BMI) is a convenient rule of thumb that is often used to broadly categorize a person as underweight, normal weight, overweight, or obese based on one's tissue mass and height. Typically, BMI is calculated with body mass divided by the square of body height. Except for BMI, body fat percentage (BFP) is also frequently used in scientific studies, which is calculated by the total mass of fat divided by total body mass, then multiplied by 100. These are all important and commonly used health indicators regarding people's physical status, and also are chosen to be the variables this case-study is going to focus on.

Prevalence and consequences of insufficient physical activity

Physical Activity Guidelines for Americans (2018) recommended that children and adolescents should do 60 minutes or more of physical activity daily, while adults should do at least 150 minutes a week of moderate - to - intensity, or 75 minutes a week of vigorous - to - intensity aerobic physical activity, or an equivalent combination of moderate - and vigorous - intensity aerobic activity, for substantial health benefits. According to World Health Organization (WHO), insufficient physical activity is one of the 10 leading risk factors for global mortality. According to the latest published Global Health Observatory (GHO) data by WHO, people who are insufficiently physically active have a 20% to 30% increased risk of all-cause mortality compared to those who met the 150-minute minimum guidelines mentioned before. Warburton et al. (2006) synthesized evidence showing that maintaining regular physical activity plays an important role in determining daily energy expenditure and also has an effect in reducing the risk of ischemic heart disease, diabetes, breast and colon cancer. Furthermore, it lowers the risk of stroke, hypertension, and depression.

WHO's GHO data report for prevalence of insufficient physical activity shows that in 2016, 23% of men and 32% of women aged 18+ years were insufficiently physically active in a global range. Over the past 15 years, levels of insufficient activity did not improve but instead, the percentage of people having sufficient activities went down from 28.5% in 2001 to 27.5% in 2016. According to Centers for Disease Control and Prevention (CDC), physical inactivity and poor diet was responsible for an average of 16% of deaths each year and may lead to diseases such as cardiovascular disease, type II diabetes and some types of cancers. Parts of the reason why overall people are engaging less physical activity may be due to the Industrial Revolution and the increased technology since then. People's lifestyle has been changed progressively, for example, it greatly reduced occupational physical activity in one hand and reduced necessary exercise for daily transportations as well. In fact, estimates of population participation of physical activity have consistently shown a progressive decrease (Brownson et al., 2005). Ng & Popkin's review (2012) indicated that these trends are not limited to the developed world since it appeared that in low- and middle- income countries (LMICs) in the developing world are also experiencing a rapid physical activity transition "associated with rapid acculturation to sedentary living". Statistically looking, the WHO GHO report shows that high income countries had more than double the prevalence of physical inactivity compared to low income countries, with 32% of men and 42% of women being insufficiently physically active for the former and 13% of men and 19% of women in low income countries. We could tell that a major public health challenge the world is facing together is to encourage the growing sedentary population to adopt a more active lifestyle and thereby improve population health status.

One of the major consequences of population decreases in physical activity (and increase in physical inactivity) is the associated increase in noncommunicable diseases (NCDs). According to the WHO, NCDs account for an estimated 60% of all deaths and are thought to be the greatest cause of morbidity and mortality in 2011. Kohl et al. (2012)

published their review where globally physical inactivity was described as a pandemic and the researchers managed to gather evidence for the benefits of physical activity for health in order to promote physical activity and improve the health of populations. However, situations have not improved so much as expected from Kohl's review. Guthold et al. (2018) carried out a pooled analysis of more than three hundred population-based surveys regarding worldwide trends in insufficient physical activity from 2001 to 2016, and found out during the 15 years, high-income countries have showed a 5% increase in inactivity levels. They concluded that if current trends continue, the 2025 global physical activity target will not be met, which is to reach a 10% relative reduction in insufficient physical activity.

Prevalence and consequences of childhood obesity

One of the most commonly used definition of obesity is described using cut points of BMI. CDC standard for adults to be overweight is if one's BMI is 25.0 to <30; if one's BMI is 30.0 or higher, it falls within the obese range. For children and teens, CDC uses percentile to help categorize their weight status: if one's BMI falls 85th to less than the 95th percentile, he/she would be considered overweight; one will be counted as obese if his/her BMI falls 95th percentile or greater. Although BMI does not measure body fat directly, it is considered to be a reliable alternative to direct measures of body fat with research shown that BMI is correlated with direct measures of body fat, such as skinfold thickness measurements, dual energy x-ray absorptiometry (DXA) and other methods (Barlow, 2007).

According to latest CDC statistics, the prevalence of obesity was 18.5% and had affected about 13.7 million children and adolescents. In 2015 – 2016, obesity prevalence was 13.9%, 18.4% and 20.6% for children among 2- to 5- year-olds, 6- to 11-year-olds, and 12- to 19-year-olds, respectively. Obesity prevalence increased in both adults and youth during the 18 years between 1999–2000 and 2017–2018 (Hales et al., 2020). Previous analyses showed no change in prevalence among youth between 2003–2004 and 2013–2014 and the observed increase in the prevalence of obesity between 2015 – 2016 and 2017 – 2018 was not significant (Hales et al., 2020).

Obesity during childhood can have negative effects on the body both immediate and also in the long-term. Specifically, children who are overweight or obese also have shown to have a greater risk for bone and joint difficulties, sleep apnea, and social or psychological issues, such as poor self-esteem. From a long-term perspective, obesity in children and adolescents has been shown to track into adulthood. Obese children and adolescents are more likely to develop high blood pressure and high cholesterol, which are risk factors for cardiovascular disease (CVD) (Umer et al., 2017); they are also likely to have increased risk of type II diabetes, stroke, several types of cancer, and osteoarthritis during adulthood according to CDC facts for childhood obesity.

BMI was chosen as one of the most important outcome variables for this analysis because it’s most widely reported in research studies as demographic information or as an intervention outcome. Children’s BMI are relatively easy to get as it requires simple equipment and also low cost financially. Commonly though, for the assessment of children body composition, BMI percentile could also be a preferred way for researchers to report as part of the results. However, for secondary data analysis, percentile data is not as interpretable as the raw score of BMI. To adjust the possible influence of age or other possible variables on the group, we also included the standard deviation score of BMI (BMI z-score). Body fat percentage is the last variable included in this set of analysis and is also relatively the most difficult one to obtain. Usually, percentage of body fat is obtained through dual-energy X-ray absorptiometry (DXA), which is more financially demanding than measuring and calculating children's BMI and z-score.

Children's weight and body composition is greatly affected by their internal hormones during adolescence period. To avoid introducing more complicated influencers during childhood development, we decided to choose children who are 12 years old and below since they are less likely to be affected by the development hormones than adolescents.

Physical activity promotion programs

Researchers from around the world are trying different approaches on solving the problem of changing children's body composition. There are simple interventions as controlling children's physical activity behavior by enrolling them in exercise programs to make sure they are engaged in certain amount of physical activity each week; or providing education courses about healthy lifestyle instructions on daily media use control, forming healthy sleep patterns, improving self-regulatory habits etc. Directly educating and encouraging children, teachers, and parents to live active and healthy lifestyles are all important aspects and should have a great positive influence in children's lives. Instructions on dietary habits and access to heathy snacks are also interventions that could have directly impact on children's body composition. Additionally, school environmental settings, for example, the physical environment, organization of school breaks, playing during school time, and sports facilities etc. could be altered to promote more physical activity in some studies. Other research combined some of the simple interventions above and made more complex and multidimensional lifestyle interventions. Theoretically, the more complex the intervention is designed, the bigger the effect size would be for all the variables. However, that is not always the case as shown in some research studies. Therefore, we're interested in finding out whether there are some kinds of intervention that could show more effectiveness than the other approaches.

Simply designed RCTs usually have fewer aspects of intervening compared to multidimensional RCT designs. The former is more specific and direct on single-arm intervention with more aspects held controlled. For mixed-arms interventions, the more aspects of change introduced to participants' lives, the lower external validity the experiment is going to have, because it is going be harder for the conditions to be all promoted at the same time together in real life. Therefore, both kinds of intervention designs have their own advantages and disadvantages.

There are plenty of strategies to increase physical activity targeting aspects including specifically designed school and youth programs, community-wide campaigns, social and individual supporting actions, providing access to more sites for physical activity, and so on. Decision-makers, including parents, teachers, school administrators, health care providers, and policymakers are encouraged to make additional efforts to facilitate opportunities for physical activity for children and youth. In fact, there have been plenty of research studies testing effectiveness of a variety of different programs from many approaches trying to promote physical activity, so it will be meaningful to pull all the studies together and carry out a systematic review to compare which paths might be better than the others.

To the best of the authors' knowledge, no previous network meta-analysis has examined the effects of physical activity, nutrition status, education contents, environmental change and the combination of two or more of the single aspects mentioned above on body composition in school-aged children. Therefore, the primary objective of this study was to conduct a systematic review with network meta-analysis of randomized controlled trials to determine the effects of multidimensional lifestyle intervention on BMI, BMI z-score, and percentage of body fat in young children population and provide a ranking of treatments as a practical conclusion.

Methods

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement was followed when preparing to conduct the network meta-analysis.

Study eligibility criteria

We included studies that are 1) data-based articles published in English between 2000 and 2020; 2) used randomized controlled trial design; 3) subjects aged between 4 to 12 years old and did not suffer from any physical or mental illness; and 4) investigated some type of intervention on body mass index (BMI), BMI z-score (BMIz), and body fat percentage among children.

