CLINICAL EPIDEMIOLOGY WORKSHOP

Workshop 11 — Sources of bias in systematic reviews with or without meta-analysis

A systematic review is an attempt to summarise information from all available studies conducted on a certain topic or area, and has been highly regarded as the first stop when retrieving information from the literature to address a defined question. It is a literature review that tries to identify, appraise, and synthesise all available evidence relevant to that question using a systematic approach, and hence should be repeatable if the same methods of search and information synthesis are followed. A systematic review starts with formulating a clinical or public health question, usually in the format of PICO (Problem/Population, Intervention [or exposure], Comparison, Outcomes),[1] followed by identifying individual studies based on predefined inclusion and exclusion criteria using an organised searching strategy. The relevant information from individual studies is then retrieved, and simply described or summarised. When appropriate, findings of individual studies can be quantitatively combined using the statistical method termed meta-analysis. This is especially useful in situations where individual studies are not adequately powered to provide statistically significant results or conflicting results are being reported by different primary studies. After appraising the individual studies, if they are regarded as heterogeneous in quality and/or in terms of other characteristics (eg study design), sometimes a stratified meta-analysis is performed.

Systematic reviews can be conducted in all the four major areas of clinical activities (diagnosis, therapy, prognosis, and harm/causation). Clinicians are often advised to place systematic reviews and meta-analyses at the top level of hierarchical evidence when making decisions about clinical interventions in the practice of evidence-based medicine. There is broad agreement that systematic reviews provide clinicians with up-to-date summaries that minimise their efforts on locating and reading individual studies and hence save time and effort in looking for evidence to support clinical decisions. However, systematic reviews are not without dispute and may vary greatly in quality. Similar to primary studies,[2-5] systematic reviews are also susceptible to all three common sources of biases, although most meta-analyses explicitly address only publication bias (a form of selection bias). In a way, a systematic review can be likened to a case series (Table) by collecting and summarising information from individuals (primary studies).

Selection bias

As in a case series, a systematic review consisting of all primary studies, published and unpublished, fulfilling predetermined inclusion and exclusion criteria should be representative of all primary study findings available at around the time the review is conducted. However, if only published studies with language constraint (eg English) are utilised from certain electronic databases, as is the case with many published meta-analyses, a large number of eligible studies may be missed out resulting in substantial selection bias. An inappropriate or non-comprehensive search strategy and including only publications with full text are other common errors that can reduce the representativeness of the identified or included studies. A comprehensive search should include the 'grey literature'.[6] On the other hand, self-selection bias may arise when researchers choose not to publish some of their primary studies for various reasons, eg as a result of null or unexpected or unexplained findings. Moreover, journal reviewers and editors are more likely to accept studies with statistically significant results, particularly if they ensue even when the sample size is relatively small. Such self-selection bias and selection bias taken together constitute the publication bias, which is frequently examined in systematic reviews and meta-analyses using the "funnel plot".[7] Not infrequently, authors of systematic reviews mention contacting authors and researchers in the related field to obtain information on studies that are not published or not included in the databases used in the search. However, the response rate and other details related to such enquiries are seldom reported. A low response rate to such queries can also introduce self-selection bias.

Information bias

The process of retrieving information from individual studies for inclusion in systematic reviews is generally more problematic than in case series, as the data to be collected or abstracted from the primary studies are more prone to misclassification than demographic and clinical data for patients in a case series. Better objectivity can be achieved if a standard form for abstracting relevant information with clear definitions and categorisations of various variables is adopted, and the information retrievers are blinded to the research question of the systematic review. Exposure/intervention, outcomes, and potential confounding factors may not be defined or categorised in the same way across individual primary studies, and re-categorisation for the purpose of performing a meta-analyses can result in misclassification. To

TABLE. Characteristics and sources of bias—comparing systematic review and meta-analysis to case series

	Case series	Systematic review (narrative)	Meta-analysis
Nature and study unit	A series of individual patients (cases)	A series of individual primary studies	A series of individual primary studies
Source and representativeness	Patients of individual clinics or hospitals; representative only of the study setting	Various databases; purports to have comprehensive coverage of all relevant studies, but depends very much on search strategy	Various databases; purports to have comprehensive coverage of all relevant studies, but depends very much on search strategy
Selection bias—by investigator	+	++	++
Selection bias—by subject	+	++	++
Information bias—exposure/ intervention	+	Depending on nature of primary studies[5]	++
Information bias—outcome	+	Depending on nature of primary studies[5]	++
Information bias—confounding	NA	Depending on nature of primary studies[5]	++
Information bias—associations	NA	Depending on nature of primary studies[5]	++
Confounding	NA	NA (with no quantitative summary)	++ (non-comparability among primary studies)

* + denotes minor source, ++ major source, and NA not applicable

minimise the potential for misclassification, it is not uncommon in a meta-analysis to have two reviewers independently abstract information from the primary studies, and resolve any disagreement by recourse to a third reviewer or through consensus. However, this process is more to ensure reliability than validity, and is not infrequently dominated by more senior or experienced reviewers. Furthermore, as a systematic review or meta-analysis utilises information presented in the primary studies, any measurement or reporting errors in the latter studies can invalidate the inferences of the review.

