

# World Journal of *Meta-Analysis*

*World J Meta-Anal* 2018 August 28; 6(3): 21-28





**EDITORIAL**

- 21 Improving the conduct of meta-analyses of observational studies  
*Lee PN*

## ABOUT COVER

Editorial Board Member of *World Journal of Meta-Analysis*, Ai-Ping Lu, MD, PhD, Professor, School of Chinese Medicine, Hong Kong Baptist University, Hong Kong, China

## AIM AND SCOPE

*World Journal of Meta-Analysis* (*World J Meta-Anal*, *WJMA*, online ISSN 2308-3840, DOI: 10.13105) is a peer-reviewed open access academic journal that aims to guide clinical practice and improve diagnostic and therapeutic skills of clinicians, with a specific focus on meta-analysis, systematic review, mixed-treatment comparison, meta-regression, overview of reviews.

*WJMA* covers a variety of clinical medical fields including allergy, anesthesiology, cardiac medicine, clinical genetics, clinical neurology, critical care, dentistry, dermatology, emergency medicine, endocrinology, family medicine, gastroenterology and hepatology, geriatrics and gerontology, hematology, immunology, infectious diseases, internal medicine, obstetrics and gynecology, oncology, ophthalmology, orthopedics, otolaryngology, pathology, pediatrics, peripheral vascular disease, psychiatry, radiology, rehabilitation, respiratory medicine, rheumatology, surgery, toxicology, transplantation, and urology and nephrology, while maintaining its unique dedication to systematic reviews and meta-analyses.

## INDEXING/ABSTRACTING

*World Journal of Meta-Analysis* (*WJMA*) is now abstracted and indexed in Emerging Sources Citation Index (Web of Science), China National Knowledge Infrastructure (CNKI), and Superstar Journals Database.

## EDITORS FOR THIS ISSUE

Responsible Assistant Editor: *Xiang Li*  
Responsible Electronic Editor: *Yun-Xiao Jian Wu*  
Proofing Editor-in-Chief: *Lian-Sheng Ma*

Responsible Science Editor: *Fang-Fang Ji*  
Proofing Editorial Office Director: *Jin-Lei Wang*

NAME OF JOURNAL  
*World Journal of Meta-Analysis*

ISSN  
ISSN 2308-3840 (online)

LAUNCH DATE  
May 26, 2013

EDITOR-IN-CHIEF  
Giuseppe Biondi-Zoccai, MD, Assistant Professor,  
Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome, Latina 04100, Italy

EDITORIAL BOARD MEMBERS  
All editorial board members resources online at <http://www.wjgnet.com/2308-3840/editorialboard.htm>

EDITORIAL OFFICE  
Jin-Lei Wang, Director  
*World Journal of Meta-Analysis*  
Baishideng Publishing Group Inc

7901 Stoneridge Drive, Suite 501,  
Pleasanton, CA 94588, USA  
Telephone: +1-925-2238242  
Fax: +1-925-2238243  
E-mail: [editorialoffice@wjgnet.com](mailto:editorialoffice@wjgnet.com)  
Help Desk: <http://www.f6publishing.com/helpdesk>  
<http://www.wjgnet.com>

PUBLISHER  
Baishideng Publishing Group Inc  
7901 Stoneridge Drive, Suite 501,  
Pleasanton, CA 94588, USA  
Telephone: +1-925-2238242  
Fax: +1-925-2238243  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <http://www.f6publishing.com/helpdesk>  
<http://www.wjgnet.com>

PUBLICATION DATE  
August 28, 2018

COPYRIGHT  
© 2018 Baishideng Publishing Group Inc. Articles published by this Open-Access journal are distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license.

SPECIAL STATEMENT  
All articles published in journals owned by the Baishideng Publishing Group (BPG) represent the views and opinions of their authors, and not the views, opinions or policies of the BPG, except where otherwise explicitly indicated.

INSTRUCTIONS TO AUTHORS  
<http://www.wjgnet.com/bpg/gerinfo/204>

ONLINE SUBMISSION  
<http://www.f6publishing.com>

## Improving the conduct of meta-analyses of observational studies

Peter N Lee

Peter N Lee, P.N. Lee Statistics and Computing Ltd., Sutton SM2 5DA, Surrey, United Kingdom

ORCID number: Peter N Lee (0000-0002-8244-1904).

Author contributions: Lee PN wrote this editorial.

Conflict-of-interest statement: The author has no relevant conflict of interest to declare.

Open-Access: This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

Manuscript source: Invited Manuscript

Correspondence to: Peter N Lee, MA, MSc, Senior Statistician, Director, P.N. Lee Statistics and Computing Ltd., 17 Cedar Road, Sutton SM2 5DA, Surrey, United Kingdom. [peterlee@pnlee.co.uk](mailto:peterlee@pnlee.co.uk)  
Telephone: +44-20-6428265  
Fax: +44-20-8642135

