

The Application of Systematic Review and Meta-analysis Concepts in Summarizing the Findings of Observational Studies

Ali-Akbar Haghdoost, MD, PhD^{1, 2}
Behnam Sadeghirad, PharmD³
Behzad Hajarizadeh, MD, MPH⁴
Ali Mirzazadeh, MD, MPH⁵

1 Department of Epidemiology, Kerman University of Medical Sciences, Physiology Research Centre, Kerman, Iran

2 London School of Hygiene and Tropical Medicine, London, UK

3 Neuroscience Research Centre, University of Medical Sciences, Kerman, Iran

4 UNAIDS country office, Tehran, Iran

5 Kerman University of Medical Sciences, Physiology Research Centre, Kerman, Iran

Corresponding author:

Ali Mirzazadeh, MD, MPH,
Vice Chancellor for Research
Department, Kerman Physiology
Research Center, Ebn-e-Sina
Ave. Jahad Blvd., Thamasabad
crossroad, Kerman, Iran.
Tel: +98-341-2263787
Fax : +98-341-2264097
Email: amirzazadeh@kmu.ac.ir

Abstract: In contrast to traditional reviews, systematic reviews explore the literature on a specific topic comprehensively using a well-defined protocol and recruit the findings of eligible studies after the assessment of their qualities, using a clear and reproducible method. Checking the assumptions, we may merge the findings of recruited studies in a systematic review and use meta-analysis techniques not only to estimate the pooled effect but also to assess the possible sources of heterogeneity and the extend of heterogeneity in the findings .

Generally, randomized clinical trials are the most common type of studies which are recruiting in systematic reviews and meta-analyses. However, we couldn't ignore the importance of observational studies particularly in sensitivity topics with ethical limitations against interventional studies. Although we are usually faced with wide variations in the methodology of observational studies even in narrow topics, systematic review and meta-analysis of their results can generate valuable findings. They are particularly useful for explaining the source of heterogeneity in the results of primary observational studies .

In this paper, we review the basic concepts of the systematic review and meta-analysis and their main applications in summarizing the findings of observational studies with respect to their advantages and limitations.

Keywords: *systematic review, meta-analysis, observational studies, descriptive studies, heterogeneity*

Iran J Psychiatry 2007; 2:132-136

Although systematic review (SR) and meta-analysis (MA) are new methodological terms and have been added to the research encyclopedia since three decades ago, nowadays they are common terms for their decisive applications.

Briefly, SR introduces simple but critical methodology to review and select the best available research findings on a specific topic to minimize the selection bias in picking up the best accurate evidences. MA, however, uses statistical techniques to help us to combine the findings of comparable studies, present the aggregated statistics and check how significant the differences between the findings of studies are (i.e.: their heterogeneity).

There is a deep controversy surrounding the eligible type of studies for a MA. Some experts only recommend Randomized Clinical Trials (RCTs) (1), while others include evidences from a diversity of

sources (2). In fact, SR and MA are much more applicable to those studies that share a similar methodology and address comparable research questions. Therefore, the principles of the SR and MA are more applicable to the findings of comparable RCTs.

Nevertheless, observational studies are very common type of studies that either describe variables (descriptive studies) or explore the relationship between variables (analytical studies).

Considering the limitations, using SR and MA, we may explore the findings of observational studies conclusively. The concepts of the SA and MA may not be easily applicable to the findings of observational studies; nonetheless, we believe that the SA and MA techniques have some additional advantages which may help to propose more appropriate conclusions through combining the findings of observational studies.

In this paper, we presented the basic concepts of the SA and MA and their main applications in summarizing the findings of observational studies. Moreover, we highlighted the limitations of these techniques.

Definition of the SR and MA

A SR is a summary of the literature and it starts with a well defined question and continues by a systematic searching protocol to find out the most relevant studies. In the next step, all evidences are critically appraised with specific appraisal tools and irrelevant or low quality studies are excluded. Hence, this process sometimes may lead to a SR with no qualified study (3, 4).

Suppose you are going to run a SR on the protective effect of male circumcision on the transmission of HIV. Reviewing the available literature, you may find an extensive controversy among the study findings. The odds ratio between male circumcision and HIV transmission in the general population varies from 0.21 and 1.90 (5). Moreover, your discussion with national and international experts might amplify your confusion. What would be the best strategy in dealing with such confusion?