We only included study participants' age between four to twelve years old to avoid possible weight change due to hormonal secretion during adolescence, as well as the rapid growth period in the first several years since infancy. As a secondary data analysis, we tried to make the original data reliable by retrieving reported results from randomized controlled trails only, because the nature of RCT could avoid lots of confounding factors during the experiment, both expected and unexpected aspects. The three variables (BMI, BMI z, and body fat percentage) were chosen because they are convenient to get and therefore became some of the most commonly reported data regarding children's general health status. Also, ratio and percentages make the results more compatible than absolute values such as weight or height, making it possible to compare a wider age range of children. Lastly, the variables chosen all have unified standards in measuring or calculating, compared with other commonly used variables such as physical activity intensity, which has several different methods to decide the specific cut points for physical activity to be categorized as light, moderate or vigorous.

Literature search methodology

We searched the following databases for published articles and conference abstracts: PubMed, Embase, Web of Science and Google Scholar. The search was performed originally in August 2019, then the results were updated in February 2020. Only papers published in English were included and date restrictions were set as less than ten years from now (2009-). Keywords searched were as follows: "children or kids or youth or child", and "BMI or body mass index or weight or body fat or obesity or overweight", and "physical activity or exercise or fitness or physical exercise or sport", and “interventions or strategies or best practices or treatment or therapy or program or management”, and “randomized controlled trials or RCT or randomised control trials”, and “multidimensional interventions or strategies or therapy or program”, or “nutrition or diet or food or nourishment or food intake or eating”, or “environment”.

Study selection and data extraction

After a primary screening of titles and abstracts, the full texts of all potentially eligible studies were retrieved. A total of 657 published studies on lifestyle intervention programs were retrieved with 61 studies meeting the inclusion criteria given above. Study investigators were contacted through email if clarification or results regarding details of the study was needed. Data was extracted from each eligible study and formatted in a way convenient for further data analysis. Information retrieved included characteristics of study, subjects and interventions details, as well as mean and standard deviations for all outcome measures. Specifically, data extraction for comparisons was completed for 10 intervention categories. The first five were all simple method of intervention, and the last five interventions were more complex and combined two or three of the simple methods in the earlier categories. Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Statistical analysis

We extracted point estimates of relevant means and standard deviations (SDs) from individual studies. This includes mean difference and its associated SD between the comparison groups after the experiment. If the SD associated with the mean difference was not reported, and a 95% confidence interval (CI) or standard error (SE) was reported instead, then the provided CI or SE were used for the computation of SD:

$$(\sigma_{\bar{x}}) = \frac{\sigma}{\sqrt{n}} \quad , \quad \left(\bar{x} - z^* \frac{\sigma}{\sqrt{n}}, \bar{x} + z^* \frac{\sigma}{\sqrt{n}}\right)$$

Moreover, if the mean difference and associated SD between the comparison groups was not reported, we extracted the arm-specific mean and SD both before and after the treatment to calculate these quantities. The formula for computing the mean change and its associated SD for a specific arm referred to Dias *et al*.

$$s_p^2 = \frac{\sum_{i=1}^{k}(n_i - 1)s_i^2}{\sum_{i=1}^{k}(n_i - 1)} = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2}{n_1 + n_2 + \cdots + n_k - k}$$

The mean of the comparison groups was assumed to follow a normal distribution, and, therefore, the calculation of mean difference (and associated SD) was based on the subtraction of two independent normal distributions (Wang *et al.* 2018). If the equal variance assumption is valid (i.e. the variances of the pairwise groups were comparable), the pooled SD was calculated as $S_{pooled} = \sqrt{\frac{S_1^2(n_1-1)+S_2^2(n_2-1)}{n_1+n_2-2}}$, where n1, n2 are the sample size for the two arms and S1, S2 are the arm-specific SDs; if the equal variance assumption does not hold, the following formula was applied to compute the pooled SD: $S_{pooled} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$. Then a network meta-analysis was performed to simultaneously compare the efficacy of different treatment options for managing children's body composition (i.e. BMI, BMI z-score and percentage of body fat) and blood pressure (SBP and DBP).

All statistical analyses were carried out using R version 3.6.1, running in RStudio version 1.1.463.

Results

Performing arm-based network meta-analysis

<u>*Convergence status for treatments on BMI, BMIz, and body fat percentage*</u>

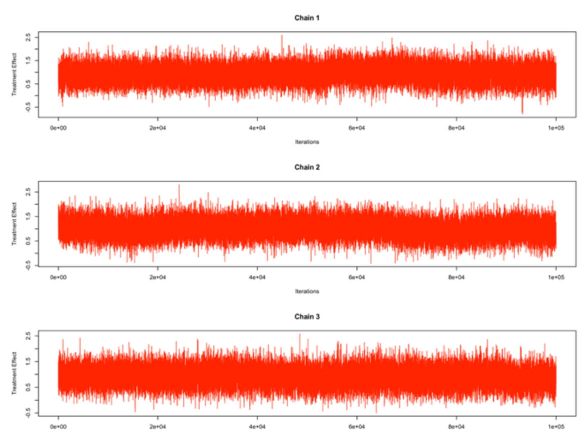
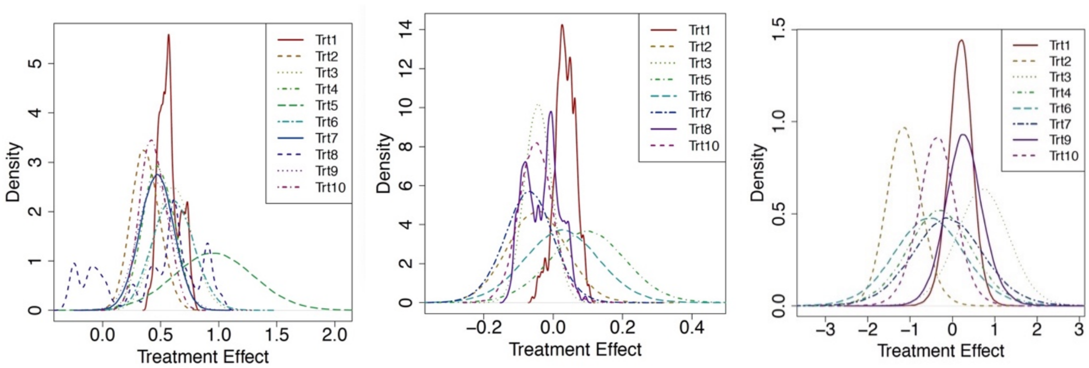


Figure 6. Trace plots generated by R for log odds ratio.

The above figure shows the trace plots of the variable BMI for treatments group five as an example. Three chains in the trace plots are drawn, and each chain shows evidence that the posterior sample of LOR are drawn from a stationary distribution. Generally speaking, most of the model simulation we used have convergence status similar with the figure shown above, however, there are several trace plots for the control group (e.g. for BMI analysis) that didn't seem as satisfactory (convergence diagnostic reported number around 2, while ideally the number should be around 1), and an increase on the iteration times (up to 750,000 times maximum during trials) didn't help solve the problem too much. Although this is not ideal, there is hardly anything more we can make a change for now.

<u>Density plots of MCMC simulation of children's BMI, BMI z and percentage body fat</u>



A B C

Figure 7. Posterior density plots generated for children's BMI, BMI z and percentage body fat.

A posterior density plot (Figure 6) for treatment-specific absolute risks is generated and smoothed by R. These plots show visualized treatment effects, and we may also evaluate treatment differences from them at the same time. The density plots could directly show the number of each treatment included in the analyses from the top right legend. It is obvious that in all three plots, usual care control treatment (red line) has the highest peak, which makes sense since most studies set usual care situation as their control group. The rest of the treatments don't have very a big difference in case of counts of inclusion. It can also visualize and evaluate treatment effects by comparing their peaks over the x-axis. For example, in Figure 6 (C), posterior density of treatment 2 has its peak located on the left of all the other treatments, and only overlaps with the density of 3 in tiny regions, which shows physical activity (trt2) has apparently better treatment effects than education (trt3) since here we treat a decrease in percentage of body fat as an expected result from the interventions.

Statistical output for outcome variables using mean difference data.

BMI

Table 2. Mean (SD) and Median (95% CI) of different treatment methods for children's BMI.

Treatment	Mean	SD	Median	95%CI
Trt1	0.575	-0.089	0.563	(0.432, 0.754)
Trt2	0.37	-0.123	0.366	(0.136, 0.618)
Trt3	0.61	-0.158	0.609	(0.304, 0.923)
Trt4	0.481	-0.134	0.479	(0.223, 0.741)
Trt5	0.955	-0.346	0.955	(0.275, 1.640)
Trt6	0.608	-0.176	0.606	(0.274, 0.959)
Trt7	0.466	-0.144	0.467	(0.180, 0.744)
Trt8	0.43	-0.387	0.565	(-0.266, 0.986)
Trt9	0.454	-0.154	0.454	(0.155, 0.760)
Trt10	0.438	-0.114	0.433	(0.225, 0.667)

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

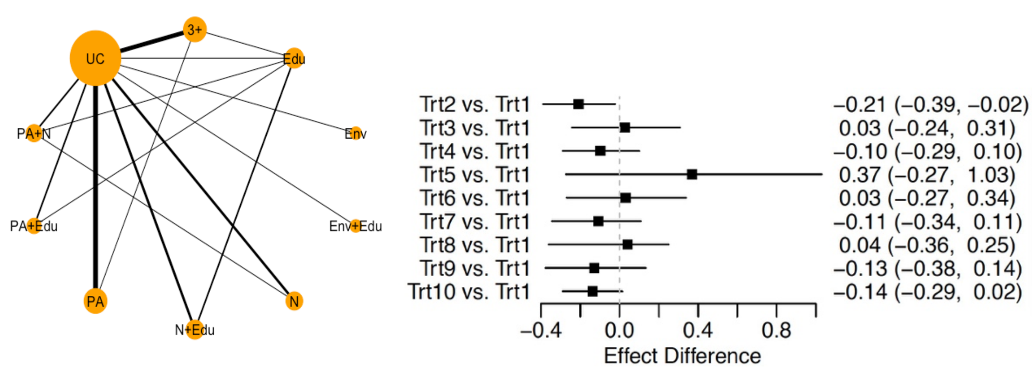
Table 3. Estimates of medians and 95% CI for the log odds ratio for each treatment.