Received: June 8, 2018

Peer-review started: June 8, 2018

First decision: July 11, 2018

Revised: July 16, 2018

Accepted: August 4, 2018

Article in press: August 4, 2018

Published online: August 28, 2018

### Abstract

The author, who has published numerous meta-analyses of epidemiological studies, particularly on tobacco, comments on various aspects of their content. While such meta-analyses, even when well conducted, are more

difficult to draw inferences from than are meta-analyses of clinical trials, they allow greater insight into an association than do simple qualitative reviews. This editorial starts with a discussion of some problems relating to hypothesis definition. These include the definition of the outcome, the exposure and the population to be considered, as well as the study inclusion and exclusion criteria. Under literature searching, the author argues against restriction to studies published in peer-reviewed journals, emphasising the fact that relevant data may be available from other sources. Problems of identifying studies and double counting are discussed, as are various issues in regard to data entry. The need to check published effect estimates is emphasised, and techniques to calculate estimates from material provided in the source publication are described. Once the data have been collected and an overall effect estimate obtained, tests for heterogeneity should be conducted in relation to different study characteristics. Though some meta-analysts recommend classifying studies by an overall index of study quality, the author prefers to separately investigate heterogeneity by those factors which contribute to the assessment of quality. Reasons why an association may not actually reflect a true causal relationship are also discussed, with the editorial describing techniques for investigating the relevance of confounding, and referring to problems resulting from misclassification of key variables. Misclassification of disease, exposure and confounding variables can all produce a spurious association, as can misclassification of the variable used to determine whether an individual can enter the study, and the author points to techniques to adjust for this. Issues relating to publication bias and the interpretation of "statistically significant" results are also discussed. The editorial should give the reader insight into the difficulties of producing a good meta-analysis.

**Key words:** Hypothesis definition; Literature searching; Heterogeneity; Publication bias; Misclassification; Confounding; Meta-analysis

© The Author(s) 2018. Published by Baishideng Publishing

Group Inc. All rights reserved.

**Core tip:** The author has published many meta-analyses of epidemiological studies, particularly on smoking, and the editorial comments on various aspects of their conduct. Areas covered include the definition of the hypothesis to be tested, literature searching and data entry, as well as methods to test for heterogeneity and investigate such issues as confounding, misclassification and publication bias. The need for well conducted meta-analyses and the difficulty in determining whether a "statistically significant" association is actually indicative of a causal relationship are discussed. The editorial should be helpful to readers inexperienced with the conduct of meta-analyses.

Lee PN. Improving the conduct of meta-analyses of observational studies. *World J Meta-Anal* 2018; 6(3): 21-28 Available from: URL: <http://www.wjgnet.com/2308-3840/full/v6/i3/21.htm> DOI: <http://dx.doi.org/10.13105/wjma.v6.i3.21>

## INTRODUCTION

Meta-analyses were originally designed to combine data from randomized controlled trials, with the Quality of Reporting of Meta-analyses statement<sup>[1]</sup> describing how the quality of such meta-analyses could be improved. Provided the trials which were being combined were of sufficiently similar design, and involved the same exposures and outcomes there was little difficulty in interpreting the overall effect estimate. Such meta-analyses clearly had greater power to detect relationships than had the individual studies being combined.

For many years attempts to summarize evidence on an association from multiple observational epidemiological studies were based on qualitative reviews. These reviews typically summarized the results of each study in a paragraph or two, and then attempted to draw an overall conclusion. International Agency for Research on Cancer monographs was often qualitative and it is sometimes difficult to see the process by which the overall conclusion had been reached.

Bringing meta-analysis techniques to the field of observational studies seemed attractive in that it provided some sort of quantitative overall assessment, but there was initially considerable concern about the validity of combining results from studies using different designs and methods, and conducted in different countries and time periods where the nature of the exposure may have varied. While there is clearly some element of truth in the criticism that one should not combine "apples and oranges", it became clear over the years that well-conducted meta-analyses can be extremely useful in assisting the judgement as to whether a relationship is a causal one. Particularly where the association is strong is consistently seen in multiple well conducted studies, and there is no source of confounding or bias that materially affects the estimates, one seems to be on safe grounds to

conclude that a causal relationship exists.

Over the years, I and my colleagues at P.N. Lee Statistics and Computing Ltd. have conducted a large number of meta-analyses relating to the health effects of tobacco. These consider effects of smoking generally<sup>[2-5]</sup>, different types of cigarette<sup>[6-8]</sup>, quitting<sup>[9-12]</sup>, smokeless tobacco<sup>[13-15]</sup>, Swedish "snus"<sup>[16-18]</sup> and nicotine replacement therapy<sup>[19]</sup>, as well as effects of parental smoking<sup>[20-22]</sup> and of environmental tobacco smoke exposure<sup>[23-28]</sup>. Mainly these meta-analyses relate to outcomes which are 1/0 variables (typically presence or absence of a disease), though some concern continuous outcomes such as forced expiratory volume<sup>[29,30]</sup> or cholesterol level<sup>[31]</sup>. While I do not have experience of conducting meta-analyses in other areas, I have also served as a reviewer for numerous meta-analyses submitted to journals and I hope that some of the knowledge I have accumulated will be of interest to others.

This editorial is not intended to describe how meta-analyses should be structured or presented. This is adequately described in the meta-analysis of observational studies in epidemiology proposals<sup>[32]</sup> and the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statements<sup>[33]</sup>, while the reporting of meta-analysis protocols is well covered by PRISMA-P<sup>[34]</sup>. Nor is it intended to cover all aspects of conducting a meta-analysis, what follows really being a collection of personal comments of mine on various aspects of meta-analyses of observational studies.

## DEFINING THE HYPOTHESIS TO BE TESTED

While some meta-analyses can be quite broad-ranging, relating a number of aspects of exposure of an agent to a number of different outcomes, others may be much more specific. It is important at the outset to clearly define the objectives of the work, and the hypotheses to be tested.

