Advances in Integrative Medicine 3 (2016) 114–120

Contents lists available at ScienceDirect

Advances in Integrative Medicine

journal homepage: www.elsevier.com/locate/aimed

Methods Series

Meta-analysis in systematic reviews of complementary and integrative medicine trials

Holger Cramer[a,b,*], Vincent C.H. Chung[a,c]

[a] Australian Research Centre in Complementary and Integrative Medicine, University Technology Sydney, Australia
[b] Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Essen, Germany
[c] Cochrane Hong Kong, Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, China

ARTICLE INFO

Article history:
Received 8 February 2017
Accepted 21 February 2017
Available online 22 February 2017

Keywords:
Meta-analysis
Methodology
How to publish
Publication
Clinical research

ABSTRACT

In evidence-based medicine, meta-analyses are often considered to provide an exceptionally high level of evidence. Beyond other systematic review formats, meta-analyses allow for statistically pooling results of original research. This approach allows us to increase statistical power, improve precision, investigate heterogeneity, and settle controversies across different clinical trials. However, meta-analysis requires rigour in methodology, in order to avoid "mixing apples and oranges". Study design, participants, interventions, comparators, and outcomes need to be clearly defined, and statistical methods need to be applied precisely. For researchers in specialized fields such as integrative medicine, additional considerations on the unique features of therapies need to be applied. This article describes practical and academic insights into preparing a meta-analysis for publication.

1. Introduction

Evidence-based medicine is the science and art of making patient care decisions which are consistent with current best evidence. Evidence-based decisions should also consider available resources and values of the individual patient in addition to evidence [1]. The evidence-based medicine movement has influenced policy, practice research direction in complementary and integrative medicine (CIM) [2]. The recent WHO Traditional Medicine Strategy also highlighted the importance of strengthening CIM evidence base in promoting the integration of CIM in health systems [3]. Randomized controlled trials (RCTs) is considered the best primary study design for investigating efficacy or effectiveness of interventions, and systematic reviews of RCTs provides synthesis of all evidence generated from RCTs in a replicable manner. When the clinical and methodological features of RCTs are similar enough, meta-analysis can be used to pool effect sizes reported from each trial. Meta-analysis has the advantages of allowing higher statistical power for detecting treatment effects, improving precision in estimating effect size, settling controversy arising from conflicting results from different trials, and facilitating exploring of heterogeneity across trials [4]. Meta-analysis is increasingly being applied in CIM clinical research [5–7], and this article aims to provide an overview on the methodologies of meta-analysis.

1.1. Advantages and disadvantages of adding meta-analyses to systematic reviews

A drawback of traditional systematic reviews, i.e. reviews which are conducted systematically following an a priori defined methodology but not statistically pooling the results of the original research, is that they have to solely rely on the findings of the original studies. While this is by no means only a problem of CIM research, the study quality in CIM randomized trials is often limited by small sample sizes, unclear reporting of methodology, and inadequate rigor in statistical analysis. With continual call for methodological improvements and endorsement of reporting guidelines by the academia, standards of future trials are expected to improve but many existing trials have shortcoming in these areas. With these limitations, both significant and insignificant results are often hard to interpret based on single trials, due to two major reasons: First, due to limited funding, trials on CIM interventions are often small and underpowered for detecting any specific effects of the intervention; resulting in an underestimation of true treatment effect. Second, randomized trials (regardless if on CIM or other interventions) are generally

* Corresponding author at: Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Essen, Germany.
E-mail address: h.cramer@kliniken-essen-mitte.de (H. Cramer).

http://dx.doi.org/10.1016/j.aimed.2017.02.004
2212-9588/

conducted to compare two or more different groups (e.g. CIM versus no treatment, CIM versus guideline recommended conventional treatment, or CIM versus placebo/sham interventions). With such design, the impact of CIM should be estimated using between-group comparisons, but unfortunately many trials only report results of within-group comparisons that do not take into account unspecific effects. With such analysis, the value of having a control group is lost and it is impossible to estimate the real specific effects of the intervention; this will lead to overestimation of actual specific efficacy.

In systematic reviews without meta-analyses, the reviewers have to handle these uncertainties on whether the reported findings and – even more subjective – the conclusions of original research can really be trusted. This subjective process can result in quite different interpretations on reviews dealing with the same pool of original studies. If correctly conducted, a meta-analysis can deal with some of the limitations of original trials by calculating pooled statistics: by pooling and weighing the findings from multiple trials, sample sizes of the comparisons can be dramatically increased and even small trials can contribute to the overall picture. Also, if usable data were available from the original trial publications, meta-analysts do not bother on whether the trialists had applied appropriate statistical tests or not – they simply re-calculate pooled effect estimates using meta-analytic techniques. This way, systematic reviews with and without meta-analyses can reach totally different conclusions even if they include the same trials, especially when only few small trials are available for analysis [8].

