CLINICAL EPIDEMIOLOGY WORKSHOP

Workshop 12 — Appraising a systematic review with meta-analysis

As mentioned in the previous workshop,[1] a systematic review is not infrequently used as the first stop when retrieving information from the literature to address a defined question. Hence, acquiring the skills for appraising a systematic review appears essential for evidence-based medical practice. Regrettably, the less straight-forward statistical techniques often used in meta-analyses can be intimidating to the average clinician, with the end result that either everything presented in a meta-analysis is trusted as truth (as advised and advocated by some 'authorities') or that the meta-analysis is simply neglected. The latter is liable to occur if the results contradict the personal experience of the clinician, especially when the meta-analysis is prepared by non-clinicians. Neither is desirable in the practice of evidence-based medicine.

Admittedly, appraising a systematic review with a meta-analysis is no easy task. First, one must have acquired the skills for appraising individual primary studies. The purposes of the critical appraisal are common, that is to ensure that the results arising from a meta-analysis are clinically important and can be applied in one's own clinical setting. Of course, before contemplating applying the results, one must ensure that they are important, valid, and reasonably precise. In this Workshop, a similar approach will be taken to addressing the four major questions in appraising a systematic review with meta-analysis.

(1) Are the results of clinical and/or public health importance?

As systematic reviews can be conducted in all the four major areas of clinical activities (diagnosis, therapy, prognosis, and harm/causation), the specific questions on clinical and/or public health importance may be worded differently.[2-5]

(2) Are the results basically valid?

The very nature of a systematic review with meta-analysis dictates that its results may be heavily subject to different sources of bias, including those present in the primary studies incorporated into the meta-analysis. The specific questions to be answered for ascertaining the validity of study results in a systematic review are structured under the three major sources of bias (Box).[6] Furthermore, heterogeneity among

BOX. Validity of study results[6]

Validity — selection bias

- Was the source of primary studies described, as well as their inclusion and exclusion criteria, and were important sources of relevant primary studies liable to be missing, eg non-English publications?
- Was the final collection of studies just an easily accessible or conveniently available collection from electronic sources on the internet only?
- If researchers in the specific field were approached, was the response rate reported and was it reasonably high?
- Was there a funnel plot (or other equivalent measure) to support the absence of publication bias? (Note: absence of publication bias suggested by a symmetrical funnel plot does not per se constitute evidence for the absence of selection or self-selection bias [see the above points]).

Validity — measurement/information bias (including misclassification)

- Were exposures/interventions similarly defined in the primary studies? If not, were the categorisations/classifications used in the meta-analysis defined a priori and clinically rational?
- Were the outcomes similarly defined in the primary studies? If not, were the categorisations/classifications used in the meta-analysis defined a priori and clinically rational?
- Was a standard form or protocol (with clear definitions and categorisations of various variables) adopted for abstracting relevant information?
- Were the information retrievers blinded to the research question of the review?
- Was double retrieval of information implemented (two reviewers independently retrieving the same set of data from the primary studies)? And if so, how was disagreement resolved?
- Were the effect measures (eg relative risk) the same across the primary studies? If not, would the estimation of the finally adopted effect measure from the different effect measures in the primary studies result in misclassification?
- Was validity of the results (freedom from biases arising from selection, information, and confounding) in the primary studies properly assessed?
- Were results from studies with doubtful validity included in the meta-analysis?

Validity — confounding and heterogeneity

- Were the backgrounds (eg demographic, clinical) of the subjects similar in the primary studies?
- Were the influences of differences in backgrounds of subjects on meta-analysis outcome(s) analysed?
- Were the same potential confounding factors being adjusted in the primary studies?
- Were there any attempts to identify sources of heterogeneity (eg study design) if present?
- Was stratified analysis carried out when heterogeneity (not necessarily identifiable by homogeneity statistics) was noted in the primary studies? (One should not count apples as oranges though they are both fruits.)

primary studies can create problems in summarising results.

(3) Are the results reasonably reliable?

Meta-analyses are usually performed with easily accessible software on the web and the 95% confidence intervals (CIs) of the summary effect measures are automatically generated. One of the benefits of meta-analysis is that the summary effect measure usually has greater precision than those of its individual studies, but one must still examine whether the 95% CI overlaps with the null value of difference between the groups being compared. Overlapping suggests that no conclusion can be drawn regarding the association and could have happened either when there is no actual association between the exposure/treatment and the outcome (with a narrow enough CI), or as happens not infrequently the meta-analysis is conducted prematurely and has a wide 95% CI.

(4) Can the results be applied in another setting?

It goes without too much emphasis that the demographic and/or clinical backgrounds of subjects should be considered. Heterogeneity of effects among subjects with different background, if identified, should provide guidance for actual applicability in different settings. The utility of results also depends on the background risk of the disease. This is pertinent to estimating the positive/negative predictive value for a screening/diagnosis test, and the number needed to treat for therapy.[2,3] The prevalence of a specific exposure (in estimating preventable disease burdens) in the population for applying the results,[5] as well as the incremental benefit over the incremental cost and harm for any contemplated intervention, would also affect utility.

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

References

1. Tse SL, Yu IT. Clinical Epidemiology Workshop 11—Sources of bias in studies of systematic reviews with or without meta-analysis. Hong Kong Med J 2013;19:156-8.
2. Yu IT, Tse SL. Clinical Epidemiology Workshop 7—Appraising a study on diagnosis. Hong Kong Med J 2012;18:325-6.
3. Yu IT, Tse SL. Clinical Epidemiology Workshop 8—Appraising a study on therapy. Hong Kong Med J 2012;18:419-20.
4. Yu IT, Tse SL. Clinical Epidemiology Workshop 9—Appraising a study on prognosis. Hong Kong Med J 2012;18:507-8.
5. Yu IT, Tse SL. Clinical Epidemiology Workshop 10—Appraising a study on risk factors or aetiology. Hong Kong Med J 2013;19:50-1.
6. Yu IT, Tse SL. Clinical Epidemiology Workshop 2—General approach to critical appraisal of a medical journal paper. Hong Kong Med J 2011;17:405-6.