	Trt1	Trt2	Trt3	Trt4	Trt5	Trt6	Trt7	Trt8	Trt9	Trt10
Trt1	--									
Trt2	-0.204 (-0.399, -0.008)	--								
Trt3	0.044 (-0.351, 0.238)	0.248 (-0.092, 0.608)	--							
Trt4	-0.094 (-0.129, 0.326)	0.110 (-0.190, 0.397)	-0.138 (-0.522, 0.210)	--						
Trt5	0.382 (-1.080, 0.292)	0.586 (-0.116, 1.310)	0.338 (-0.404, 1.080)	0.477 (-0.231, 1.210)	--					
Trt6	0.040 (-0.365, 0.277)	0.244 (-0.126, 0.619)	-0.004 (-0.431, 0.412)	0.134 (-0.243, 0.531)	-0.342 (-1.110, 0.405)	--				
Trt7	-0.111 (-0.122, 0.382)	0.093 (-0.241, 0.388)	-0.155 (-0.546, 0.177)	-0.017 (-0.361, 0.301)	-0.494 (-1.250, 0.213)	-0.152 (-0.579, 0.234)	--			
Trt8	-0.189 (-0.981, 0.468)	0.041 (-0.317, 1.240)	-0.175 (-0.615, 1.010)	-0.070 (-0.433, 1.180)	-0.422 (-1.250, 0.951)	-0.170 (-0.631, 1.050)	-0.052 (-0.439, 1.230)	--		
Trt9	-0.120 (-0.168, 0.382)	0.084 (-0.242, 0.435)	-0.163 (-0.542, 0.211)	-0.027 (-0.346, 0.336)	-0.501 (-1.240, 0.232)	-0.161 (-0.573, 0.267)	-0.009 (-0.351, 0.396)	0.018 (-1.170, 0.433)	--	
Trt10	-0.121 (-0.038, 0.282)	0.084 (-0.166, 0.329)	-0.165 (-0.488, 0.134)	-0.026 (-0.303, 0.252)	-0.502 (-1.220, 0.189)	-0.161 (-0.524, 0.191)	-0.009 (-0.286, 0.302)	0.050 (-1.150, 0.382)	-0.001 (-0.329, 0.300)	--

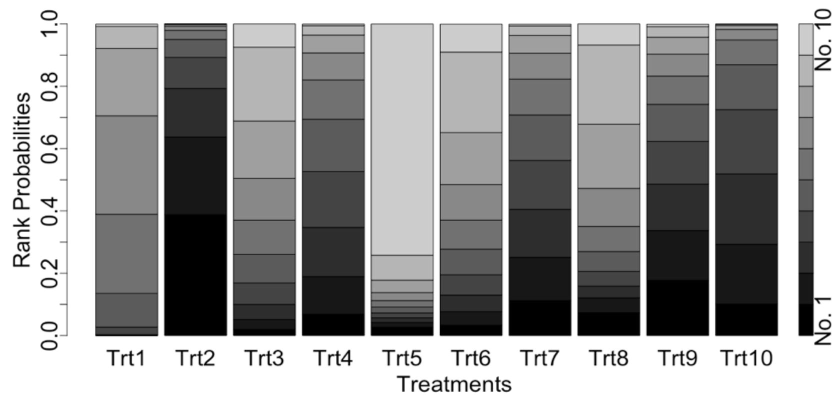
Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 2 shows the mean (SD) and median (95% CI) of different treatment methods for children's BMI. The Mean (SD) variable contains posterior sample means followed with sample standard deviations; while the Median (95% CI) variable contains posterior medians followed with 95% CIs. In general, mean and median of each treatment group look quite similar with each other. Since the data is representing the average change of BMI from post intervention to baseline, no matter the number is positive or negative, the smaller the number is, the more efficient the treatment would be. Therefore, it appears in the table that treatment 2 and 8 are two of the lowest among all the other treatments.

The estimates of median and 95% CI for the log odds ratio of each treatment are displayed in a 10 * 10 matrix in Table 3. The estimates of mean and standard deviations could also be obtained through software and is attached in the appendix. Using the median (95% CI) table here could help examine whether the effect difference between each pair of the treatments are true or due to chance. Under the null hypothesis, ORs and RRs should be one, while LORs, LRRs, and RDs should be 0. The element in the *ith* row and *jth* column is the estimated log odds ratio of treatment *i* compared to treatment *j*. From what the output shows, the 95% CI of the log odds ratio for treatment 1 vs. 2 is 0.204 (95% CI: 0.008, 0.399), the confidence interval of which does not contain 0; therefore, we have 95% of confidence to reject the null hypothesis and say that physical activity (Trt2) has a better intervention effect on reducing children's BMI compared with usual care control group (Trt1). This appears to be the only pair of comparison that has shown significance in Table 3. In conclusion, results from Table 3 showed physical activity has a significant impact on the children's BMI decrease.



A. B



C

Figure 8. (A) Network plotting for treatment used in BMI observations. (B) Treatment effect differences between control group and all intervention groups. (C) Plots of treatment rank probabilities for BMI.

Figure 7 (A) is the obtained network plot for the single variable BMI using mean difference data change from post-intervention to baseline. It appears that all ten different intervention arms were used in BMI observations, and all intervention arms have made direct comparisons with the usual care control group, since every type of treatment node is connected to the node of treatment 1. Expect from those, there are six other sets of direct comparisons, which fall between education (Edu) versus education plus physical activity (PA + Edu), education (Edu) versus physical activity plus nutrition (PA + N), education (Edu) versus nutrition plus education (N + Edu), education (Edu) versus multiple component treatment (3+), nutrition (N) versus physical activity plus nutrition (PA + N), and lastly between physical activity (PA) versus multiple component treatment (3+). For other pairs of treatment, e.g. education versus environment, since there was no edge between them, meaning no study directly compared treatments 3 and 5. Therefore, only indirect evidence is available when comparing treatments 3 and 5 in this analysis, for example, using the comparisons of treatments 2 vs. 1 and 3 vs. 1 as one of the indirect paths.

The Contrast figure (Fig. 7 B) above shows the differences of effect sizes comparing each intervention arms with the usual care group. The column left to the effect size lines indicates each pair of comparison, while the right column indicates the statistical significance of the differences by providing the 95% confidence interval. According to the contrast figure of BMI, only physical activity (Trt 2) has a significant impact on the decrease of children's BMI, compared with the usual care control group, with a larger effect size of -0.21 (95% CI: -0.39, -002), on average.

The rank probability figure (Fig. 7 C) shows predicted rankings of each intervention arm including the no intervention control arm. In plot Fig. 7 (C), each treatment falls on the x-axis and the vertical bars represent the probabilities for each specific treatment to be ranked from 1 to 7. The legend laid on the right side indicated that a darker area represents the probability of having a higher rank, i.e. the black areas show the probabilities of being the best ranked treatment while the white areas show the probabilities to be ranked last. Therefore, from Figure 7 (C), treatments 2 and 9 have much higher probabilities of being the best treatment, compared with other treatments in their respective studies. Therefore, for BMI, it appears physical activity (Trt 2) and physical activity plus nutrition (Trt 9) are predicted to be more effective than the other interventions, with physical activity to be even higher ranked than physical activity and nutrition combined. Similar with each other, the rest of the treatments have lower possibility to be the best intervention for decreasing children's BMI.

BMI z-score

Table 4. Mean (SD) and Median (95% CI) of different treatment methods for children's BMI z-score.

Treatment	Mean	SD	Median	(95% CI)
Trt1	0.034 (0.030)	-0.03	0.034	(-0.033, 0.090)
Trt2	-0.047(0.086)	-0.086	-0.047	(-0.216, 0.121)
Trt3	-0.044(0.040)	-0.04	-0.044	(-0.121, 0.035)
Trt5	0.100 (0.111)	-0.111	0.099	(-0.119, 0.320)
Trt6	0.024 (0.110)	-0.11	0.025	(-0.193, 0.242)
Trt7	-0.071(0.072)	-0.072	-0.07	(-0.216, 0.071)
Trt8	-0.032(0.049)	-0.049	-0.025	(-0.113, 0.055)
Trt10	-0.049(0.051)	-0.051	-0.05	(-0.147, 0.054)

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 5. Estimates of medians and 95% CI for the log odds ratio of each treatment for BMI z-score.