Searching all the available evidences and checking their validities is extremely necessary. It is not wise to rely on the findings of only a few papers or to accept the findings of those studies that did not follow a clear and justified methodology.

In fact, SR helps you to formulate the steps that you need to follow to address your question.

With a well-defined systematic search strategy you will gather all the available evidences and check their qualities. (Figure 1)

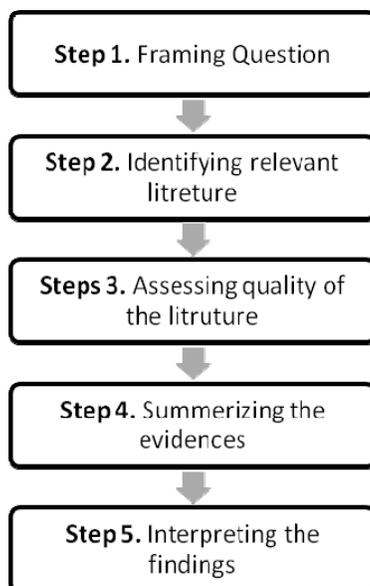


Figure 1. The main steps in a systematic review

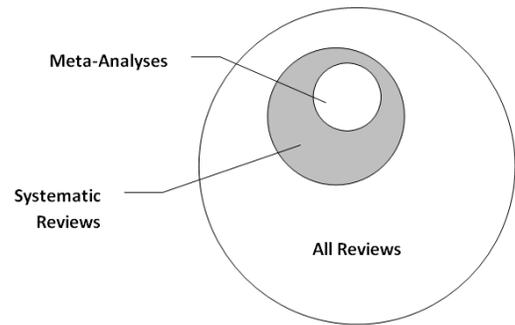


Figure 2. Reviews, Systematic reviews and meta-analysis

In contrast to SR, traditional reviews do not follow a clear data collection method that may lead to their findings to section bias. In addition, the selected studies are interpreted based on the researcher opinions (6) while SRs are more robust to these possible errors (7) (Figure 2).

Having collected the findings of well-qualified evidences, you may recognize that different studies reported diverse OR's regarding the effect of male circumcision on the odds of HIV transmission. In the next step, you should combine the findings of different studies to calculate the most appropriate pooled OR. MA methods help you to compute such an estimate using weighted average of the eligible study's findings. MA is defined as the statistical analysis of different results (effects) for the purpose of integrating the findings (Pooled effects)(8).

The term "effect" refers to any measure of association between exposure and outcome (e.g. odds ratio, risk ratio). Individual studies may be too small to produce precise effects (i.e. explore the significant effects) while MA improves the precision by combining the findings of comparable studies. However, we should mention that in descriptive studies the mean (the average of age at first sex) or prevalence (the frequency of condom use) are our effect and we may like to estimate the overall mean or prevalence based on the observed effects in individual descriptive studies.

In addition, MA explores the variations of findings (i.e.: the heterogeneity among the effects estimated in different studies). With MA in subgroups or some advanced statistical methods (e.g. Meta regression models) we explore the source of heterogeneity and interpret the differences in the findings of individual studies. For example, the wide variation of the ORs between male circumcision and HIV can be explained by the difference in their study designs (for example, the ORs would be closer to each other if you compare the findings of studies with similar designs; i.e., case-control, cross-sectional or cohort design) (Figure 3). Based on this finding, we can imply that the study design is one of the main sources of heterogeneity.