While there are benefits in conducting meta-analyses, their appropriateness depend on the availability of at least two clinically and methodologically homogeneous trials [9]. The studies need to be similar in the aspects of patient characteristics, intervention, comparators, outcome measures, and trial designs [10]. Early CIM meta-analyses have simply pooled all available trials, e.g. all randomised trials on homeopathy, regardless of the specific remedy used, the condition to be treated or which outcomes were assessed [11,12]. Variations in comparators are also not uncommon, for instance, different conventional treatments are often lumped inappropriately as “western treatment” in many meta-analyses of Chinese herbal medicine trials [13].

But what does it mean if we find a significant or non-significant meta-analytic result after pooling such a heterogeneous group of trials? It can mean all or nothing and it will definitely not give you any guidance whether a specific intervention might work for a specific treatment goal in clinical practice. We now know that this approach of “mixing apples and oranges” does not make sense when we plan to reach a single indicator of efficacy or effectiveness as in most meta-analyses [10] (see example in Box 1). The appropriate approach of handling variations in interventions is to apply network meta-analysis technique [14], but this require sharing of a common comparator (e.g. placebo) and outcome among at least some of the trials (see example in Box 2).

When clinical judgement suggest that pairwise or network meta-analysis is inappropriate, a systematic review without meta-analysis can include all those heterogeneous trials and discuss the implications of the heterogeneity. Such discussion can drive future research directions, and may guide treatment decision when there are few other options left for certain patients (see example in Box 3). Thus, as always in research, the choice of a specific methodology strongly depends on the research question we intend to answer.

2. Locating and reviewing eligible trials

2.1. Defining the research question

As outlined above, the clinical relevance and usefulness of a meta-analysis stands and falls with its research question – much stronger than for other forms of reviews. For instance, a topic like “Is CIM effective in women’s health?” can be a great question for a non-meta-analytic systematic review, perhaps even more for a critical review, it is however not suitable for a meta-analysis. In order to allow statistical pooling of results from different trials, they need to be adequately homogeneous (remember the apples and oranges problem), otherwise results of the meta-analysis will not be interpretable at all. In order to reduce such clinical heterogeneity, an a priori defined research question is necessary and the PICOS (participants, intervention, comparison, outcomes, study design) approach for formulating clinically answerable question can be extremely useful for clarifying scope of meta-analysis [24].

2.1.1. Participants

Who will be the target of the intervention under review? This includes the basic question on whether healthy participants at risk (e.g. trials to evaluate the effect of manual acupuncture for smoking cessation in the general population) or those with a specific condition (e.g. trials to assess the effect of moxibustion for fatigue among patients with lung cancer after completion of chemotherapy) will be included. In a meta-analysis on cancer, will you include non-metastatic or metastatic cancer or both? Will both genders be eligible? Will there be a defined age range or will there be ethnic or geographic criteria to consider?

2.1.2. Intervention

What intervention, or intervention package, are to be studied? As we outlined above, something like ‘homeopathy’ or ‘Ayurveda’ might be too broad. If enough studies are available, it might be more useful to focus on a specific remedy or formula. But it goes beyond that: if the intervention has any subtypes, which of them

Box 1. “Mixing apples and oranges” in meta-analysis of different herbal medicines

In Europe and North American, herbal medicine is used for Benign Prostate Hyperplasia (BPH) [15]. Sources of phytotherapeutic agents used for treatment of BPH include saw palmetto (*Serenoa repens*) fruit [16,17], African plum tree (*Pygeum africanum*) bark [18], and stinging nettle (*Urtica dioica*) roots in the form of β-sitosterol, South African stargrass (Hypoxis roopieri), Pumpkin seed (*Cucurbita pepo*), and rye pollen (*Secale cereale*) [19].
According to recommendations made by the American Urological Association (AUA) Practice Guideline Committee, change in the AUA Symptom Index (AUASI) should be regarded as the primary outcome measure for evaluating outcomes of treatment for BPH [20]. If a systematic review of these herbal treatments were conducted using AUASI as a common outcome, meta-analysis should only be performed among trials comparing the same herb on a common control (e.g. a placebo). The “mixing apple and oranges” approach of lumping different herbs together will not allow identification of phytotherapeutic agents that contribute to the pooled effect.

Box 2. Evaluating comparative effectiveness of different Chinese herbal medicines using network meta-analysis approach

In China, different Chinese herbal medicines (CHMs) are often used in addition to conventional medications in managing various chronic conditions [21], including chronic obstructive pulmonary disease (COPD). Clinicians are often interested in comparative effectiveness of different CHMs, but pairwise meta-analysis could not provide such evidence. Network meta-analysis (NMA) allows simultaneous evaluation of relative performance among different CHMs, when there is a common comparator. An NMA with comparisons of CHMs plus salmeterol and fluticasone propionate (SFP), versus SFP alone has recently been published.
In this NMA, eleven RCTs (n = 925) assessing 11 different CHMs were included. Results suggested that Runfeijianpibushen decoction and Renshenbufei pills performed best in improving quality of life among COPD patients, as measured by the St George's Respiratory Questionnaire, while the sole use of SFP has the lowest probability of delivering such outcome [22].