In a simple case, there may be one specific outcome of interest, and the study protocol should make clear what definitions of that outcome are allowed. For some diseases this may cause few problems, but for others this requires thought. In other cases, there may be several related outcomes, or specific subsets of the outcome, which are of interest. For example, in our review of the evidence relating smoking to chronic obstructive pulmonary disease (COPD), chronic bronchitis and emphysema<sup>[2]</sup>, we had to be careful to define what could be regarded as satisfactorily equivalent diseases, since COPD is a relatively recently used term, and we did not wish to exclude relevant older studies. We were also careful to record the basis of definition used in each study (e.g., symptoms reported on a questionnaire, mortality records), so that we could compare effect estimates according to this definition.

Similar considerations apply to the definition of exposure. First, we have to define what the exposure is - for smoking, for example, are we limiting attention to cigarettes, or do we include cigars and pipes? Are we



considering only exposure above a certain minimum level or any exposure? Are we considering ever exposure or current exposure? If we are considering current exposure are we comparing this with non-current exposure or with never exposure? Should we accept those who have ceased exposure very recently as part of the currently exposed group? Should we accept those with only a minimum lifetime exposure among the never exposed group? Often it may be useful to meta-analyse effect estimates for various exposure definitions. However, it is, in principle, a good idea to define in advance the main exposure of interest, to avoid being accused of trying various alternative definitions and then only reporting or emphasising the one that best shows the association of interest.

For both outcome and exposure, a balance has to be struck between using narrow definitions which may seriously limit the number of eligible studies, or allowing broader definitions which will increase the number of studies (and thus the workload and costs) and may hamper interpretation of the results.

In some situations, the hypothesis of interest is to be tested among a subset of the population. For example, when studying the relationship of environmental tobacco smoke exposure to a disease, it is usual to restrict attention to those who have never smoked (as exposure to tobacco smoke constituents from smoking is typically two orders of magnitude higher than from environmental tobacco smoke exposure). Here, one needs to define whether it is acceptable to include results from studies which include those with minimum lifetime cigarette consumption among the definition of never smoking.

One also has to define study inclusion and exclusion criteria. Are we restricting attention to certain study designs, perhaps only considering cohort studies, or certain sub-populations, such as employed persons? Are we excluding studies in children, or in adults who have relevant co-existing diseases or conditions, or who work in high-risk occupations? Are we only interested in studies which provide dose-response results? There are many possibilities depending on the detail of the study protocol. It may be useful to keep a list of those studies where the decision to reject was a marginal one, partly so that this list can be presented, together with the reason for rejection, in a supplementary file to the paper reporting the results of the meta-analysis, and partly so that results from such rejected papers may be included in sensitivity analyses.

## LITERATURE SEARCHING

As discussed elsewhere<sup>[33]</sup> it is necessary to make it absolutely clear exactly what the search criteria used are, so that others can repeat the searches, perhaps at a later date. Whether one limits attention to Medline searches, on the basis that they are quite comprehensive and free, or to studies published in English, to avoid the costs of translation, is up to the researcher. Especially where such restricted searches provide substantial numbers of relevant

studies, extending to other literature databases or studies in other languages may add little useful.

It is sometimes suggested that attention should be restricted to studies published in peer-reviewed journals. I disagree with this view for two reasons. Firstly, my personal experience suggests that peer-review is not necessarily a guarantee of quality. Second, it is the quality of the study that matters, so why should one necessarily reject results from a good study published in a journal which is not peer-reviewed?

Similar considerations apply to unpublished data. In my 50 yr as a practising epidemiologist/medical statistician I have accumulated and filed a number of unpublished reports. If they contain relevant data, why should I not use them? On some occasions, the reviewer may be able to add useful material to his review by conducting analyses on public databases. While the methods used will need to be clearly described, perhaps in a supplementary file to the publication presenting the results of the meta-analyses, there seems in principle to be no good reason to exclude such evidence.

## IDENTIFICATION OF STUDIES AND DOUBLE COUNTING

Once a set of suitable papers has been identified from the literature search it will be necessary to draw up a list of studies. Some papers will present results from multiple studies, which it is advisable to keep separate in data entry for proper assessment of between-study heterogeneity. More commonly results from some studies will be presented in multiple publications. If one publication clearly supersedes another (*e.g.*, reporting results from 20 rather than 10 year follow-up from a cohort study), the superseded publication can be omitted from the meta-analysis to avoid double-counting. However, if two publications present independent results (*e.g.*, for different sexes or age groups) then they should both be considered in the meta-analyses.

Complete avoidance of overlap may not be the most desirable solution. For example, a national study based on outcomes occurring in, say, 1990 may include some individuals also considered in a study in a smaller region based on outcomes in 1985 to 1995. Similarly one paper may publish results from a study involving cases in 2000 to 2005 while another may publish results from the same study involving cases in 2004 to 2008. In both examples, complete avoidance would require exclusion of one of the studies, whereas, given the minor overlap, it would seem acceptable to include both sets of results.

## ENTERING DATA

For complex meta-analysis projects, we have found it useful to have two linked databases, one containing the characteristics of each study and the other the detailed results, typically containing multiple records for each study.

The study database would include a single record per study and contain such information as the relevant publication(s), the sexes considered, the age range of the population, the location of the study and its timing and length of follow-up, the nature of the population studied, any study weaknesses, the definition of the outcome, the numbers of cases and of subjects, the types of controls and matching factors used in case-control studies, the confounding variables studied, and the availability of results for each index of exposure and outcome studied.