The Application of Systematic Review and Meta-analysis

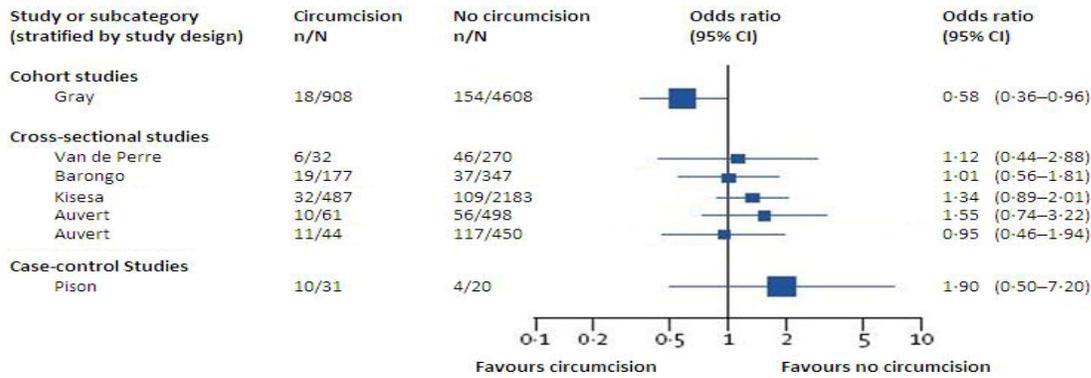


Figure 3. Forest plot adopted from *Lancet Infect Dis.* 2005; 5: 165–73 (5). ORs greater than 1 indicates increased risk of HIV infection with circumcision, and ORs less than 1 indicates decreased risk of HIV infection with circumcision. A wide heterogeneity in the ORs indicates that there are some underlying factors which cause ORs to differ from one study to another. Making strata based on the study design showed closer ORs. Therefore, we can conclude that study design is one of the main sources of heterogeneity and we can quantify the impact of the study design using meta-analytic techniques. (We modified the original graph to make it easier for the readers to understand how an extraneous factor like the Study Design can induce heterogeneity in the effects, such as ORs.)

In general, MA is a statistical method which only aggregates the findings of comparable and eligible studies selected in a SR. However, some limitations may force us to report the findings of a SR without using MA methods. Sometimes we cannot combine the findings of the selected studies due to their methodological differences. For instance, studies might measure their variables using different definitions or tools. In addition, you may even use SR principals to search qualitative studies, while MA only combines the findings of quantitative studies. Lastly, SR may select few eligible studies while for a meaningful MA we need at least a minimum number of comparable studies.

The application of SR and MA in RCTs

A RCT (Randomized Clinical Trial) is considered the most reliable (Gold Standard) method to determine which medical laboratory test/ interventions work the best. The participants are allocated randomly between different intervention arms while they do not know which intervention they are receiving (blindness). Therefore, the results are less prone to selection and information biases and confounding errors. (9). All the unique characteristics of RCTs, in compare to observational studies, suggest that choosing RCTs is the gold standard and RCTs are considered the best evidence in primary researches(10).

SR and MA on randomized clinical trials with limited biases and low level of heterogeneity, produce the top

level of evidences in medicine. That's why most of the SRs only included such study types (4). MA of clinical trials can recommend effective treatments on time or lead to the timely identification of adverse effects (11).

However, some biases such as loss to follow up (withdrawal) after randomization, affect RCT's validity; moreover, the results of negative trails have a less chance of being published (Publication bias). Thus, SR and MA on published trials are prone to show the effects stronger than the real ones (12). Moreover, due to ethical considerations, we cannot test every intervention following the general rules of RCTs strictly.

The application of SR and MA in observational studies

Based on the above explanation, you can use the SR principles to search and select any type of studies including observational studies. Without any doubt, SR is an efficient method for presenting the findings of comparable observational studies with a clear and straight forward approach to convince readers that you had a comprehensive search. A well-conducted SR allows a more objective appraisal of the evidence than the traditional narrative reviews.

Observational designs may lack the experimental element of a random allocation to an intervention and may rely on the observed associations (e.g., the association between male circumcision and risk of HIV infection) (13). These designs have long been used in the evaluation of educational programs (14) and exposures that might cause disease or injury (15).

In fact, observational studies are intending to cover some limitations of RCTs. For instance, we cannot randomize some risk factors because they relate to inherent human characteristics or practices; and we should bear in mind that exposing subjects to harmful risk factors is unethical (16). Referring to our first example, to perform a study on the preventive effects of male circumcision on the transmission of HIV, you have to run a case control, cross-sectional or cohort

study since RCT design will be practically impossible (17). Observational data may also be needed to assess the effectiveness of an intervention in a community as opposed to the special setting of a controlled trial (18). They estimate the prevalence or incidence of diseases or their determinants in communities, and SR and even MA methods may help us to explore their findings more comprehensively (e.g. prevalence of hypertension (19), chronic kidney disease (20), Hepatitis B virus infection (21), etc).

Although MA is merely preferred for RCTs (22-24), the number of published MAs concerning observational studies in health has increased substantially during the past 4 decades (678 in 1955-1992, 525 in 1992-1995, and more than 400 in 1996 alone) (13, 25). Although the results of MAs based on observational studies have particular challenges due to inherent biases in individual studies and differences in their study designs, they may help us to understand and quantify sources of variability in results across studies. MA of observational studies also is the only available method for assessing the efficacy and effectiveness of some interventions; therefore, their publication is increasing in numbers (13). Thus, a clear understanding of the advantages and limitations of statistical syntheses of observational data is needed (26, 27).