Box 3. Narrative approach in synthesis in systematic review

In a systematic review on herbal medicines for idiopathic Parkinson's disease, nine trials were included but each of them investigated a different herbal product. Since the controls used were also variable, network meta-analysis is not appropriate and a narrative approach was used to summarize current evidence on effectiveness and safety. The results addressed methodological limitations of existing trials, and offered practical advice on the designs of inclusion criteria, disease staging, and outcome assessments in future trials [23].

will be included (e.g. choices among manual acupuncture, electroacupuncture, laser acupuncture etc.). If broad categories or even whole medical systems (e.g. a wide array of modalities in Chinese medicine practice) are to be studied, we should consider using a more narrative approach or at least plan for subgroup analyses, which focus on specific remedies or interventions. If a common comparator exist for specific intervention, network meta-analysis maybe applicable if the trial designs were similar enough [25].

2.1.3. Comparison

A specific CIM intervention might be compared to no treatment, or to placebo, when there is no recommended conventional treatment. A CIM intervention can also be compared against a conventional/CIM intervention with limited effectiveness or significant side effects, as be evaluated as an add-on to conventional care. Given the wide range of possible control, it would be inappropriate to include trials comparing an intervention to different control interventions in a single pairwise meta-analysis. Here, we have to decide on which comparator is of interest for our research question (e.g. by comparing St John's Wort to placebo, or to antidepressant drugs). If we would like to cover the whole picture (e.g. if we want to know whether St John's Wort is more effective than placebo, and whether it is equally effective as antidepressant drugs) we need to perform separate meta-analyses for both research questions, or to conduct network meta-analysis.

In many trials, CIM are often used in conjunction with conventional care due to clinical or ethical reasons. The systematic reviewer should clarify the specific components of each CIM and conventional interventions to be included. For example, in smoking cessation trials, acupuncture plus nicotine replacement patch could be compared against sham acupuncture plus nicotine replacement patch, or nicotine replacement patch alone. In other trials electroacupuncture maybe used. The systematic reviewer should specify which type of comparison they would be the focus, and provide separate analysis for each specific comparison.

2.1.4. Outcomes

What does effectiveness mean? If we study the effects of an intervention in cancer patients, do we want to know whether it actually cures cancer or whether it helps dealing with side effects of the conventional treatment? Again, we can include multiple outcomes but we should compute separate meta-analyses for each outcome. Gold standard outcomes are often not used in CIM trials, of which authors tends to use self-developed outcome measures with doubtful reliability, validity and responsiveness [26,27]. In systematic review protocols, investigators should pre-specific core outcomes to be covered for a certain condition, and include trials which report patient centred outcome recommended by relevant expert panels or authorities. It is important to state clearly how the pooled effect size should be interpreted, in accordance to established minimal clinically important difference values [28].

2.1.5. Study types

Which type of study is of interest for our research question? Generally most meta-analyses are limited to randomized trials but sometimes non-randomized controlled trials or even uncontrolled trials are also included. Including those trials increases methodological heterogeneity, and it may also limit validity of findings. Nevertheless, inclusion of studies with less than ideal design might be necessary if too few randomized trials were available. When non-randomized studies are included, appraisals of their risk of bias should be conducted using appropriate instruments, with particular attention to selection bias. If meta-analysis were performed using data from non-randomized studies, reviewers should consider the impact of residual confounding and other bias on the pooled results [29].

As we see, inclusion criteria can be broad or narrow, and the broader they are the more we risk to increase clinical and methodological heterogeneity, while the narrower they are, the more we risk to end up with no or too few usable trials. This can generally be handled by defining broader inclusion criteria and computing several meta-analyses for more distinct subgroups or adopt network meta-analysis, but this will increase complexity of your systematic review.

2.2. Locating and appraising the literature

Most meta-analyses start with a systematic review of the literature, basically following the same approach as a traditional systematic review: based on the research question, inclusion and exclusion criteria are defined (these often are more narrow than for other types of reviews), based on these criteria a literature search is performed, and the papers are filtered and critically appraised

Table 1
Risk of bias assessment.

Bias	Criteria for low risk of bias
Random sequence generation	Use of a random sequence generation (random number table, throwing dice, coin tossing).
Allocation concealment	Participants and investigators cannot foresee assignment (central randomization, sequentially numbered drug containers or envelopes).
Blinding of participants and personnel	Blinding of participants and key study personnel OR no blinding but the outcome is not likely to be influenced by lack of blinding.
Blinding of outcome assessment	Blinding of outcome assessment OR no blinding but the outcome is not likely to be influenced by lack of blinding.
Incomplete outcome data	Drop-outs and reasons for drop-outs are balanced across groups and/or missing data are adequately imputed.
Selective reporting	All pre-specified outcomes that are of interest in the review have been reported in the pre-specified way
Other bias	The study appears to be free of other sources of bias.

using standardised assessments tools (see [10] for a more detailed description of systematic review methodology). Critical appraisal of the literature is especially important as it facilitate sensitivity analysis within meta-analyses. Sensitivity analysis allows the comparison of pooled results calculated from only trials with low risk of bias, versus all trials included in a meta-analysis [30]. This will allow assessment on how risk of bias may impact effect size [31].

