Systematic Reviews and Meta-Analyses

Lindsay S. Uman PhD[1]

With an ever-increasing plethora of studies being published in the health sciences, it is challenging if not impossible for busy clinicians and researchers alike to keep up with the literature. Reviews summarizing the outcomes of various intervention trials are therefore an extremely efficient method for obtaining the "bottom line" about what works and what doesn't.

Key Terms Defined

Systematic reviews differ from traditional narrative reviews in several ways. Narrative reviews tend to be mainly descriptive, do not involve a systematic search of the literature, and thereby often focus on a subset of studies in an area chosen based on availability or author selection. Thus narrative reviews while informative, can often include an element of selection bias. They can also be confusing at times, particularly if similar studies have diverging results and conclusions. Systematic reviews, as the name implies, typically involve a detailed and comprehensive plan and search strategy derived a priori, with the goal of reducing bias by identifying, appraising, and synthesizing all relevant studies on a particular topic. Often, systematic reviews include a meta-analysis component which involves using statistical techniques to synthesize the data from several studies into a single quantitative estimate or summary effect size (Petticrew & Roberts, 2006). In contrast to traditional hypothesis testing which can give us information about statistical significance (i.e., did the intervention group differ from the control group) but not necessarily clinical significance (i.e., was this difference clinically meaningful or large), effect sizes measure the strength of the relationship between two variables, thereby providing information about the magnitude of the intervention effect (i.e., small, medium, or large). The type of effect size calculated generally depends on the type of outcome and intervention being examined as well as the data available from the published trials; however, some common examples include odds ratios (OR), weighted/standardized mean differences (WMD, SMD), and relative risk or risk ratios (RR). Although systematic reviews are published in academic forums, there are also organizations and databases specifically developed to promote and disseminate them. For example, the Cochrane Collaboration (www.cochrane.org) is a widely recognized and respected international and not-for-profit organization that promotes, supports, and disseminates systematic reviews and meta-analyses on the efficacy of interventions in the health care field.

8 Stages of a Systematic Review and Meta Analysis

1. Formulate the review question.

The first stage involves defining the review question, forming hypotheses, and developing a review title. It is often best to keep titles as short and descriptive as possible, by using the following formula: *Intervention* for *population* with *condition* (e.g., Dialectical behavior therapy for adolescent females with borderline personality disorder). Reviews published with the Cochrane Collaboration do not need to be identified as such, but reviews published in other sources should also indicate in the title that they represent a systematic review and/or meta-analysis. If authors chose to conduct their review through the Cochrane Collaboration, they will also be required to register their title to the appropriate review group, which in essence "saves their spot" for this topic and provides access to further Cochrane support (e.g. , assistance running search strategies).

2. Define inclusion and exclusion criteria.

The Cochrane acronym PICO (or PICOC) which stands for population, intervention, comparison, outcomes (and context) can be useful to ensure that one decides on all key components prior to starting the review. For example, authors need to decide a priori on their population age range, conditions, outcomes, and type(s) of interventions and control

[1]Clinical Psychologist, IWK Health Centre, Halifax, Nova Scotia

Corresponding email: lindsay.uman@iwk.nshealth.ca

groups. It is also critical to operationally define what types of studies to include and exclude (e.g., randomized controlled trials-RCTs only, RCTs and quasi-experimental designs, qualitative research), the minimum number of participants in each group, published versus unpublished studies, and language restrictions. For Cochrane Reviews, this information gets prepared, peer-reviewed, and published in a Protocol format first, which is then replaced with the full Review once it is completed.

3. Develop search strategy and locate studies.

This is the stage where a reference librarian can be extremely helpful in terms of helping to develop and run electronic searches. Generally, it is important to come up with a comprehensive list of key terms (i.e., "MeSH" terms) related to each component of PICOC to be able to identify all relevant trials in an area. For example, if the age range is 13-18 year old females, search terms may need to include any of the following: adolescents, teenagers, youth, female, women, girls, etc. The key in developing an optimal search strategy is to balance sensitivity (retrieving a high proportion of relevant studies) with specificity (retrieving a low proportion of irrelevant studies). Searches generally include several relevant electronic databases but can also include checking article reference lists, hand-searching key journals, posting requests on listservs, and personal communication with experts or key researchers in the field.