	Trt1	Trt2	Trt3	Trt5	Trt6	Trt7	Trt8	Trt10
Trt1	--							
Trt2	-0.088 (-0.258, 0.082)	--						
Trt3	-0.090 (-0.182, -0.008)	-0.003 (-0.189, 0.181)	--					
Trt5	0.058 (-0.161, 0.277)	0.146 (-0.125, 0.416)	0.148 (-0.079, 0.380)	--				
Trt6	-0.017 (-0.151, 0.112)	0.071 (-0.141, 0.280)	0.073 (-0.069, 0.221)	-0.075 (-0.327, 0.176)	--			
Trt7	-0.115 (-0.262, 0.030)	-0.027 (-0.244, 0.189)	-0.024 (-0.174, 0.131)	-0.172 (-0.430, 0.082)	-0.098 (-0.287, 0.092)	--		
Trt8	-0.033 (-0.137, 0.056)	0.049 (-0.135, 0.239)	0.054 (-0.076, 0.178)	-0.097 (-0.326, 0.140)	-0.021 (-0.180, 0.141)	0.076 (-0.089, 0.250)	--	
Trt10	-0.091 (-0.194, 0.011)	-0.003 (-0.192, 0.188)	0.000 (-0.115, 0.125)	-0.148 (-0.383, 0.088)	-0.075 (-0.226, 0.085)	0.024 (-0.142, 0.196)	-0.052 (-0.186, 0.079)	--

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 4 shows the mean (SD) and median (95% CI) of different treatment methods for children's BMI z-score. Similarly, with BMI, the mean and median of each treatment group look quite similar with each other for BMI z-score as well. Since the data is representing the average change of BMI z-score from post intervention to baseline, no matter the number is positive or negative, the smaller the number is, the more efficient the treatment would be. Therefore, it appears in the table that treatment 7 and 2 are two of the lowest among all the other treatments.

The estimates of medians and 95% CI of the log odds ratio of children's BMI z-score are displayed in an 8 * 8 matrix in Table 5. The estimates of mean and SD could also be obtained through software and is attached in the appendix. As mentioned earlier in BMI section, the element in the *ith* row and *jth* column is the estimated log odds ratio of treatment *i* compared to treatment *j*. The only pair of comparison with a significant difference is colored in red in Table 5, which is -0.090 (95% CI: -0.182, -0.008) between treatment 3 (Environment) and 1 (Control). Therefore, we have 95% of confidence to reject the null hypothesis and say that general education (Trt3) may have a better intervention effect on reducing children's BMI z-score compared with usual care control group (Trt1).

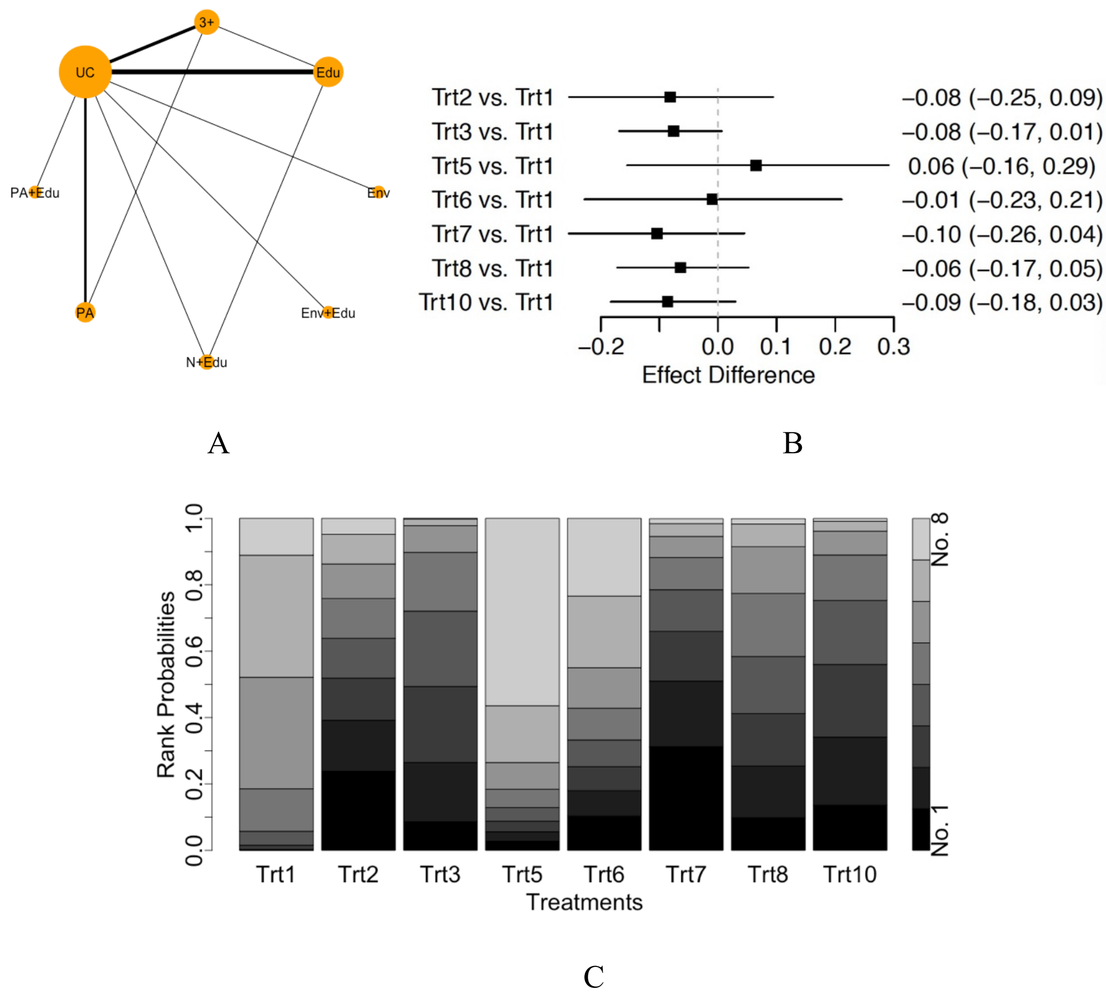


Figure 9. (A)Network plotting for treatment used in BMI z-score observations. (B)Treatment effect differences between control group and all intervention groups. (C) Plots of treatment rank probabilities for BMI z-score.

Figure 8 (A) is the obtained network plot for the single variable BMI z-score using the mean change of children's BMI z-score from post-intervention to baseline. It appears that there are 8 different intervention arms observed in BMI z-score studies, and all intervention arms have made direct comparisons with the usual care control group. Except from those, there are three other sets of direct comparisons, which are between education (Edu) versus multiple component (3+), education (Edu) versus nutrition plus education (N + Edu), and physical activity (PA) versus multiple component treatment (3+). For other pairs of treatment, only indirect evidence is available for this case study.

The Contrast figure (Fig. 8 B) above shows the differences of effect sizes between each intervention arms and usual care group. According to the contrast figure of BMI z-score, there does not appear to be any treatment category that might have a significant impact on the decrease of children's BMI z-score, compared with the usual care control group.

The rank probability figure shows predicted rankings of each intervention arm including no intervention control group. From Figure 8 (C), treatments 2 and 7 have relatively higher probabilities of being the best treatment, compared with other treatments in their respective studies. Therefore, for the variable of BMI z-score, it is estimated that physical activity (Trt 2) and nutrition plus education (Trt 7) are likely to be more effective than the other six interventions included in the analysis, with physical activity to be slightly lower ranked than nutrition and education combined. The rest of the treatments have similar lower possibility to be ranked as the best intervention for decreasing children's BMI z-score using this analysis.

Body fat percentage

Table 6. Mean (SD) and Median (95% CI) of different treatment methods for children's body fat percentage.

Treatment	Mean	SD	Median	95%CI
Trt1	0.189	-0.289	0.194	(-0.396, 0.750)
Trt2	-1.14	-0.426	-1.14	(-1.950, -0.270)
Trt3	0.706	-0.65	0.711	(-0.587, 1.980)
Trt4	-0.302	-0.811	-0.303	(-1.910, 1.310)
Trt6	-0.543	-0.867	-0.544	(-2.250, 1.190)

Trt7	-0.104	-0.874	-0.108	(-1.830, 1.640)
Trt9	0.278	-0.457	0.271	(-0.611, 1.210)
Trt10	-0.36	-0.461	-0.359	(-1.280, 0.553)

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

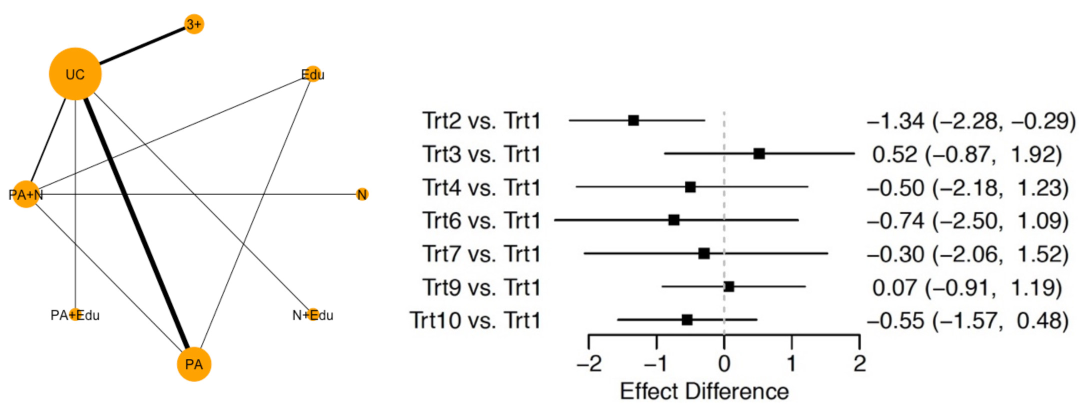
Table 7. Estimates of medians and 95% CI for the log odds ratio of each treatment for children's body fat percentage.