Most meta-analyses now use the Cochrane criteria for assessing risk of bias in randomised trials.[1] The Cochrane tool assesses risk of bias on the following domains: selection bias (random sequence generation, allocation concealment), performance bias (blinding of participants and personnel), detection bias (blinding of outcome assessment), attrition bias (incomplete outcome data), reporting bias (selective reporting), and other bias as: (1) low risk of bias, (2) unclear, or (3) high (Table 1) [32].

3. Computing the meta-analysis

3.1. Software

There are numerous meta-analyses softwares available but most analyses use the Review Manager (RevMan) software provided by the Cochrane Collaboration (http://tech.cochrane.org/revman/download) because it is easy to use, freely available and it is also used in Cochrane Reviews which are often considered to be of exceptional high quality. RevMan has built-in routines which will be enough for most basic meta-analyse while other software include more elaborated functions which might be needed when addressing technical requirement. For instance, the free Meta-Analyst software has a user friendly, excel based interface that allows quick data input. Also, it has dedicated functions for meta-regression, cumulative meta-analysis, and leave one out sensitivity analysis, and enhanced graphic editing modules [33]. The commercially available Comprehensive Meta-Analysis software has strength in calculating effect sizes using an efficient bulit-in program, reducing the efforts for reviewers to convert differently reported statistical results into a common metric [https://www.meta-analysis.com/]. Finally, STATA is an efficient generic statistical software for performing network meta-analysis [34]

3.2. Analyzing the data

The most basic meta-analyses compare outcomes between two different interventions. Therefore, for continuous outcomes measures on post-intervention central tendency (normally post intervention means or change scores, whatever is available) and dispersion (normally standard deviation, but standard errors, confidence intervals, or even t-values will also work) and group sample sizes are needed. Based on these data, the difference between groups and its confidence interval are calculated. There are two general types of group differences: mean differences (MD) and standardised mean differences (SMD). MDs are calculated if all studies use the same scale for assessing outcomes while SMDs standardise the difference in means by the pooled standard deviation and can thus be used when different scales are used. For example, blood pressure is almost always measured as mmHg, thus in a meta-analysis differences between groups can be expressed as MD which makes them much easier to interpret. However, for the outcome of depression it may be measured by different instruments, and calculation of MD is not appropriate. The pooled outcomes can instead be expressed with SMD, and its interpretation can follow Cohen's conventions for interpreting SMDs:1) SMD = 0.2–0.5: small effect; 2) SMD = 0.5–0.8: moderate effect and 3) SMD >0.8: large effect [35].

On the other hand, when using dichotomous outcomes we want to know whether there are more patients in one treatment group that experience a specific outcome (adequate relief, adverse events or even death) than in the other group. Dichotomous outcomes answer a slightly different question than continuous ones do as they put the number of patients (rather than the mean of a patient-related variable) in focus. Here, meta-analyses are mainly computed for ratios, i.e. we want to know how many patients in the treatment are likely to experience a specific event *in relation* to patients in the control group. The most straight-forward dichotomous metric is the risk ratio. The risk of a specific event in a given group is computed as the number of patients with this event divided by the total number of patients in this group. I.e. if 2 out of 10 patients using a specific intervention receive adequate relief, the incidence of relief is 2/10 or 0.2; in the control group, only 1 out of 10 patients might experience relief, here the incidence is 0.1. The risk ratio is computed by simply dividing one incidence by the other: 0.2/0.1 = 2. This simply means that the incidence of relief in the treatment group is twice as high as in the control group. There are ratios such as odds ratio that are less easy to interpret but can have other advantages when used in a meta-analysis [32].

Beyond calculating effect sizes or ratios for individual studies, meta-analyses calculate pooled effect sizes across studies; the influence of each study is weighed based on its sample size and/or precision. This way we end up with a single indicator of an intervention's effectiveness across all available studies. Most meta-analysis softwares mentioned above are capable in creating a forest plot that numerically and graphically displays all relevant

[1] Risk of bias is a similar concept as study quality but they are not to be used interchangeably. While study quality evaluated whether a study meets a specific criterion (e.g. blinding of patients), risk of bias assesses whether meeting or not meeting this criterion is likely to bias the results of the study (e.g. even an unblinded trial can have a low risk of bias if two credible and well-accepted interventions are compared and both interventions are judged equally by the study participants).