4. Select studies.

Once a comprehensive list of abstracts has been retrieved and reviewed, any studies appearing to meet inclusion criteria would then be obtained and reviewed in full. This process of review is generally done by at least two reviewers to establish inter-rater reliability. It is recommended that authors keep a log of all reviewed studies with reasons for inclusion or exclusion, and it may be necessary to contract study authors to obtain missing information needed for data pooling (e.g., means, standard deviations). Translations may also be required.

5. Extract data.

It can be helpful to create and use a simple data extraction form or table to organize the information extracted from each reviewed study (e.g., authors, publication year, number of participants, age range, study design, outcomes, included/excluded). Data extraction by at least two reviewers is important again for establishing inter-rater reliability and avoiding data entry errors.

6. Assess study quality.

There has been a movement in recent years to better assess the quality of each RCT included in systematic reviews. Although there are brief check-lists available such as the 5-point Oxford Quality Rating Scale (Jadad et al., 1996) commonly used in Cochrane reviews, this measure is heavily influenced by double-blinding which is appropriate for drug trials but generally not for psychological or non-pharmacological interventions. There are other more comprehensive recommended guidelines and standards available such as the Consolidated Standards of Reporting Trials (CONSORT Statement; http://www.consort-statement.org/), as well as articles providing recommendations for improving quality in RCTs and meta-analyses for psychological interventions (e.g., Uman et al., 2010).

7. Analyze and interpret results.

There are various statistical programs available to calculate effects sizes for meta-analyses, such as the Review Manager (RevMan) program endorsed by the Cochrane Collaboration. Effect sizes are stated along with a 95 % confidence interval (CI) range, and presented in both quantitative format and graphical representation (e.g., forest plots). Forest plots visually depict each trial as a horizontal diamond shape with the middle representing the effect size (e.g., SMD) and the end points representing both ends of the CI. These diamonds are presented on a graph with a centre line representing the zero mark. Often the left side of the graph (< zero) represents the side favoring treatment, while the right side (> zero) represents the side favouring the control condition. At the bottom of the graph is a summary effect size or diamond representing all of the individual studies pooled together. Ideally, we would like to see this entire diamond (effect size and both anchors of the CI) falling below zero, indicating that the intervention is favoured over the control. In addition, most programs also calculate a heterogeneity value to indicate whether the individual studies are similar enough to compare. In this case, it is preferable to have non-significant findings for heterogeneity. It is still possible to pool studies when significant heterogeneity exists, although these results should be interpreted with caution or reasons for the heterogeneity should be explored. As with all papers, the last step in the writing process involves summarize the findings, and providing recommendations for clinical work (e.g., which interventions are efficacious, for whom, and under what conditions) and research (e.g., what areas/topics/interventions require further research.

8. Disseminate findings.

Although reviews conducted through the Cochrane Collaboration get published in the online Cochrane Database of Systematic Reviews, they are often quite lengthy and detailed. Thus, it is also possible and encouraged to publish abbreviated versions of the review in other relevant academic journals, as long as they are clearly indicated as such (e.g., Uman et al., 2008). Plain language summaries for families and patients are also commonly provided, and there is an expectation that reviews should be regularly updated to ensure they are always up-to-date and relevant. Indeed, participating in a review update or joining a well-established review team, can be a helpful way of getting involved in the systematic review process.

References

Jadad, A. R, Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials, 17*, 1-12.

Petticrew, M. & Roberts, H. (year). *Systematic reviews in the social sciences: A practical guide.* Malden, MA: Blackwell Publishing.

Uman, L.S., Chambers, C. T., McGrath, P. J., & Kisely, S. (2010). Assessing the quality of randomized controlled trials examining psychological interventions for pediatric procedural pain: Recommendations for quality improvement. *Journal of Pediatric Psychology, 35*, 693-703.

Uman, L. S., Chambers, C. T., McGrath, P. J., & Kisely, S. (2008). A systematic review of randomized controlled trials examining psychological interventions for needle-related procedural pain and distress in children and adolescents: An abbreviated Cochrane review. *Journal of Pediatric Psychology, 33*, 842-854.