	Trt1	Trt2	Trt3	Trt4	Trt6	Trt7	Trt9	Trt10
Trt1	--							
Trt2	-1.200 (-2.020, -0.319)	--						
Trt3	0.552 (-0.779, 1.870)	1.760 (0.304, 3.130)	--					
Trt4	-0.468 (-2.060, 1.160)	0.735 (-0.974, 2.410)	-1.020 (-2.960, 0.966)	--				
Trt6	-0.733 (-2.410, 0.991)	0.470 (-1.330, 2.240)	-1.280 (-3.310, 0.795)	-0.268 (-2.490, 1.990)	--			
Trt7	-0.289 (-1.970, 1.440)	0.914 (-0.897, 2.700)	-0.840 (-2.880, 1.260)	0.182 (-2.080, 2.430)	0.441 (-1.870, 2.770)	--		
Trt9	0.098 (-0.840, 1.120)	1.300 (0.193, 2.400)	-0.454 (-1.860, 1.040)	0.561 (-1.100, 2.300)	0.835 (-1.030, 2.710)	0.392 (-1.500, 2.280)	--	
Trt10	-0.538 (-1.490, 0.418)	0.666 (-0.477, 1.750)	-1.090 (-2.580, 0.431)	-0.069 (-1.840, 1.680)	0.191 (-1.660, 2.020)	-0.252 (-2.110, 1.590)	-0.639 (-1.900, 0.571)	--

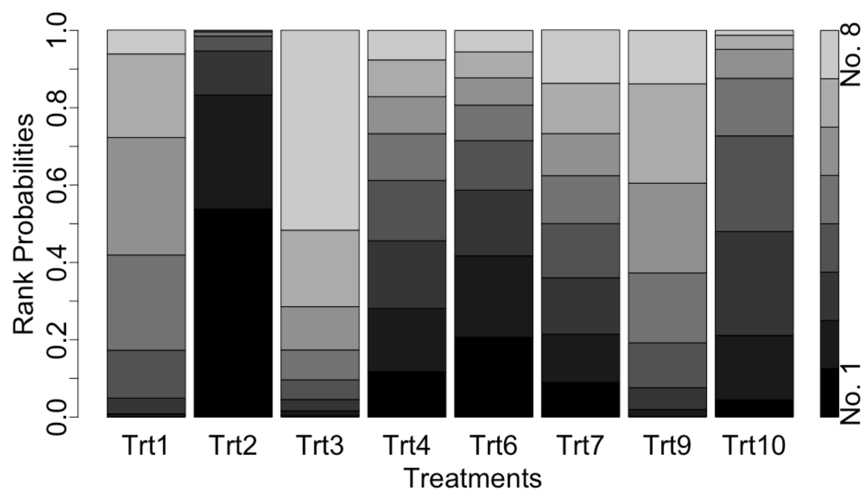
Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 6 shows the mean (SD) and median (95% CI) of different treatment methods for children's percentage of body fat. Not surprisingly, similar with previous occasions in BMI and BMI z-score, the mean and median of each treatment group look quite similar with each other for percentage of body fat as well. Since the data is representing the average change of the percentage of children's body fat from post intervention to baseline, no matter the number is positive or negative, the smaller the number is, the more efficient the treatment would be. Therefore, it appears in the table that treatment 2 is apparently the lowest among all the other treatments, indicating that physical activity may possibly show more positive impact on children's body fat percentage than other approaches.

The estimates of medians and 95% CI for the log odds ratio of each treatment are displayed in an 8 * 8 matrix in Table 7 for children's body fat percentage. The estimates of means and SDs could also be obtained through software and are attached in the appendix. Physical activity (Trt 2) is proved to have significantly more impact on reducing children's percentage of body fat compared with control group (Trt 1) (LOR = -1.200, 95% CI: -2.020 ~ -0.319) and education group (Trt 3) (LOR = 1.760, 95% CI: 0.304 ~ 3.130).



A B



C

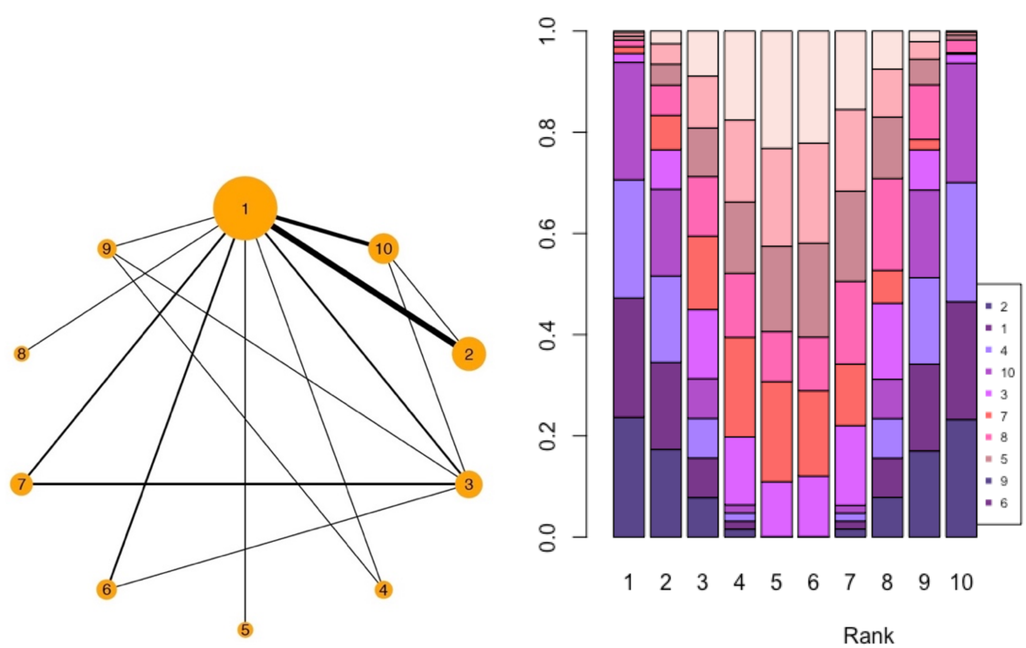
Figure 10. (A) Network plotting for treatment used in body fat percentage observations. (B)Treatment effect differences between control group and all intervention groups. (C) Plots of treatment rank probabilities for body fat percentage.

Figure 9 (A) is the obtained network plot for the single variable body fat percentage. It appears that there are 8 different intervention arms were used investigating body fat percentage observations. All intervention arms have made direct comparisons with the usual care control group, except for education and nutrition intervention groups. Among other intervention arms, there are four sets of direct comparisons, which are between education (Edu) versus physical activity (PA), and physical activity plus nutrition (PA + N), respectively; physical activity plus nutrition (PA + N) versus nutrition (N), and physical activity (PA), respectively. For other pairs of treatment, only indirect evidence is available.

The contrast figure (Fig. 9 B) above shows the differences of effect sizes between each intervention arms and usual care group for children's body fat percentage. According to the plot, only physical activity (Trt 2) has a significant impact on the decrease of children's body fat percentage, compared with the usual care control group, with a larger effect size of -1.34 (95% CI: -2.28 ~ -0.29), on average.

The rank probability figure shows predicted rankings of each intervention arm for children's body fat percentage. From Figure 9 (C), treatment 2 has a much higher probability to be the best treatment, compared with all the other treatments. Therefore, for the variable of children's body fat percentage, physical activity (Trt 2) is predicted to be the most effective treatment than the other six interventions included in the analysis. The rest of the treatments have lower possibility similar to each other to be a better intervention for decreasing children's body fat percentage.

Statistical output for outcome variables using pre-post combined data analysis.



A **B**

Figure 11. (A) Network plot for all treatment used in BMI, BMI z-score and body fat percentage observations. (B) Plot of treatment rank probabilities for combined effects of three variables.

Table 8. Mean and median rank, SUCRA rank and the probabilty to be the best treatment for each intervention.

	Mean Rank	SD	Rank Median	Prob(Best)	SUCRA	SD	SUCRA_rank
2	5.6672	1.9018	6	0.008	0.4814	0.2113	9
1	5.4127	1.5964	5	0.0029	0.5097	0.1774	2
4	5.5081	3.8393	7	0.2337	0.4991	0.4266	6
10	4.9583	1.801	5	0.0135	0.5602	0.2001	1
3	5.4553	1.799	5	0.0071	0.505	0.1999	3
7	5.9426	2.2816	6	0.0128	0.4508	0.2535	10
8	5.4713	3.8399	4	0.2366	0.5032	0.4267	4
5	5.4834	3.8396	4	0.2358	0.5018	0.4266	5
9	5.5192	3.8389	7	0.2319	0.4979	0.4265	7
6	5.5818	2.2693	6	0.0178	0.4909	0.2521	8

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Results in the previous section mostly used pre-processed data to obtain the mean change of different variables from post-intervention to baseline, except for some studies which reported the mean differences and standard deviation in their original paper (procedure described in Methods section 2.4), and the variables were treated separately to see treatment effects for different variables. After this, we used a different method to carry out the analysis, by using the original baseline and post-intervention data published in papers, and focus on the combined effects of each treatment on children's body composition as a whole (still including BMI, BMI z-score and percentage body fat as variables).