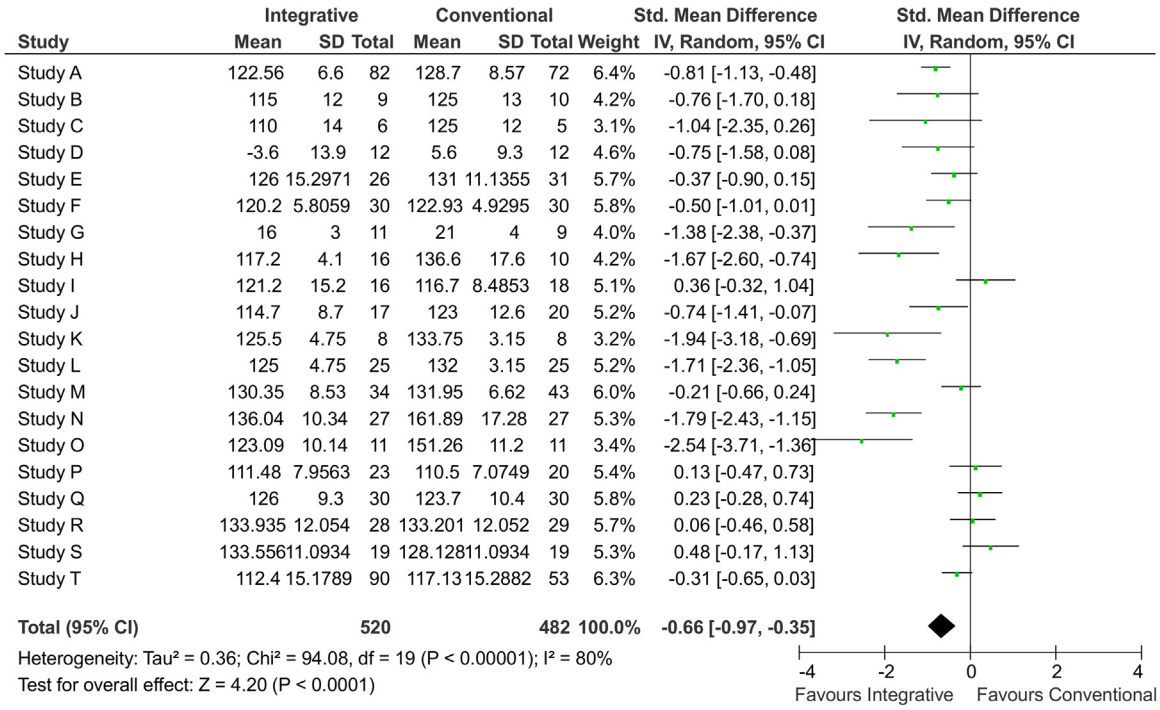

Fig 1. Example of a forest plot. The integrative intervention was more effective than the conventional one, overall effect was of moderate size, and there was considerable heterogeneity.

measures of individual study effects, overall effect, and statistical heterogeneity (Fig. 1).

3.3. Assessing statistical heterogeneity

As outlined above, clinical heterogeneity (differences between included trials in terms of PICOS) and methodological heterogeneity (difference between included trials in terms of study design and risk of bias) can increase variability of effect size across trials and thus reduce the interpretability of results. There is uncertainty on whether the positive pooled effects really are driven by all included studies, or just by a few of them with certain PICOS features or risk of bias. This will affect decision making among clinicians – for instance, to what patient population the findings can be applied if the pooled effect were generated partly from trials which recruited low risk population, as well as trials with high risk patients? Likewise, the lack of overall effects might still mean that there were positive effects in some of the included trials with certain features. While perfect homogeneity might possibly be expected in highly standardised placebo-controlled pharmacological trials, for CIM trials – especially for trials dealing with behavioural, mind-body medical or body-based interventions

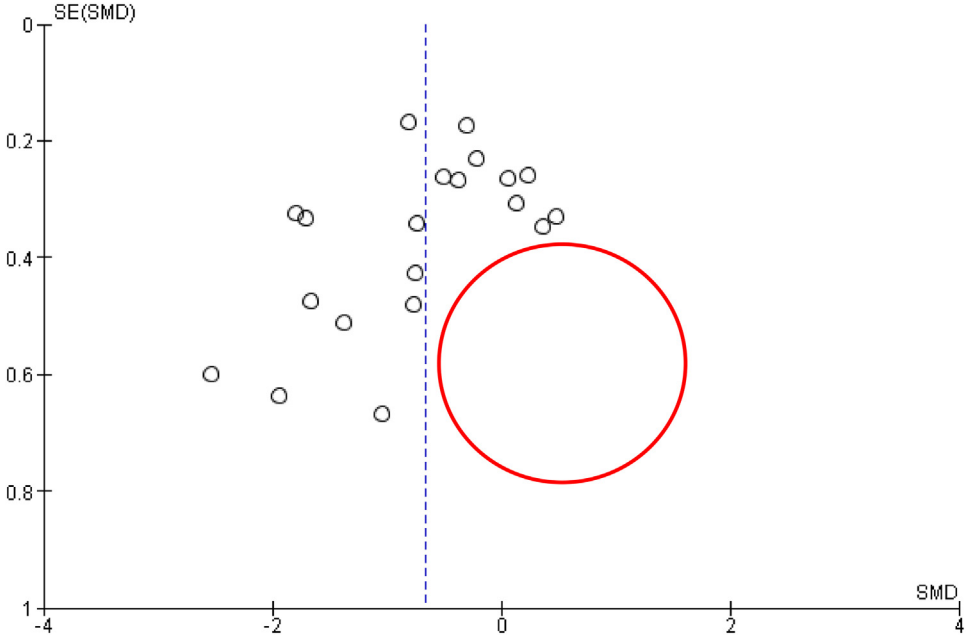

Fig. 2. Example of a funnel plot using the same data as in Fig. 1. Funnel plot asymmetry suggests potential publication bias.