Each record on the other database would be linked to the relevant study and refer to a specific effect estimate, recording the comparison made and the results. This record would include such details as the outcome, the sex, details of the exposure considered (including the level of exposure for dose-related indices), the source of the effect estimate (e.g., source publication, with page or table number), the type of effect estimate (e.g., relative risk, hazard ratio or odds ratio for 1/0 outcomes, or means or medians for continuous outcomes), the method of derivation (see below) and the adjustment variables taken into account. It would also include the effect estimate itself and its 95%CI or standard deviation, and the numbers of exposed and unexposed cases and controls (or at risk). It is also advisable to look routinely for errors in reported results. Some years ago I described<sup>[35]</sup> some simple methods to do this for odds ratios, relative risks and CI, and used these methods to give some examples of seriously erroneous published data, which unless corrected could seriously distort the results of the meta-analyses.

It is also necessary to have a clear set of rules for identifying which effect estimates are to be entered from each study. Is it planned to enter estimates by sex, age or other stratifying variables, or only overall estimates? Are there types of estimate that should not be entered, such as those which are adjusted for symptoms of the disease of interest?

Consideration should also be given to how to handle incompletely reported results. Where studies simply report results as "non-significant", without providing an effect estimate, one at least should mention this in a paper reporting on a meta-analysis. Ideally, an attempt to obtain quantitative estimates from the author should be made.

In many cases the effect estimates can be taken directly from the source publication, but in other cases it will be necessary to calculate them from the material provided (or, if practicable, from raw data supplied by the author of the publication). Often the effect estimates can be calculated using standard methods<sup>[36]</sup>, but there is a situation I commonly come across, where more sophisticated techniques are required. This is where a study presents effect estimates and 95%CI for a range of different exposures (e.g., dose levels) relative to a specific exposure (e.g., unexposed), and one wishes to derive effect estimates and 95%CI for a different comparison (e.g., all exposed vs unexposed). Here the important thing to note is that the effect estimates and 95%CI are not independent, as they have a common base, so that the combined estimate cannot be derived by simple

meta-analysis of the individual estimates (as would be the situation given simple stratified data, e.g., by age). Fortunately a method to derive an appropriate combined estimate is available<sup>[37]</sup> and should be used. A method is also available<sup>[38]</sup> to derive estimates of the increase in effect per unit dose from such a table. Note that when deriving such estimates one will need a method to estimate the mean level of exposure from ranges, including open-ended intervals.

Most of the meta-analyses my colleagues and I have carried out over the years have been based on software we have written ourselves. Simple fixed-effect and random-effects meta-analysis can be programmed quite rapidly in Excel, the relevant methodology being succinctly described in the Appendix to a paper by Fleiss and Gross<sup>[39]</sup>. More commonly we use software incorporated into the ROEELE system developed by my colleague John Fry. While programming one's own software gives better insight into the methodology, John Fry advises me that 'meta for', the meta-analysis package for R, is a convenient one to use for those who do not wish to get so involved.

## STUDY QUALITY

While there are published methods for assessing study quality, such as the Cochran Collaboration Risk of Bias Tool and Effective Public Health Practice Project Quality Assessment Tool<sup>[40]</sup>, or the Newcastle-Ottawa Scale<sup>[41]</sup> which I have on occasion used, I have always been somewhat sceptical of them, because they seem to be trying to quantify what is essentially multi-dimensional into a single dimension. Even where study quality assessments are made, it is usually advisable to also carry out heterogeneity tests to see how effect estimates vary by those specific study characteristics which contribute to the assessment of quality.

## HETEROGENEITY TESTS

Where there are a reasonable number of independent effect estimates to be combined, analyses of heterogeneity should be conducted. If  $Q$  is Cochran's heterogeneity statistic, and  $df$  is the number of degrees of freedom (one less than the number of estimates combined), then heterogeneity is often expressed by the  $I^2$  statistic which is equal to  $100\% \times (Q - df) / Q$ . Negative values of  $I^2$  are set equal to zero, so that  $I^2$  lies in the range 0 to 100%, with values of 0% indicating no obvious heterogeneity, larger values indicating increased heterogeneity.

Apart from conducting standard fixed- effect and random-effects meta-analyses (see<sup>[39]</sup>), a systematic review should also include more detailed tests of heterogeneity, where  $Q$  is shown to be statistically significant (at  $P < 0.05$ ) and the number of estimates is sufficiently large (usually at least 10). These more detailed tests would involve separate fixed-effect meta-analyses for different levels of relevant study characteristic - such as sex, location, study type, definition of outcome, definition of exposure, number of confounding variables adjusted for, study size and

**Table 1** In a longitudinal study (often referred to as a prospective or cohort study) the data may be expressed

		Predictor variable		Total
		Exposed	Unexposed	
Outcome	Yes	A	B	A + B
	No	C	D	C + D
	Total	A + C	B + D	N

presence of a study weakness. These analyses serve two main purposes - first, to see whether an association seen in the overall meta-analysis is consistently seen in study subsets, and to see whether any factors are the cause of any heterogeneity seen. If a study characteristic has  $m$  levels ( $i = 1, \dots, m$ ) and if  $Q_i$  is the Cochran heterogeneity statistic for level  $i$ , then the statistic  $Q^* = Q - \sum_{i=1}^m Q_i$  is a test of heterogeneity between levels of the characteristic on  $m - 1$  degrees of freedom. If  $Q^*$  is close to its degrees of freedom, it implies that the study characteristic explains little or none of the heterogeneity. If, on the other hand, it is close to  $Q$ , it suggests that the characteristic is a major determinant of the heterogeneity. Where data permit it is useful to carry out meta-regression analyses in which a model is fitted simultaneously relating the effect estimate to a set of study characteristics. Because of correlation between characteristics, this should give greater insight into which are the important sources of heterogeneity and which are not. Variation in the effect estimate by levels of a study characteristic may arise for different reasons. For example, higher effect estimates in one location may be because of greater exposure to (or differing metabolism of) the exposure of interest by the population there. Or it may be due to differing biases in different situations. For example, higher effect estimates in case-control studies than in cohort studies may suggest that recall bias in case-control studies may be relevant, or for other reasons as described in the next section.