Figure 10 (A) shows the network plot for all treatments used in BMI, BMI z-score and body fat percentage observations. This network plot should show the same result as a stack up of all three network plots drew before for the variables separately. All intervention arms were used, and all intervention arms have made direct comparisons with the usual care control group. Among other intervention arms, there are six sets of direct comparisons, which are between multicomponent intervention (3+) versus physical activity (PA), education (Edu) versus physical activity plus education (PA + Edu), nutrition plus education (N + Edu), physical activity plus nutrition (PA + N), and multicomponent interventions (3+) respectively, and the last pair is between nutrition (N) and physical activity plus nutrition (PA + N). For other pairs of treatment, only indirect evidence is available for ranking estimation.

The rank probability figure shows in Figure 10 (B) predicted rankings of each intervention arm including usual care control arm. In the plot, ranks from 1 to 10 fall on the x-axis. Each vertical bar contains different colors where each color represents an intervention category, details provided in the legend beside the plot. The y-axis represents probability for each different color bar at each rank. However, in Figure 10 (B), treatments 2 has the longest bar in both the first rank as well as the last rank, making it difficult to decide the probability for treatment 2 to be a better intervention or not. Besides, the plot appears to be symmetric for the rest of the ranks as well.

Therefore, final decision needs to be made using results provided in Table 8 as the main reference. The further left column shows 10 different treatment arms. Mean rank seems to be quite similar among all the arms (all between 5 to 6), except for treatment 10 to be 4.95 (1.80) which is the lowest estimation as well as the best prediction. In the column of “probability to be the best treatment”, nutrition, environmental changes, nutrition plus education, environment change plus education were predicted to be slightly higher than the rest of the treatments. Finally, surface under the cumulative ranking curve (SUCRA) estimations and SUCRA rank were provided in the table. SUCRA is a numeric presentation of the overall ranking and presents a single number associated with each treatment. SUCRA values should range from 0 to 100%, while the ones generated in the table showed all the values are among 40% - 60% (?). According to the SUCRA rank column, the best intervention predicted using this method should be the multiple component treatment (trt 10).

Test for inconsistency

The current version of R package *pcnetmeta* does not have the function to detect inconsistency in arm-based network meta-analysis. Future work would add functions for network consistency assessment (Zhao, Hodges, Ma, Jiang, & Carlin, 2016).

We may use different R packages for inconsistency check such as *gemtc* and *netmeta*, or try different software such as STATA.

Discussion

Two different analysis methods were used to address whether one or more intervention approaches are more effective than the other(s) in promoting children's body composition. The first method treated all interested variables separately and used mean difference from post-intervention to baseline as input data; the second method combined different variables together, used pre- and post- intervention data separately and focused on the total effect on children's body composition of each intervention.

Results of the two methods did not completely agree with each other. For the first method, conclusions are for children's BMI reduction, physical activity (Trt 2) and physical activity plus nutrition (Trt 9) were predicted to be more effective than the other interventions, with physical activity to be even higher ranked than physical activity and nutrition combined; for the effect on children's BMI z-score, it is estimated that physical activity (Trt 2) and nutrition plus education (Trt 7) are likely to be more effective than the other interventions, with physical activity to be slightly lower ranked than nutrition and education combined; for children's body fat percentage reduction, physical activity (Trt 2) is predicted to be the most effective treatment than all the other interventions. In general, physical activity is estimated to be one of the most effective intervention in all of our interested variables. It is reasonable to say that physical activity is the best intervention predicted by the mean difference method. However, for the other method, the SUCRA rank of physical activity intervention was only 8 out of 10. While the best ranked intervention was multiple component treatment (Trt 10), which also makes sense since it engaged more aspects of interventions and is expected to have larger effects than simpler interventions. Personally, I think the estimations both make sense in their own way and could both give instructive guidance regarding children body composition management. The only difference may be results of the first method were provided with more details, so that people with specific aims could have different views and options compared with the second method. In conclusion, if applicable, changing more aspects of children's lifestyle and living environment to improve physical activity and healthy eating may have a great amount of impact on children's body composition management;

if there are difficulty aiming at different aspects at one time, focus on being physically active could be very effective as well.

However, this study bears with some limitations and future studies may benefit from further addressing to problems embedded. Firstly, we did not include the frequency, intensity and duration of any intervention while carrying out the analysis. These could be some important aspects to be added into considerations. However, due to the variety of the interventions, especially with multicomponent interventions, it is often hard to find a standard way to quantify a specific intensity of each of the intervention. Besides, ways to deal with the combination of frequency and duration is also hard to decide. It might be legit to consider the overall intervention time as a whole, while the assigned time for each bout of the intervention is also important to address. Not to mention that some studies did not provided detailed and clear information for secondary categorization for their intervention type or to tell other details like the actual intensity in the intervention protocol. Secondly, the analysis only included per-protocol data and only focused on comparing direct post-test results with baseline data. Some studies provided long-term follow-up data, though very helpful, were not included as part of the results in this case study. Follow-up results are hard to be included because the variety of follow-up time and difficulty in seeking an appropriate cut point to fit all the data. In addition, reliability of the reported results and subject process of secondary data, like categorization, can be another limitation. The interventions in all the studies were all categorized into 10 groups, the criteria could be relatively subjective and less than ideal. For example, in the first three intervention categories, namely, physical activity, education, and nutrition, the nutrition group included interventions with actual food or drink supply, as well as programs only focused on education on knowledge about nutrition and health eating; while for the general health education category, education about nutrition was excluded and it mainly focused on other aspects of healthy living, such as sleep patterns, screen time manage, etc., but included education encouragement on being physically active. Therefore, the physical activity group only contained intervention that actually had children involved in exercise and not only receiving education on exercise related topics. Judgement like this could be a relatively subjective decision and may lead to different ways for intervention categorization with other analysis studies and thus may have different results. Another important part that we are currently missing is the consistency and homogeneity in checking for the network meta-analysis since the current version of pcnetmeta does not have the function to detect inconsistency in arm-based network meta-analysis. Future work could add functions for network consistency assessment or choose other package in R to do the tests. Lastly, since our last intervention category was classified as "multiple component treatment" with more than three simple intervention of any kind, it is hard to determine which component made the most contribution to the final effect, nor could we separate a more effective combination better than the others. Future studies could focus on multiple component treatment only and further categorize then into different treatment groups to clarify their effects.

This case attempted to apply network meta-analysis into the field of physical activity and health promotion. It has shown the feasibility and acceptability for doing so, especially when previous NMA studies only included exercise as one of the intervention arms among other medicine treatments. There are still several aspects future experimental studies could focus on in order to make it more accessible and easier for other researchers to carry out a similar secondary-data analyses. Firstly, it is helpful for researchers to report their adjusted and unadjusted original data (pre-test and post-test), as well as secondary data after simple calculation (mean difference and standard deviation). This could help reduce data pre-processing procedure, although the calculations are only conservative, it is still beneficial to use accurately calculated data instead of using mathematical formula for estimation. Secondly, it is also helpful to set standard or at least comparable quantifications for more commonly used variables, such as moderate-to-vigorous intensity physical activity (MVPA), and light intensity physical activity (LPA). This will help involve more aspects of physical activity field by making their objectively determined data to be even more meaningful at a higher level. Introducing more variables besides BMI and body fat percentage into future analysis would definitely produce many other interesting findings. Lastly, we could observe from the results that the most commonly chose situation of control group is usual care without any intervention. Only a small part of the studies included participants in other status to be control group in those studies. If applicable (sufficient participants), it would be beneficial to intentionally add one or more groups in treatment conditions other than usual care. Not only the single study will benefit a more comprehensive result, but they will contribute more direct comparisons for further network meta-analysis as well. With the rapid growth of NMA studies in the past few years, we have reasons to believe that this is going to be a promising field of study and will show significance in promoting practical analysis methods to more health-related fields like the field of physical activity and health promotion.

Bibliography

Andreato, L. V., Esteves, J. V., Coimbra, D. R., Moraes, A. J. P., & De Carvalho, T. (2019). The influence of high-intensity interval training on anthropometric variables of adults with overweight or obesity: a systematic review and network meta-analysis. *Obesity Reviews*, *20*(1), 142–155. https://doi.org/10.1111/obr.12766

Barlow, S. E. (2007). Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. *Pediatrics*, *120*(4), 164-192. https://doi.org/10.1542/peds.2007-2329C

Boreham, C., & Riddoch, C. (2001). The physical activity, fitness and health of children. *Journal of Sports Sciences*, *19*(12), 915–929. https://doi.org/10.1080/026404101317108426

Brignardello-Petersen, R., Bonner, A., Alexander, P. E., Siemieniuk, R. A., Furukawa, T. A., Rochwerg, B., Hazlewood, G. S., Alhazzani, W., Mustafa, R. A., Murad, M. H., Puhan, M. A., Schünemann, H. J., & Guyatt, G. H. (2018). Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis. *Journal of Clinical Epidemiology*, *93*, 36–44. https://doi.org/10.1016/j.jclinepi.2017.10.005

Brownson, R. C., Boehmer, T. K., & Luke, D. A. (2005). Declining rates of physical activity in the United States: what are the contributors? *Annual Review of Public Health*, *26*(1), 421–443. https://doi.org/10.1146/annurev.publhealth.26.021304.144437

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. https://doi.org/10.1136/bmj.331.7521.897

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. https://doi.org/10.1136/eb-2015-102275

Cowles, M.K. & Carlin, B.P. (2016). Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association, 91*(434), 883-904.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2013). Developing and evaluating complex interventions: The new Medical Research Council guidance. *International Journal of Nursing Studies*, *50*(5), 587–592. https://doi.org/10.1016/j.ijnurstu.2012.09.010

Dias, S., & Ades, A. E. (2016). Absolute or relative effects? Arm-based synthesis of trial data. *Research Synthesis Methods*, *7*(1), 23–28. https://doi.org/10.1002/jrsm.1184

Dias, S., Ades, A.E., Welton, N.J., Jansen, J. & Sutton, A. (2018). *Network Meta-analysis for Decision-making*. John Wiley & Sons Ltd.