where treatment packages can be highly variable – heterogeneity can normally be expected even in rigorous meta-analyses.

To test how much outcome variance between studies can really be attributed to PICOS differences between studies as well as bias, rather than chance, the I^2 statistics can be used (Fig. 1). The I^2 value is expressed as % and is often interpreted as: 1) I^2 = 0–25%: low heterogeneity; 2) I^2 = 26–50%: moderate heterogeneity; 3) I^2 = 51–75%: substantial heterogeneity; and 4) I^2 = 76–100%: considerable heterogeneity [36]. This existence of considerable heterogeneity does not render the findings of a meta-analysis worthless, but reviewers should explore sources of heterogeneity in subgroup and sensitivity analyses.

3.4. Subgroup and sensitivity analyses

Beyond calculating the overall effect size, it is often useful to include subgroup analyses, particularly when there are variations in patient and interventional features across trials. This way, differences in effects between specific patient groups (e.g. women vs. men; metastatic vs. non-metastatic cancer), interventional features (e.g. pills vs. liquids; aerobic exercise vs. resistance exercise) can be calculated. This allows for comparison of effect sizes between groups; and facilitate exploration of potential sources of heterogeneity in the overall sample. If heterogeneity (as reflected by the I^2 values) is lower in the subgroups than in the overall sample, then heterogeneity might be driven by differences between patient and interventional features across trials.

Sensitivity analyses are conducted by excluding all studies from the analysis that were judged to have high or unclear risk of bias in a specific domain. This analysis evaluate whether the pooled effect is robust against bias. For instance, a significant pooled effect may become insignificant when studies with low risk of selection bias are exclusively included. In this case, the effect is likely to be biased by inadequate randomization and/or allocation concealment.

3.5. Assessing risk of publication bias

Beyond risk of bias in individual studies, a well-conducted meta-analysis should also assess the presence of risk of bias across studies, i.e. publication bias. Publication bias describes the well-known fact that negative trials are less likely to be published than positive one [37]. This phenomenon is attributable to a number of reasons: (i) researchers, and especially clinician-researchers, are often reluctant to publish findings that contradict what they believe in and what they observe in clinical practice; (ii) manufacturers have spent a lot of time and money in the development and testing of a new drug or intervention devices, and would not welcome the idea that all this effort and investment would be becoming worthless; (iii) last but not least, scientific journals' editors are often more willing to publish positive trials than negative ones, because the former are considered more informative and/or interesting to the target readership, and would probably attract more citations which help improving impact factors of the journal.

Consequently, the absence of negative results will bias a meta-analysis towards overly positive results. There however is a simple procedure for detecting the presence of publication bias in meta-analyses. The funnel plot method makes use of the fact that smaller negative trials have a higher likelihood of not being published than large positive ones (it is much easier to hide the conduction of an n = 10 trial from the scientific community than that of an n = 10,000 trial). Asymmetry of funnel plots can be used to detect publication bias, and essentially this plot is a simple scatter plot of the effect sizes from individual studies against sample size or precision [38]. If there is publication bias because small negative trials remained unpublished, there will be an asymmetry of the funnel plot there of a gap at the bottom of the graph (Fig. 2) [39]. While there also are statistical tests (e.g Egger's test) for publication bias available, they often lack statistical power. Simple visual inspection of funnel plots will be sufficient for most basic meta-analyses.

4. Summary

Meta-analysis is a well-established method of pooling individual trial data into a summary estimate and the recent applications of network meta-analysis has advanced researchers' toolbox in handling trials evaluating heterogeneous interventions. However, the implementation of meta-analysis results in routine CIM practice is not straight forward as CIM students and practitioners have variable intention is applying such evidence in their practice [40]. Aside from teaching CIM students and clinicians basic competency in evidence based healthcare [41,42], the implementation of evidence in CIM will require stronger facilitation and dissemination efforts, let alone methodological innovations that allow creation of evidence that is more compatible to the complex nature of CIM practices [43,44].