### Combining relative risks and odds ratios

Suppose we are studying the relationship of a predictor variable to an outcome, each with two levels. In a longitudinal study (often referred to as a prospective or cohort study) the data may be expressed as in Table 1.

The relationship of outcome to exposure is typically expressed by the relative risk (RR), the ratio of the probability of the outcome given exposure,  $A / (A + C)$ , to that given no exposure,  $B / (B + D)$ , or  $RR = A(B + D) / B(A + C)$ , the variance of its logarithm being given by  $1/A + 1/B - 1/(A + C) - 1/(B + D)$ .

In a cross-sectional or case-control study, the data may be similarly expressed, but here the relationship is typically expressed by the odds ratio (OR), the ratio of the odds of the outcome given exposure,  $A/C$ , to that given no exposure,  $B/D$  or  $OR = AD/BC$ , the variance of its logarithm being given by  $1/A + 1/B + 1/C + 1/D$ .

Where the outcome is relatively rare, it can be shown that RR and OR are very similar. Thus, for example, with  $A = 10$  and  $B = 20$ , and a true RR of 2, the OR will be 2.04

when comparing probabilities of 2% and 4%, and even closer to 2 for smaller probabilities. Even comparing 10% and 20% the OR of 2.25 is not that far from 2.

This suggests that when conducting meta-analysis of a reasonably rare outcome, one can combine RRs and ORs without worrying. Where this is not the case, e.g. when comparing 20% and 40% (where the OR is 2.67), this is less valid and it is preferable either to report separate combined results for ORs and RRs, or to try to convert one into the other. This is simple when the data are in the form of a  $2 \times 2$  table, but not possible for adjusted estimates without access to the raw data.

I note that often in longitudinal studies, where RRs are in principle more appropriate, ORs are often presented in publications. This is related to the simplicity of adjusting for multiple variables simultaneously using logistic regression analysis.

## ADJUSTMENT FOR CONFOUNDING VARIABLES

Especially where the association between the exposure and disease of interest is quite modest, one needs to bear in mind that the association may not be a causal one, and may be due to confounding by one or more variables which are correlated both with the exposure and the disease. Individual study authors are usually well aware of the problem and often present effect estimates adjusted for one or more sets of potential confounders. There are various approaches to investigate confounding in meta-analyses.

One possibility is to extract most-adjusted and least-adjusted effect estimates from each study. Most-adjusted estimates are those estimates reported in the source publication which have been adjusted for the most potential confounding variables, while least-adjusted estimates may include estimates that are totally unadjusted or adjusted only for age. Given these estimates, one can either compare results of meta-analyses based on the alternative estimates, or meta-analyse the ratio of estimates (perhaps using a weight based on the confidence limits of the most-adjusted estimates). Some studies may of course only provide one estimate, and can be excluded from such meta-analyses.

An additional method which may provide insight is to look for heterogeneity of the effect estimate according to the grouped number of confounders adjusted for, or to compare estimates adjusted or unadjusted for specific potential confounding variables.

Where an association substantially reduces following adjustment for confounding, but remains statistically significant, the possibility of bias arises.

Though beyond the scope of most meta-analyses, it is on some occasions worth formally investigating the extent to which effect estimates from meta-analyses may be biased by such uncontrolled confounding. The interested reader may wish to study the techniques used in our systematic review of the relation between environmental



tobacco smoke exposure and lung cancer<sup>[23]</sup> which concluded that bias due to uncontrolled confounding by four factors (fruit, vegetable and dietary fat consumption, and education) explains a substantial part of the observed association.

Another possibility to be borne in mind is “residual confounding”, arising because relevant confounders have not been adjusted for. It is well documented that “misclassification of a confounder” leads to “partial loss of ability to control confounding”<sup>[42]</sup> while “even misclassification rates as low as 10% can prevent adequate control of confounding”<sup>[43]</sup>. It has even been noted that if X is an inaccurately measured true cause of disease, and if Y, which is precisely measured but not a cause, is correlated with X, one may incorrectly conclude that Y, not X, is the cause (e.g.<sup>[44-46]</sup>).

## MISCLASSIFICATION

Apart from bias arising due to misclassification of confounding variables, bias may also arise because of other forms of misclassification. Random misclassification of the exposure or outcome variable will tend to dilute any relationship, but misclassification may not be random, and can lead to underestimation of the relationship. For example, when studying a relatively weak association of smoking to cancer at one site, the inclusion of some individuals who actually have cancer of a site known to be strongly related to smoking (such as lung cancer) will bias upward the association being studied. Misdiagnosis of lung cancer certainly exists<sup>[47-49]</sup>. Similarly, upward bias will arise if some of those classified as having the exposure of interest actually have an exposure which is more strongly related to the disease.

While random misclassification of exposure or outcome should not produce an association when no true causal relationship exists, this is certainly not so for random misclassification of the variable used to determine whether an individual should be included in the study. This applies, for example, to the study of the relationship of spousal smoking to lung cancer in never smokers. As I have demonstrated<sup>[50,51]</sup>, the inclusion of some true ever smokers among the reported never smokers, can cause bias. This bias arises because spouses tend to have smoking habits in common, so that the exposed group (with spouses who smoke) are likely to include more misclassified smokers than will the comparison group (with spouses who do not smoke). Because of the very high risk of lung cancer this bias can be substantial, and the interested reader may wish to study the techniques which my colleagues and I used to adjust for misclassification bias<sup>[23]</sup>.