The most important piece of information used in a meta-analysis is the measurement of the association or effect size in its primary studies. Different effect measures (eg relative risk, odds ratio, standardised mortality ratio) may have been used in the primary studies, and may not be correctly translated into the single effect measure adopted for a meta-analysis. Furthermore, wrong or invalid measurements of association due to the presence of different types of bias (selection, information, confounding) in the primary studies can invalidate the results of a meta-analysis and cannot be adjusted or compensated for. Hence, appraising the quality of individual studies to be included in a systematic review is of utmost importance. Only studies with reasonably valid results of associations should be included. Relevant guidelines for appraising primary studies have already been discussed in previous workshops in this series.[8-11] It is prudent to note that such appraisals will not be adequate if done mechanically using simple scores or scales, and are best conducted by experienced researchers in the particular area who are also competent in critical appraisal.

Confounding

Provided high-quality randomised controlled trials are the main source of information, confounding is usually not an issue for narrative systematic reviews, as the comparison groups in the primary studies should be very similar with respect to prognostic factors other than the intervention.[2] However, when a quantitative summary is contemplated in a meta-analysis, heterogeneity or non-comparability of the exposures/interventions, outcomes and study subjects (eg age, disease stage, co-morbidities, co-interventions) in the primary studies could interfere with the validity or interpretation of the final summary effect measures. Regrettably, such non-comparability cannot be easily adjusted or compensated for. Hence, adopting more stringent inclusion/exclusion criteria to improve homogeneity across primary studies should help, but at the expense of external generalisation and applicability. Recourse to stratified analyses may improve comparability and/or homogeneity. The issue of non-comparability is even more complicated with meta-analyses that include observational studies or non-randomised trials in which the adjusted confounding factors can be very different, in terms of numbers, definitions, and categorisations, in the different primary studies. A narrative summary with in-depth critical appraisals of individual primary studies may be more informative.

We have to agree that systematic reviews represent an improvement, because the methods for searching the relevant literature and synthesising information are being described explicitly; traditional reviews could have been done haphazardly and not infrequently by purposively identifying primary studies in support of the a priori conclusion of the

authors. However, the quality of the output from a meta-analysis is determined by the quality of input (the raw materials of the primary studies), which is based on the analogy "garbage in, garbage out". Although frequently hailed as the gold standard or placed on the top of the hierarchy of evidence in evidence-based medicine, one must understand that the validity of results from a meta-analysis can be no better than that of the lowest quality study it includes. This fact may be hard to swallow for many 'meta-analysts', but clinicians should be made aware of this potential fallacy of the 'gold standard'.

Shelly LA Tse
Ignatius TS Yu
Clinical Epidemiology Group
Hong Kong Medical Journal

References

1. Richardson WS, Wilson MC, Nishikawa J, Hayward RS. The well-built clinical question: a key to evidence-based decisions. ACP J Club 1995;123:A12-3.
2. Yu IT, Tse SL. Clinical Epidemiology Workshop 3—Sources of bias in case series, patient cohorts, and randomised controlled trials. Hong Kong Med J 2011;17:478-9.
3. Yu IT, Tse SL. Clinical Epidemiology Workshop 4—Sources of bias in case-referent studies. Hong Kong Med J 2012;18:46-7.
4. Yu IT, Tse SL. Clinical Epidemiology Workshop 5—Sources of bias in cohort studies. Hong Kong Med J 2012;18:137-8.
5. Yu IT, Tse SL. Clinical Epidemiology Workshop 6—Sources of bias in cross-sectional studies; summary on sources of bias for different study designs. Hong Kong Med J 2012;18:226-7.
6. Fourth International Conference on Grey Literature: New Frontiers in Grey Literature. GreyNet, Grey Literature Network Service, Washington DC, USA, 4-5 October 1999.
7. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.
8. Yu IT, Tse SL. Clinical Epidemiology Workshop 7—Appraising a study on diagnosis. Hong Kong Med J 2012;18:325-6.
9. Yu IT, Tse SL. Clinical Epidemiology Workshop 8—Appraising a study on therapy. Hong Kong Med J 2012;18:419-20.
10. Yu IT, Tse SL. Clinical Epidemiology Workshop 9—Appraising a study on prognosis. Hong Kong Med J 2012;18:507-8.
11. Yu IT, Tse SL. Clinical Epidemiology Workshop 10—Appraising a study on risk factors or aetiology. Hong Kong Med J 2013;19:50-1.