References

[1] J.M. Last, A Dictionary of Public Health, Oxford University Press, Oxford, 2007.
[2] K. Hansen, K. Kappel, The proper role of evidence in complementary/alternative medicine, J. Med. Philos. 35 (1) (2010) 7–18.
[3] World Health Organization, WHO Traditional Medicine Strategy: 2014–2023, (2013) .
[4] S.L. Normand, Meta-analysis: formulating, evaluating, combining, and reporting, Stat. Med. 18 (3) (1999) 321–359.
[5] R.S.T. Ho, C.H.L. Wong, V.C.H. Chung, Medical synopsis: can acupuncture be an alternative treatment option for patients with refractory functional dyspepsia, Adv. Integr. Med. 2 (3) (2015) 143–145.
[6] J.T.M. Chow, C.H.L. Wong, V.C.H. Chung, Is Tian Jiu Therapy effective in reducing asthma symptoms? Adv. Integr. Med. 3 (2) (2016) 64–65.
[7] C.H.L. Wong, J.T.M. Chow, V.C.H. Chung, Should Tai Chi be part of cardiac rehabilitation programme for patients with chronic heart failure? Adv. Integr. Med. 3 (2) (2016) 62–63.
[8] H. Cramer, Why meta-analyses are important for complementary and alternative medicine research. Editorial, World J. Meta-Anal. 3 (2015) 1–3.
[9] P. Sedgwick, Meta-analyses: what is heterogeneity, BMJ (2015) h1435.
[10] J. Wardle, A. Steel, Review: systematic reviews in integrative medicine: a clinician's guide to publication, Adv. Integr. Med. 2 (2) (2015) 103–109.
[11] K. Linde, N. Clausius, G. Ramirez, D. Melchart, F. Eitel, L.V. Hedges, W.B. Jonas, Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials, Lancet (London Engl.) 350 (9081) (1997) 834–843.
[12] A. Shang, K. Huwiler-Muntener, L. Nartey, P. Juni, S. Dorig, J.A. Sterne, D. Pewsner, M. Egger, Are the clinical effects of homoeopathy placebo effects? Comparative study of placebo-controlled trials of homoeopathy and allopathy, Lancet (London Engl.) 366 (9487) (2005) 726–732.
[13] V.C. Chung, R.S. Ho, X. Wu, D.H. Fung, X. Lai, J.C. Wu, S.Y. Wong, Are meta-analyses of Chinese herbal medicine trials trustworthy and clinically applicable? A cross-sectional study, J. Ethnopharmacol. 162 (2015) 47–54.
[14] G. Salanti, Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool, Res. Synth. Methods 3 (2) (2012) 80–97.
[15] C.L. Foley, S. Tadtayey, T.J. Greenwell, Alternative or complementary treatments in urology, Med. Ther. Urol. (2010) 221–232.
[16] S. Bent, C. Kane, K. Shinohara, J. Neuhaus, H.E.S.H. Goldberg, et al., Saw palmetto for benign prostatichyperplasia, N. Engl. J. Med. 354 (6) (2006) 557–566.
[17] T. Wilt, A. Ishani, R.M. Donald, Serenoa repens for benign prostatic hyperplasia, Cochrane Database Syst. Rev. 3 (3) (2002) D001423.
[18] T. Wilt, A. Ishani, R.M. Donald, I. Rutks, G. Stark, Pygeum africanum for benign prostatic hyperplasia (Review), Cochrane Collab. 1 (1) (1998) D001044.
[19] T. Wilt, R.M. Donald, A. Ishani, I. Rutks, G. Stark, Cernilton for benign prostatic hyperplasia, Cochrane Database Syst. Rev. 2 (2) (2000) D001042.
[20] S.A. Kaplan, Update on the American Urological Association guidelines for the treatment of benign prostatic hyperplasia, Rev. Urol. 8 (Suppl. 4) (2006) S10.
[21] V.C. Chung, P.H. Ma, H.H. Wang, J.J. Wang, L.C. Hong, X. Wei, S.Y. Wong, J.L. Tang, S.M. Griffiths, Integrating traditional chinese medicine services in community health centers: insights into utilization patterns in the pearl river region of china, Evid.-Based Complementary Altern. Med.: eCAM 2013 (2013) 426360.
[22] V.C. Chung, X. Wu, P.H. Ma, R.S. Ho, S.K. Poon, D.S. Hui, S.Y. Wong, J.C. Wu, Chinese herbal medicine and salmeterol and fluticasone propionate for chronic obstructive pulmonary disease: systematic review and network meta-analysis, Medicine 95 (20) (2016) e3702.