## PUBLICATION BIAS

Publication bias occurs if the published data are not representative of all the data that exist on a topic. It is well documented (e.g.<sup>[52,53]</sup>) that positive findings are published more often than negative findings, so meta-analyses of data drawn from the literature tend to overestimate true

relationships. Inasmuch as large studies are more likely than small studies to publish their findings regardless of the result, one can compare effect estimates from larger and smaller studies as some sort of test of publication bias. More formal tests are available, but tend to involve assumptions that are difficult to justify. Furthermore, they are based on the published results, and ignore what may be known about unpublished results. What should one conclude if a very large cohort study has published evidence demonstrating a statistically significant relationship between an exposure and various common diseases, but has not reported results relating that exposure to other common diseases? It seems to me quite likely that the authors would have looked at these other diseases, found no significant association, and decided not to publish their findings. The existence of such studies should at least be pointed out in the discussion section of a paper describing a meta-analysis of the exposure to one of these other diseases.

Publication bias can also arise in the meta-analysis of dose-response relationships. It is certainly plausible that authors will be more likely to report dose-response results where there is a strong association in the first place. This can be tested by comparing effect estimates for overall exposure in studies reporting and not reporting dose-response results.

## STATISTICAL SIGNIFICANCE

An effect estimate derived from a meta-analysis that is not statistically significant ( $P > 0.1$ ) clearly cannot be interpreted as supporting a true causal relationship. Nor can it rule it out, as one cannot prove a negative, but it can suggest an upper limit to any true effect. Additional studies may clarify the situation, especially where the original meta-analysis had little power, being based on relatively few studies.

On the other side of the coin, a significant association alone does not demonstrate that a true causal effect exists.  $P$ -values less than 0.05 but greater than 0.1 may be due to chance, and even where the probability is very low, so that chance can be excluded for practical purposes, confounding or bias may be relevant. Before concluding that a causal effect is likely, it is up to the meta-analyst to demonstrate that confounding or bias cannot explain the relationship, which may be difficult, especially where the relationship is weak.

## CONCLUSION

Meta-analysis is an interesting subject and quite difficult to do well. If it is done well it can act as an extremely useful tool to aid the epidemiologist in reaching a conclusion. However, it is very important for the meta-analyst to be aware of the limitations of meta-analysis, and of the epidemiological studies on which it is based.

## REFERENCES

- 1 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF.

- Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999; **354**: 1896-1900 [PMID: 10584742 DOI: 10.1016/S0140-6736(99)04149-5]
- 2 **Forey BA**, Thornton AJ, Lee PN. Systematic review with meta-analysis of the epidemiological evidence relating smoking to COPD, chronic bronchitis and emphysema. *BMC Pulm Med* 2011; **11**: 36 [PMID: 21672193 DOI: 10.1186/1471-2466-11-36]
  - 3 **Lee PN**, Forey BA, Coombs KJ. Systematic review with meta-analysis of the epidemiological evidence in the 1900s relating smoking to lung cancer. *BMC Cancer* 2012; **12**: 385 [PMID: 22943444 DOI: 10.1186/1471-2407-12-385]
  - 4 **Fry JS**, Lee PN, Forey BA and Coombs KJ. Dose-response relationship of lung cancer to amount smoked, duration and age starting. *World J Meta-Anal* 2013; **1**: 57-77 [DOI: 10.13105/wjma.v1.i2.57]
  - 5 **Lee PN**, Forey BA, Thornton AJ and Coombs KJ. The relationship of cigarette smoking in Japan to lung cancer, COPD, ischemic heart disease and stroke: a systematic review [version 1; referees: awaiting peer review]. *F1000Research* 2018; **7**: 204 [DOI: 10.12688/f1000research.14002.1]
  - 6 **Lee PN**. Lung cancer and type of cigarette smoked. *Inhal Toxicol* 2001; **13**: 951-976 [PMID: 11696868 DOI: 10.1080/089583701753210353]
  - 7 **Lee PN**. Systematic review of the epidemiological evidence comparing lung cancer risk in smokers of mentholated and unmentholated cigarettes. *BMC Pulm Med* 2011; **11**: 18 [PMID: 21501470 DOI: 10.1186/1471-2466-11-18]
  - 8 **Lee PN**. Tar level of cigarettes smoked and risk of smoking-related diseases. *Inhal Toxicol* 2018; **30**: 5-18 [PMID: 29488428 DOI: 10.1080/08958378.2018.1443174]
  - 9 **Lee PN**, Fry JS, Hamling JS. Using the negative exponential distribution to quantitatively review the evidence on how rapidly the excess risk of ischaemic heart disease declines following quitting smoking. *Regul Toxicol Pharmacol* 2012; **64**: 51-67 [PMID: 22728684 DOI: 10.1016/j.yrtph.2012.06.009]
  - 10 **Fry JS**, Lee PN, Forey BA, Coombs KJ. How rapidly does the excess risk of lung cancer decline following quitting smoking? A quantitative review using the negative exponential model. *Regul Toxicol Pharmacol* 2013; **67**: 13-26 [PMID: 23764305 DOI: 10.1016/j.yrtph.2013.06.001]
  - 11 **Lee PN**, Fry JS, Thornton AJ. Estimating the decline in excess risk of cerebrovascular disease following quitting smoking--a systematic review based on the negative exponential model. *Regul Toxicol Pharmacol* 2014; **68**: 85-95 [PMID: 24291341 DOI: 10.1016/j.yrtph.2013.11.013]
  - 12 **Lee PN**, Fry JS, Forey BA. Estimating the decline in excess risk of chronic obstructive pulmonary disease following quitting smoking - a systematic review based on the negative exponential model. *Regul Toxicol Pharmacol* 2014; **68**: 231-239 [PMID: 24361344 DOI: 10.1016/j.yrtph.2013.12.006]
  - 13 **Lee PN**. Circulatory disease and smokeless tobacco in Western populations: a review of the evidence. *Int J Epidemiol* 2007; **36**: 789-804 [PMID: 17591642 DOI: 10.1093/ije/dym039]
  - 14 **Weitekunt R**, Sanders E, Lee PN. Meta-analysis of the relation between European and American smokeless tobacco and oral cancer. *BMC Public Health* 2007; **7**: 334 [PMID: 18005437 DOI: 10.1186/1471-2458-7-334]
  - 15 **Lee PN**, Hamling J. Systematic review of the relation between smokeless tobacco and cancer in Europe and North America. *BMC Med* 2009; **7**: 36 [PMID: 19638245 DOI: 10.1186/1741-7015-7-36]
  - 16 **Lee PN**. Summary of the epidemiological evidence relating snus to health. *Regul Toxicol Pharmacol* 2011; **59**: 197-214 [PMID: 21163315 DOI: 10.1016/j.yrtph.2010.12.002]
  - 17 **Lee PN**. Epidemiological evidence relating snus to health--an updated review based on recent publications. *Harm Reduct J* 2013; **10**: 36 [PMID: 24314326 DOI: 10.1186/1477-7517-10-36]
  - 18 **Lee PN**, Thornton AJ. The relationship of snus use to diabetes and allied conditions. *Regul Toxicol Pharmacol* 2017; **91**: 86-92 [PMID: 29061372 DOI: 10.1016/j.yrtph.2017.10.017]
  - 19 **Lee PN**, Fariss MW. A systematic review of possible serious adverse health effects of nicotine replacement therapy. *Arch Toxicol* 2017; **91**: 1565-1594 [PMID: 27699443 DOI: 10.1007/s00204-016-1856-y]
  - 20 **Thornton AJ**, Lee PN. Parental smoking and risk of childhood cancer: a review of the evidence. *Indoor Built Environ* 1998; **7**: 65-86 [DOI: 10.1159/000024566]
  - 21 **Thornton AJ**, Lee PN. Parental smoking and sudden infant death syndrome: a review of the evidence. *Indoor Built Environ* 1998; **7**: 87-97 [DOI: 10.1159/000024567]
  - 22 **Thornton AJ**, Lee PN. Parental smoking and middle ear disease in children: a review of the evidence. *Indoor Built Environ* 1999; **8**: 21-39 [DOI: 10.1159/000024607]
  - 23 **Lee PN**, Fry JS, Forey B, Hamling JS and Thornton AJ. Environmental tobacco smoke exposure and lung cancer: a systematic review. *World J Meta-Anal* 2016; **4**: 10-43 [DOI: 10.13105/wjma.v4.i2.10]
  - 24 **Lee PN**, Thornton AJ, Hamling JS. Epidemiological evidence on environmental tobacco smoke and cancers other than lung or breast. *Regul Toxicol Pharmacol* 2016; **80**: 134-163 [PMID: 27321059 DOI: 10.1016/j.yrtph.2016.06.012]
  - 25 **Lee PN**, Hamling JS. Environmental tobacco smoke exposure and risk of breast cancer in nonsmoking women. An updated review and meta-analysis. *Inhal Toxicol* 2016; **28**: 431-454 [PMID: 27541291 DOI: 10.1080/08958378.2016.1210701]
  - 26 **Lee PN**, Thornton AJ, Forey BA, Hamling JS. Environmental Tobacco Smoke Exposure and Risk of Stroke in Never Smokers: An Updated Review with Meta-Analysis. *J Stroke Cerebrovasc Dis* 2017; **26**: 204-216 [PMID: 27765554 DOI: 10.1016/j.jstrokecerebrovasdis.2016.09.011]
  - 27 **Lee PN**, Forey BA, Hamling JS and Thornton AJ. Environmental tobacco smoke exposure and heart disease: A systematic review. *World J Meta-Anal* 2017; **5**: 14-40 [DOI: 10.13105/wjma.v5.i2.14]
  - 28 **Lee PN**, Forey BA, Coombs KJ, Hamling JS and Thornton AJ. Epidemiological evidence relating environmental smoke to COPD in lifelong non-smokers: a systematic review [version 1; referees: awaiting peer review]. *F1000Research* 2018; **7**: [DOI: 10.12688/f1000research.13887.1]
  - 29 **Lee PN**, Fry JS. Systematic review of the evidence relating FEV1 decline to giving up smoking. *BMC Med* 2010; **8**: 84 [PMID: 21156048 DOI: 10.1186/1741-7015-8-84]
  - 30 **Fry JS**, Hamling JS, Lee PN. Systematic review with meta-analysis of the epidemiological evidence relating FEV1 decline to lung cancer risk. *BMC Cancer* 2012; **12**: 498 [PMID: 23101666 DOI: 10.1186/1471-2407-12-498]
  - 31 **Forey BA**, Fry JS, Lee PN, Thornton AJ, Coombs KJ. The effect of quitting smoking on HDL-cholesterol - a review based on within-subject changes. *Biomark Res* 2013; **1**: 26 [PMID: 24252691 DOI: 10.1186/2050-7771-1-26]
  - 32 **Stroup DF**, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB. 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 [PMID: 10789670 DOI: 10.1001/jama.283.15.2008]
  - 33 **Liberati A**, Altman DG, Tetzlaff J, Mulrow C, Götzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700 [PMID: 19622552 DOI: 10.1136/bmj.b2700]
  - 34 **Shamseer L**, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, Shekelle P, Stewart LA; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ* 2015; **350**: g7647 [PMID: 25555855 DOI: 10.1136/bmj.g7647]
  - 35 **Lee PN**. Simple methods for checking for possible errors in reported odds ratios, relative risks and confidence intervals. *Stat Med* 1999; **18**: 1973-1981 [PMID: 10440880 DOI: 10.1002/(SICI)1097-0258(19990815)18:15<1973::AID-SIM166>3.0.CO;2-G]
  - 36 **Gardner MJ**, Altman DG, editors. Statistics with confidence: Confidence intervals and statistical guidelines. London: British Medical Journal 1989; 140