Dias, Sofia, & Caldwell, D. M. (2018). Network meta-analysis explained. *Archives of Disease in Childhood. Fetal and Neonatal Editoin. 104*(1), F8–F12.

Gao, Z., Zeng, N., Pope, Z. C., Wang, R., & Yu, F. (2019). Effects of exergaming on motor skill competence, perceived competence, and physical activity in preschool children. *Journal of Sport and Health Science*, *8*(2), 106–113. https://doi.org/10.1016/j.jshs.2018.12.001

George, K., Kristi, K., & Russell, P. (2018). Exercise and BMI z-score in overweight and obese children and adolescents: a systematic revie and network meta-analysis of randomized trials. *Journal of Evidence-based Medicine, 10*(2), 108-128. https://doi.org/10.1111/jebm.12228.

Gilks, W.R., Thomas, A., & Spiegelhater, D.J. (1994). A language and program for complex Bayesian Modelling. *Journal of the Royal Statistical Society, 43,* 169-177.

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. *The Lancet Global Health*, *6*(10), e1077–e1086. https://doi.org/10.1016/S2214-109X(18)30357-7

Hales, C., Carroll, M., Fryar, C., & Ogden, C. (2020). Prevalence of Obesity and Severe Obesity Among Adults: United States, 2017-2018. *NCHS Data Brief*, *360*(360), 1–8. http://www.ncbi.nlm.nih.gov/pubmed/26633046

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*(2), 98–110. https://doi.org/10.1002/jrsm.1044

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

Higgins, J. P.T., Sterne, J. A. C., Savović, J., Page, M. J., Hróbjartsson, A., Boutron, I., Reeves, B. C., & Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. *Cochrane Methods. Cochrane Database of Systematic Reviews 2016*, *10*(52). https://doi.org/10.1002/14651858.CD201601

Hong, H., Chu, H., Zhang, J., & Carlin, B. P. (2016a). A Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons. *Research Synthesis Methods*, *7*(1), 6–22. https://doi.org/10.1002/jrsm.1153

Hong, H., Chu, H., Zhang, J., & Carlin, B. P. (2016b). Rejoinder to the discussion of "a Bayesian missing data framework for generalized multiple outcome mixed treatment comparisons," by S. Dias and A.E. Ades. *Research Synthesis Methods*, *7*(1), 29–33. https://doi.org/10.1002/jrsm.1186

Hong, C., Duan, R., Zeng, L., Hubbard, R.A., Lumley, T., Riley, R., Chu, H., Kimmel, S.E. & Chen, Y. (2020). Galaxy plot: a new visualization tool of bivariate meta-analysis studies. *American Journal of Epidemiology*, *2*(2), 1–20.

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P. A., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catala-Lopez, F., Gotzsche, P. C., Dickersin, K., Boutron, I., Altman, D. G., & 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*(11), 777–784. https://doi.org/10.7326/M14-2385

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

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*(1). https://doi.org/10.1186/1741-7015-11-159

Kohl, H. W., Craig, C. L., Lambert, E. V., Inoue, S., Alkandari, J. R., Leetongin, G., Kahlmeier, S., Andersen, L. B., Bauman, A. E., Blair, S. N., Brownson, R. C., Bull, F. C., Ekelund, U., Goenka, S., Guthold, R., Hallal, P. C., Haskell, W. L., Heath, G. W., Katzmarzyk, P. T., … Wells, J. C. (2012). The pandemic of physical inactivity: Global action for public health. *The Lancet*, *380*(9838), 294–305. https://doi.org/10.1016/S0140-6736(12)60898-8

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*(79). https://doi.org/10.1186/1741-7015-9-79

Lian, Q., Hodges, J. S., & Chu, H. (2019). A Bayesian Hierarchical Summary Receiver Operating Characteristic Model for Network Meta-Analysis of Diagnostic Tests. *Journal of the American Statistical Association*, *114*(527), 949–961. https://doi.org/10.1080/01621459.2018.1476239

Lin, L., Chu, H., & Hodges, J. S. (2016). Sensitivity to excluding treatments in network meta-analysis. *Epidemiology*, *27*(4), 562–569. https://doi.org/10.1097/EDE.0000000000000482

Lin, L., Zhang, J., Hodges, J. S., & Chu, H. (2017). Performing arm-based network meta-analysis in R with the pcnetmeta package. *Journal of Statistical Software*, *80*(5). https://doi.org/10.18637/jss.v080.i05

Lin, L., Shi, L., Chu, H., Murad, M.H. (2020). The magnitude of small-study effects in the Cochrane Database of Systematic Reviews: an empirical study of nearly 30 000 meta-analyses. *BMJ Evidence-Based Medicine, 25*(1), 27-32.

Lox, C. L., Martin Ginis, K. A., & Petruzzello, S. J. (2014). *The Psychology of Exercise. Routledge Fourth Edition.*

Lu, G., & Ades, A. E. (2004). Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine*, *23*(20), 3105–3124. https://doi.org/10.1002/sim.1875

Lu, G., & Ades, A. E. (2006). Assessing evidence inconsistency in mixed treatment comparisons. *Journal of the American Statistical Association*, *101*(474), 447–459. https://doi.org/10.1198/016214505000001302

Lumley, T. (2002). Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine*, *21*(16), 2313–2324. https://doi.org/10.1002/sim.1201

Ma, X., Lian, Q., Chu, H., Ibrahim, J. G., & Chen, Y. (2018). A Bayesian hierarchical model for network meta-analysis of multiple diagnostic tests. *Biostatistics*, *19*(1), 87–102. https://doi.org/10.1093/biostatistics/kxx025

Martín-García, M., Alegre, L. M., García-Cuartero, B., Bryant, E. J., Gutin, B., & Ara, I. (2019). Effects of a 3-month vigorous physical activity intervention on eating behaviors and body composition in overweight and obese boys and girls. *Journal of Sport and Health Science*, *8*(2), 170–176. https://doi.org/10.1016/j.jshs.2017.09.012

Matthews, C. E., Keadle, S. K., Troiano, R. P., Kahle, L., Koster, A., Brychta, R., Van Domelen, D., Caserotti, P., Chen, K. Y., Harris, T. B., & Berrigan, D. (2016). Accelerometer-measured dose-response for physical activity, sedentary time, and mortality in US adults. *American Journal of Clinical Nutrition*, *104*(5), 1424–1432. https://doi.org/10.3945/ajcn.116.135129

Molloy, G. J., Noone, C., Caldwell, D., Welton, N. J., & Newell, J. (2018). Network meta-analysis in health psychology and behavioural medicine: a primer. *Health Psychology Review*, *12*(3), 254–270. https://doi.org/10.1080/17437199.2018.1457449

Naci, H., & John, P. A. (2013). Comparative effectiveness of exercise and drug interventions on mortality outcomes: Metaepidemiological study. *BMJ (Online)*, *347*(7929), 1–14. https://doi.org/10.1136/bmj.f5577

Naci, H., Salcher-Konrad, M., Dias, S., Blum, M. R., Sahoo, S. A., Nunan, D., & Ioannidis, J. P. A. (2019). How does exercise treatment compare with antihypertensive medications? A network meta-analysis of 391 randomised controlled trials assessing exercise and medication effects on systolic blood pressure. *British Journal of Sports Medicine*, *53*(14), 859–869. https://doi.org/10.1136/bjsports-2018-099921

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. https://doi.org/10.1371/journal.pone.0115065

Ng, S. W., & Popkin, B. M. (2012). Time use and physical activity: A shift away from movement across the globe. *Obesity Reviews*, *13*(8), 659–680. https://doi.org/10.1111/j.1467-789X.2011.00982.x

Owen, R. K., Bradbury, N., Xin, Y., Cooper, N., & Sutton, A. (2019). MetaInsight: An interactive web-based tool for analyzing, interrogating, and visualizing network meta-analyses using R-shiny and netmeta. *Research Synthesis Methods*, *10*(4), 569–581. https://doi.org/10.1002/jrsm.1373

Palmer, K. K., Chinn, K. M., & Robinson, L. E. (2019). The effect of the CHAMP intervention on fundamental motor skills and outdoor physical activity in preschoolers. *Journal of Sport and Health Science*, *8*(2), 98–105. https://doi.org/10.1016/j.jshs.2018.12.003

Pan, B., Ge, L., Xun, Y. qin, Chen, Y. jing, Gao, C. yun, Han, X., Zuo, L. qian, Shan, H. qian, Yang, K. hu, Ding, G. wu, & Tian, J. hui. (2018). Exercise training modalities in patients with type 2 diabetes mellitus: A systematic review and network meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 1–14. https://doi.org/10.1186/s12966-018-0703-3

Papakonstantinou, T., Nikolakopoulou, A., Higgins, J. P. T., Egger, M., & Salanti, G. (2020). CINeMA: Software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell Systematic Reviews*, *16*(1). https://doi.org/10.1002/cl2.1080

Puhan, M. A., Schünemann, H. J., Murad, M. H., Li, T., Brignardello-Petersen, R., Singh, J. A., Kessels, A. G., & Guyatt, G. H. (2014). A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ (Online)*, *349*, 1–10. https://doi.org/10.1136/bmj.g5630