MA may explain the observed heterogeneity between the results of individual studies (28); for instance, it may help us to find the source of heterogeneity; most readers pay less attention to this function of MA. However, exploring the source of heterogeneity is the most important function in MA of descriptive studies. In addition, it could help us to discuss why the prevalence or incidence of a disease was reported differently.

Zhang and Rothenbacher performed a SR design to assess the prevalence of chronic kidney disease (CKD) using available population-based studies with published data in MEDLINE (20). This SR design provides useful information on equation, age, gender and ethnic-specific prevalence of CKD in various population-based studies. Almost the same method was used by Goodman et al to estimate the prevalence and the mean score for child mental health status in different ethnic groups (29). Considering the main pitfalls of prevalence studies, in addition to the usual criteria used to assess the quality of studies, they also used different inclusion/exclusion criteria from methodological point of view (e.g. calculating a minimum sample size for studies or their subgroups, measurements of mental health, etc).

Usually in SR and MA of descriptive studies, a minimum sample size is considered as one of the inclusion criteria because null findings involving very small numbers become 'uninformative', being better interpreted as an absence of evidence rather than evidence of an absence. Moreover, because a null finding based on so few individuals is so uninformative such findings are unlikely to be published, and

publication bias is likely to become particularly acute (29).

Surprisingly, MAs belong to observational studies (28) even when they are applied to RCTs (30).

Running SR and MA using such data can provide useful information not only for health care planning but also for future researchers to perform studies that are more succinct and less prone to errors. Recently such designs have been used in different studies (19-21, 31).

Limitations of SR and MA in observational studies

Based on the above explanation, it is important to take into account the limitations of SR and MA in exploring the data of observational studies. These limitations are applicable in SA and MA of interventional studies but with less extend.

1- Analytic observational studies yield estimates of associations which may deviate from true underlying relationships beyond the play of chance. This may be due to the effects of confounding factors, biases or both. Accordingly, MA of a number of biased studies will result in a biased pooled estimation. In the case of confounding factors, the usual approach is to adjust the effect size for confounding factors using multivariate analysis. However, it is very difficult to believe that all observational studies included in a MA report an effect size adjusted for exactly similar confounding factors. On the other hand, the associations resulted from RCTs are less prone to such effects because of random allocation. These limitations in methodology may lead to contradictory results. For instance, in a MA evaluating association between beta carotene intake and cardiovascular mortality, combining cohort studies showed a significantly protective effect for beta carotene while the pooled result from RCTs showed a moderate adverse effect of beta carotene (26).

2- Diversity of methodology in observational studies per se may lead to contradictory results even when different types of observational studies enter into analysis. For instance, a comprehensive MA study on whether the higher intake of saturated fat is associated with the increased risk of breast cancer indicated an association for the case control but not for cohort studies (31). Similar result is apparent in figure 3.

3- MA of descriptive observational studies is exposed to some other limitations in addition to biases. Combining the results of descriptive studies which apply the classic MA methods may not be the most appropriate approach. For instance, coming back to one of the mentioned examples, estimating the prevalence of hypertension in Iran weights of each study in MA is usually the reverse of their standard errors (SE) which is a function of sample sizes. However, as the goal is to estimate the prevalence of hypertension in the whole country, it is more logical to weight the reported provincial estimates based on the population size of the provinces and not based on the size of the studies (21). However, such a limitation is less important if you aim

to find the source of heterogeneity between the findings of descriptive studies.

Conclusion

Despite these challenges, following the rules of SR helps us to present a fair and comprehensive representation of the best available evidences in all types of studies including observational studies. In addition, accepting the limitations in observational studies, MA of observational studies may not only illustrate the best summary estimate but may also explore the source of heterogeneity which is even more important than the summary estimate in most of the cases.