[23] V. Chung, L. Liu, Z. Bian, Z. Zhao, W. Leuk Fong, W.F. Kum, J. Gao, M. Li, Efficacy and safety of herbal medicines for idiopathic Parkinson's disease: a systematic review, Mov. Disord.: Off. J. Mov. Disord. Soc. 21 (10) (2006) 1709–1715.
[24] A. Liberati, D.G. Altman, J. Tetzlaff, C. Mulrow, P.C. Gotzsche, J.P. Ioannidis, M. Clarke, P.J. Devereaux, J. Kleijnen, D. Moher, The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration, Ann. Intern. Med. 151 (4) (2009) W65–94.
[25] S. Donegan, P. Williamson, U. D'Alessandro, C. Tudur Smith, Assessing key assumptions of network meta-analysis: a review of methods, Res. Synth. Methods 4 (4) (2013) 291–323.
[26] V.C. Chung, X. Wu, P. Lu, E.P. Hui, Y. Zhang, A.L. Zhang, A.Y. Lau, J. Zhao, M. Fan, E. T. Ziea, B.F. Ng, S.Y. Wong, J.C. Wu, Chinese herbal medicine for symptom management in cancer palliative care: systematic review and meta-analysis, Medicine 95 (7) (2016) e2793.
[27] V.C. Chung, X. Wu, E.P. Hui, E.T. Ziea, B.F. Ng, R.S. Ho, K.K. Tsoi, S.Y. Wong, J.C. Wu, Effectiveness of Chinese herbal medicine for cancer palliative care: overview of systematic reviews with meta-analyses, Sci. Rep. 5 (2015) 18111.
[28] F. Angst, A. Aeschlimann, J. Angst, The minimal clinically important difference (MCID) raised the significance of outcome effects above the statistical level, with methodological implications for future studies, J. Clin Epidemiol . 82 (2017) 128–136.
[29] J.P. Higgins, C. Ramsay, B.C. Reeves, J.J. Deeks, B. Shea, J.C. Valentine, P. Tugwell, G. Wells, Issues relating to study design and risk of bias when including non-randomized studies in systematic reviews on the effects of interventions, Res. Synth. Methods 4 (1) (2013) 12–25.
[30] L. Thabane, L. Mbuagbaw, S. Zhang, Z. Samaan, M. Marcucci, C. Ye, M. Thabane, L. Giangregorio, B. Dennis, D. Kosa, V. Borg Debono, R. Dillenburg, V. Fruci, M. Bawor, J. Lee, G. Wells, C.H. Goldsmith, A tutorial on sensitivity analyses in clinical trials: the what, why, when and how, BMC Med. Res. Methodol. 13 (2013) 92.
[31] V. Chung, Assessing internal validity of clinical evidence on effectiveness of CHinese and integrative medicine: proposed framework for a CHinese and Integrative Medicine Evidence RAting System (CHIMERAS), Eur. J. Integr. Med. 7 (4) (2015) 332–341.
[32] J.P.T. Higgins, S. Green, Cochrane Handbook for Systematic Reviews of Interventions, John Wiley & Sons Ltd, West Sussex, 2008.
[33] B.C. Wallace, C.H. Schmid, J. Lau, T.A. Trikalinos, Meta-analyst: software for meta-analysis of binary, continuous and diagnostic data, BMC Med. Res. Methodol. 9 (2009) 80.
[34] A. Chaimani, J.P. Higgins, D. Mavridis, P. Spyridonos, G. Salanti, Graphical tools for network meta-analysis in STATA, PloS One 8 (10) (2013) e76654.
[35] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.
[36] J.P. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman, Measuring inconsistency in meta-analyses, BMJ 327 (7414) (2003) 557–560.
[37] P. Sedgwick, What is publication bias in a meta-analysis? BMJ 351 (2015) h4419.
[38] P. Sedgwick, L. Marston, How to read a funnel plot in a meta-analysis, BMJ 351 (2015) h4718.
[39] M. Egger, G. Davey Smith, M. Schneider, C. Minder, Bias in meta-analysis detected by a simple, graphical test, BMJ 315 (7109) (1997) 629–634.
[40] B.J. Anderson, B. Kligler, H.W. Cohen, P.R. Marantz, Survey of chinese medicine students to determine research and evidence-based medicine perspectives at pacific college of oriental medicine, Explore (New York, N.Y.) 12 (5) (2016) 366–374.
[41] H. Zwickey, H. Schiffke, S. Fleishman, M. Haas, A. Cruser d, R. LeFebvre, B. Sullivan, B. Taylor, B. Gaster, Teaching evidence-based medicine at complementary and alternative medicine institutions: strategies, competencies and evaluation, J. Altern. Complementary Med. (New York, N.Y.) 20 (12) (2014) 925–931.
[42] S. Jackson, G. Scambler, Perceptions of evidence-based medicine: traditional acupuncturists in the UK and resistance to biomedical modes of evaluation, Sociol. Health Illness 29 (3) (2007) 412–429.
[43] V. Chung, R.S. Ho, X.W. Wu, J.C. Wu, Incorporating traditional Chinese medicine syndrome differentiation in randomized trials: methodological issues, Eur. J. Integr. Med. 8 (6) (2016) 898–904.
[44] S.R. Shuldiner, V. Chung, X. Wu, J. Ching, R.S. Ho, P.K. Cheong, W. Wong, J.C. Wu, Methodological challenges in mapping chinese medicine syndrome with conventional diagnosis: implications for multi-centre trials in integrative medicine, Eur. J. Integr. Med. 7 (4) (2015) 358–364.