Physical activity guidelines for Americans, 2nd Edition (2018). *U.S. Department of Health and Human Services.*

Rochwerg, B., Neupane, B., Zhang, Y., Garcia, C. C., Raghu, G., Richeldi, L., Brozek, J., Beyene, J., & Schünemann, H. (2016). Treatment of idiopathic pulmonary fibrosis: A network meta-analysis. *BMC Medicine*, *14*(1). https://doi.org/10.1186/s12916-016-0558-x

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. https://doi.org/10.1002/jrsm.1037

Schwingshackl, L., Dias, S., Strasser, B., & Hoffmann, G. (2013). Impact of different training modalities on anthropometric and metabolic characteristics in overweight/obese subjects: A systematic review and network meta-analysis. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0082853

Schwingshackl, L., Missbach, B., Dias, S., König, J., & Hoffmann, G. (2014). Impact of different training modalities on glycaemic control and blood lipids in patients with type 2 diabetes: A systematic review and network meta-analysis. *Diabetologia*, *57*(9), 1789–1797. https://doi.org/10.1007/s00125-014-3303-z

Sekercioglu, N., Veroniki, A. A., Thabane, L., Busse, J. W., Akhtar-Danesh, N., Iorio, A., Lopes, L. C., & Guyatt, G. H. (2017). Effects of different phosphate lowering strategies in patients with CKD on laboratory outcomes: A systematic review and NMA. *PLoS ONE*, *12*(3), 1–26. https://doi.org/10.1371/journal.pone.0171028

Shim, S. R., & Lee, J. (2019). Dose-response meta-analysis: application and practice using the R software. *Epidemiology and Health*, *41*, e2019006. https://doi.org/10.4178/epih.e2019006

Thomas, A., & Spiegelhalter, D. J. (1925). Journal of the Royal Statistical Society. *The Economic Journal*, *35*(140), 661–668. https://doi.org/10.1093/ej/35.140.661

Tonin, F. S., Rotta, I., Mendes, A. M., & Pontarolo, R. (2017). Network meta-analysis: A technique to gather evidence from direct and indirect comparisons. *Pharmacy Practice*, *15*(1), 1–11. https://doi.org/10.18549/PharmPract.2017.01.943

Umer, A., Kelley, G. A., Cottrell, L. E., Giacobbi, P., Innes, K. E., & Lilly, C. L. (2017). Childhood obesity and adult cardiovascular disease risk factors: A systematic review with meta-analysis. *BMC Public Health*, *17*(1), 1–24. https://doi.org/10.1186/s12889-017-4691-z

Uthman, O. A., Van Der Windt, D. A., Jordan, J. L., Dziedzic, K. S., Healey, E. L., Peat, G. M., & Foster, N. E. (2013). Exercise for lower limb osteoarthritis: Systematic review incorporating trial sequential analysis and network meta-analysis. *BMJ (Online)*, *347*(7928), 1–13. https://doi.org/10.1136/bmj.f5555

Wang, Z., Lin, L., Hodges, J. S., & Chu, H. (2020). The impact of covariance priors on arm-based Bayesian network meta-analyses with binary outcomes. *Statistics in Medicine*. https://doi.org/10.1002/sim.8580

Warburton, Darren E.R.; Nicol,Crystal Whitney; Bredin, S. S. D. (2006). Health benefits of physical activity: the evidence. *Family Medicine and Primary Care Review*, *8*(3), 1110–1115.

White, I. R., Barrett, J. K., Jackson, D., & Higgins, J. P. T. (2012). Consistency and inconsistency in network meta-analysis: model estimation using multivariate metaregression. *Research Synthesis Methods*, *3*(2), 111–125. https://doi.org/10.1002/jrsm.1045

White, I. R., Turner, R. M., Karahalios, A., & Salanti, G. (2019). A comparison of arm-based and contrast-based models for network meta-analysis. *Statistics in Medicine*, *38*(27), 5197–5213. https://doi.org/10.1002/sim.8360

Xia, T. li, Huang, F. yang, Peng, Y., Huang, B. tao, Pu, X. bo, Yang, Y., Chai, H., & Chen, M. (2018). Efficacy of Different Types of Exercise-Based Cardiac Rehabilitation on Coronary Heart Disease: a Network Meta-analysis. *Journal of General Internal Medicine*, *33*(12), 2201–2209. https://doi.org/10.1007/s11606-018-4636-y

Yamaoka, K., Nemoto, A., & Tango, T. (2019). Comparison of the effectiveness of lifestyle modification with other treatments on the incidence of type 2 diabetes in people at high risk: A network meta-analysis. *Nutrients*, *11*(6). https://doi.org/10.3390/nu11061373

Zhang, J., Carlin, B. P., Neaton, J. D., Soon, G. G., Nie, L., Kane, R., Virnig, B. A., & Chu, H. (2014). Network meta-analysis of randomized clinical trials: Reporting the proper summaries. *Clinical Trials*, *11*(2), 246–262. https://doi.org/10.1177/1740774513498322

Zou, T. T., Zhang, C., Zhou, Y. F., Han, Y. J., Xiong, J. J., Wu, X. X., Chen, Y. P., & Zheng, M. H. (2018). Lifestyle interventions for patients with nonalcoholic fatty liver disease: A network meta-analysis. *European Journal of Gastroenterology and Hepatology*, *30*(7), 747–755. https://doi.org/10.1097/MEG.0000000000001135

Appendices

Table 9. Estimates of means and SDs for the log odds ratio of each treatment children's BMI.

	Trt1	Trt2	Trt3	Trt4	Trt5	Trt6	Trt7	Trt8	Trt9	Trt10
Trt1	--									
Trt2	-0.205 (0.095)	--								
Trt3	0.036 (0.143)	0.241 (0.170)	--							
Trt4	-0.094 (0.104)	0.112 (0.141)	-0.129 (0.176)	--						
Trt5	0.380 (0.339)	0.586 (0.352)	0.345 (0.367)	0.474 (0.354)	--					
Trt6	0.034 (0.158)	0.239 (0.184)	-0.002 (0.208)	0.127 (0.190)	-0.347 (0.372)	--				
Trt7	-0.109 (0.118)	0.096 (0.150)	-0.144 (0.171)	-0.016 (0.158)	-0.489 (0.359)	-0.143 (0.198)	--			
Trt8	-0.145 (0.362)	0.060 (0.376)	-0.180 (0.392)	-0.052 (0.375)	-0.525 (0.497)	-0.178 (0.398)	-0.036 (0.384)	--		
Trt9	-0.120 (0.132)	0.085 (0.162)	-0.156 (0.184)	-0.027 (0.162)	-0.501 (0.363)	-0.154 (0.205)	-0.011 (0.178)	0.024 (0.386)	--	
Trt10	-0.137 (0.080)	0.068 (0.122)	-0.172 (0.153)	-0.043 (0.131)	-0.517 (0.348)	-0.170 (0.176)	-0.028 (0.139)	0.008 (0.373)	-0.016 (0.154)	--

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 10. Estimates of means and SDs for the log odds ratio of each treatment for children's BMI z-score.

	Trt1	Trt2	Trt3	Trt5	Trt6	Trt7	Trt8	Trt10
Trt1	--							
Trt2	-0.081 (0.089)	--						
Trt3	-0.078 (0.044)	0.004 (0.096)	--					
Trt5	0.066 (0.113)	0.147 (0.140)	0.143 (0.118)	--				
Trt6	-0.010 (0.112)	0.072 (0.139)	0.068 (0.116)	-0.075 (0.156)	--			
Trt7	-0.105 (0.076)	-0.024 (0.112)	-0.027 (0.078)	-0.170 (0.132)	-0.095 (0.131)	--		
Trt8	-0.066 (0.060)	0.016 (0.098)	0.012 (0.068)	-0.131 (0.121)	-0.056 (0.121)	0.039 (0.089)	--	
Trt10	-0.083 (0.054)	-0.002 (0.098)	-0.006 (0.063)	-0.149 (0.121)	-0.074 (0.120)	0.021 (0.087)	-0.018 (0.069)	--

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).

Table 11. Estimates of means and SDs for the log odds ratio of each treatment for children's body fat percentage.

	Trt1	Trt2	Trt3	Trt4	Trt6	Trt7	Trt9	Trt10
Trt1	--							
Trt2	-1.320 (0.501)	--						
Trt3	0.517 (0.709)	1.840 (0.768)	--					
Trt4	-0.491 (0.857)	0.834 (0.912)	-1.010 (1.030)	--				
Trt6	-0.731 (0.904)	0.593 (0.962)	-1.250 (1.090)	-0.240 (1.180)	--			
Trt7	-0.293 (0.906)	1.030 (0.967)	-0.810 (1.090)	0.198 (1.190)	0.439 (1.220)	--		
Trt9	0.089 (0.532)	1.410 (0.612)	-0.428 (0.769)	0.580 (0.895)	0.820 (0.987)	0.382 (0.997)	--	
Trt10	-0.549 (0.514)	0.776 (0.623)	-1.070 (0.799)	-0.058 (0.933)	0.183 (0.975)	-0.256 (0.976)	-0.638 (0.659)	--

Note: Intervention categories are listed below: 1) control (T1; no intervention); 2) physical activity (PA)/exercise only (T2); 3) knowledge education in various dimensions (T3); 4) nutrition (T4); 5) environment changes (T5); 6) PA plus education (T6); 7) nutrition plus education (T7); 8) environment changes plus education (T8); 9) PA with nutrition (T9); and 10) multiple component treatment (more than three single treatment combined) (T10).