- 37 **Hamling J**, Lee P, Weitkunat R, Ambühl M. Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category. *Stat Med* 2008; **27**: 954-970 [PMID: 17676579 DOI: 10.1002/sim.3013]
- 38 **Berlin JA**, Longnecker MP, Greenland S. Meta-analysis of epidemiologic dose-response data. *Epidemiology* 1993; **4**: 218-228 [PMID: 8512986 DOI: 10.1097/00001648-199305000-00005]
- 39 **Fleiss JL**, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. *J Clin Epidemiol* 1991; **44**: 127-139 [PMID: 1995774 DOI: 10.1016/0895-4356(91)90261-7]
- 40 **Armijo-Olivo S**, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012; **18**: 12-18 [PMID: 20698919 DOI: 10.1111/j.1365-2753.2010.01516.x]
- 41 **Wells GA**, Shea B, O'Connell D, Peterson J, Welch V, Losos M and Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: URL: [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp?status=print](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp?status=print)
- 42 **Greenland S**. The effect of misclassification in the presence of covariates. *Am J Epidemiol* 1980; **112**: 564-569 [PMID: 7424903 DOI: 10.1093/oxfordjournals.aje.a113025]
- 43 **Tzonou A**, Kaldor J, Smith PG, Day NE, Trichopoulos D. Misclassification in case-control studies with two dichotomous risk factors. *Rev Epidemiol Sante Publique* 1986; **34**: 10-17 [PMID: 3726220]
- 44 **Greenland S**, Robins JM. Confounding and misclassification. *Am J Epidemiol* 1985; **122**: 495-506 [PMID: 4025298 DOI: 10.1093/oxfordjournals.aje.a114131]
- 45 **Savitz DA**, Barón AE. Estimating and correcting for confounder misclassification. *Am J Epidemiol* 1989; **129**: 1062-1071 [PMID: 2705426 DOI: 10.1093/oxfordjournals.aje.a115210]
- 46 **Fewell Z**, Davey Smith G, Sterne JA. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study. *Am J Epidemiol* 2007; **166**: 646-655 [PMID: 17615092 DOI: 10.1093/aje/kwm165]
- 47 **Lee PN**. Comparison of autopsy, clinical and death certificate diagnosis with particular reference to lung cancer. A review of the published data. *APMIS Suppl* 1994; **45**: 1-42 [PMID: 7811528]
- 48 **Faccini JM**. The role of histopathology in the evaluation of risk of lung cancer from environmental tobacco smoke. *Exp Pathol* 1989; **37**: 177-180 [PMID: 2637150 DOI: 10.1016/S0232-1513(89)80042-8]
- 49 **Sterling TD**, Rosenbaum WL, Weinkam JJ. Bias in the attribution of lung cancer as cause of death and its possible consequences for calculating smoking-related risks. *Epidemiology* 1992; **3**: 11-16 [PMID: 1554804 DOI: 10.1097/00001648-199201000-00004]
- 50 **Lee PN**, Forey BA. Misclassification of smoking habits as a source of bias in the study of environmental tobacco smoke and lung cancer. *Stat Med* 1996; **15**: 581-605 [PMID: 8731002 DOI: 10.1002/(SICI)1097-0258(19960330)15:6<581::AID-SIM182>3.0.CO;2-B]
- 51 **Lee PN**, Forey BA and Fry JS. Revisiting the association between environmental tobacco smoke exposure and lung cancer risk. III. Adjustment for the biasing effect of misclassification of smoking habits. *Indoor Built Environ* 2001; **10**: 384-398 [DOI: 10.1177/1420326X0101000605]
- 52 **Sterling TD**. Publication decisions and their possible effects on inferences drawn from tests of significance - or vice versa. *J Am Stat Assoc* 1959; **54**: 30-34 [DOI: 10.2307/2282137]
- 53 **Easterbrook PJ**, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991; **337**: 867-872 [PMID: 1672966 DOI: 10.1016/0140-6736(91)90201-Y]

**P- Reviewer:** Roy PK, Velasco I **S- Editor:** Dou Y  
**L- Editor:** A **E- Editor:** Wu YXJ





Published by **Baishideng Publishing Group Inc**  
7901 Stoneridge Drive, Suite 501, Pleasanton, CA 94588, USA  
Telephone: +1-925-223-8242  
Fax: +1-925-223-8243  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <http://www.f6publishing.com/helpdesk>  
<http://www.wjgnet.com>