References

1. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;312: 71-72.
2. Greenhalgh T. *How to Read a Paper: The Basics of Evidence-based Medicine*. London: BMJ Publishing; 1997.
3. Montori VM, Swiontkowski MF, Cook DJ. Methodologic issues in systematic reviews and meta-analyses. *Clin Orthop Relat Res* 2003 ; 413: 43-54.
4. Higgins JPT, Green S . *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.0 [updated February 2008]. The Cochrane Collaboration 2008.
5. Siegfried N, Muller M, Deeks J, Volmink J, Egger M, Low N, et al. HIV and male circumcision- a systematic review with assessment of the quality of studies. *Lancet Infect Dis* 2005; 5:165-173.
6. Macbeth F, Overgaard J. Expert reviews, systematic reviews and meta-analyses. *Radiother Oncol* 2002;64 : 233-234.
7. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. *World J Surg* 2005; 29: 582-587.
8. Khoshdel A, Attia J, Carney SL. Basic concepts in meta-analysis: A primer for clinicians. *Int J Clin Pract* 2006; 60 :1287-1294.
9. Kadar N. The randomized clinical trial methodology is the most efficient and effective in evaluating new treatment strategies. *Gynecol Oncol* 1997;64 :185.
10. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ* 2001; 323: 42-46.
11. Zwahlen M, Renehan A, Egger M. Meta-analysis in medical research: potentials and limitations. *Urol Oncol* 2008; 26: 320-329.
12. Khan KS. *Systematic reviews to support evidence-based medicine : how to review and apply findings of healthcare research*. RSM Press; 2003.
13. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of Observational Studies in Epidemiology: a Proposal for reporting. Meta-analysis of Observational Studies in Epidemiology(MOOSE) group. *JAMA* 2000; 283 : 2008-2012.
14. Sipe TA, Curlette WL. A meta-synthesis of factors related to educational achievement: a methodological approach to summarizing and synthesizing meta-analyses. *Int J Educ Res* 1996; 25: 583-698.
15. Ioannidis JP, Lau J. Pooling research results: benefits and limitations of meta-analysis. *Jt Comm J Qual Improv* 1999; 25:462-469.
16. Lipsett M. Occupational exposure to diesel exhaust and lung cancer: a meta-analysis. *Am Public Health Assoc* ;1999.P.1009-1017.
17. Vickers A, Cassileth B, Ernst E, Fisher P, Goldman P, Jonas W, et al. How should we research unconventional therapies? Apanel report from the conference on complementary and alternative medicine research methodology, National Institutes of Health. *Int J Technol Assess Health Care*1997;13: 111-121.
18. Mann CC. Can meta-analysis make policy? *Science* 1994; 266 :960-962.
19. Haghdoost AA, Sadeghirad B, Rezazadehkermani M. Epidemiology and Heterogeneity of Hypertension in Iran: A Systematic Review. *Arch Iran Med* 2008;11:444-452.
20. Zhang QL , Rothenbacher D. Prevalence of chronic kidney disease in population-based studies: Systematic review. *BMC Public Health* 2008; 8:117-130.
21. Alavian S, Hajarizadeh B, Ahmadzad Asl M, Kabir A, Bagheri Lankarani K. Hepatitis B virus infection in Iran: A systematic review. . *Hep Mon* 2008; 8: 281-294.
22. Greenland S. Invited Commentary: Re: HA Critical Look at Some Popular Meta-Analytic Methods. *Am J Epidemiol* 1994;140 : 390-396.
23. Shapiro S. Meta-analysis/Shmeta-analysis. *Am J Epidemiol* 1994 ;140: 771-778.
24. Lau J, Ioannidis JP, Schmid CH. Summing up evidence. *Lancet* 1998; 351:123-127.
25. Stroup DF, Thacker SB, Olson CM, Glass RM. Characteristics of meta-analyses submitted to a medical journal. *International Congress on Biomedical Peer Review and Scientific Publication*. *Am Med Assoc* 1997. p. 17-21.
26. Egger M, Schneider M, Smith GD. Meta-analysis spurious precision? Meta-analysis of observational studies. *BMJ* 1998; 316:140-144.
27. Huston P, Naylor CD. Health services research: reporting on studies using secondary data sources. *CMAJ* 1996;155:1697-1709.
28. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ* 1997; 315:1533-1537.
29. Goodman, A, Patel, V, and Leon, D. Child mental health differences amongst ethnic groups in Britain: a systematic review. *BMC Public Health* 2008; 25: 258-268.
30. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of

- randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; 352: 609-613.
31. Boyd FN, Martin LJ, Noffel M, Lockwood GA, Tritchler DL. A meta-analysis of studies of dietary fat and breast cancer. *Br J Cancer* 1993; 68: 627-636.