Practical Clinical Procedures
A USER’S GUIDE TO THE CHIROPRACTIC AND
OSTEOPATHIC LITERATURE
How to Read a Systematic Review

Simon French  B.App.Sc.(Chiropractic), M.P.H.  *

INTRODUCTION

What is the best way of keeping up with the literature? Systematic reviews are an efficient means of keeping up with health care literature offering the practitioner an easier way of making evidence-based decisions. It is estimated that there are approximately 30,000 journals published worldwide. For any health practitioner to keep up with this information would be impossible. Just to read, let alone critically appraise such a vast amount of information would not allow any time to see patients!

WHY SYSTEMATIC REVIEWS

A systematic review is considered a summary of the health care literature using explicit methods to guide systematic searches and critical appraisal1. Results from studies are often conflicting, the disparity of these results being difficult to assess. This is especially true in the manual therapy literature, with different clinical trials showing different results. Health care is riddled with controversies and reviews are often seen as the end of an argument. However, reviews that disregard the systematic process may have biases that skew the results of the review. For example, the conclusions drawn may be due more to the reviewer’s specialty rather than the available data2.

SYSTEMATIC VERSUS NARRATIVE REVIEW

The systematic review differs considerably from the traditional ‘narrative review’. Narrative reviews are generally written by an author who has an interest in a particular area of health care. He/she browses through journals and textbooks, searching for information that supports their argument. Anything that is contrary to this argument is omitted. Selective inclusion of studies that support the reviewer’s point of view is all too common. The process of writing the narrative review does not require guidance to any formal rules.

Today, many health care review articles are still written in the narrative or journalistic style, selectively choosing studies that support the author’s argument. This has been demonstrated by Knipschild and colleagues who describe how a Nobel Prize winning biochemist used selective quotes from the medical literature to ‘prove’ his theory that vitamin C helps you live longer and feel better3,4. When Knipschild and his colleagues searched the literature systematically, they also found that 1 or 2 trials did strongly suggest that vitamin C could prevent the onset of the common cold. However, there were far more studies that did not show any beneficial effect. A systematic approach was able to find evidence both for and against a hypothesis and then this evidence was assessed to determine what the current recommendation should be.

HOW TO PERFORM A SYSTEMATIC REVIEW

A systematic review is a standardised and objective procedure. Like any research, a systematic review contains an explicit statement of objectives, materials, and methods and has been conducted according to explicit and reproducible methodology. Like any good research procedure, if different investigators in a different setting conduct the review then the same results should occur.

One of the qualities of a good systematic review is that it assesses the quality of the primary research, and demonstrates that there is good quality and bad quality research. As for primary research, systematic reviews can be published that are of poor quality. For the discerning practitioner to rely upon a review’s conclusions as evidence for or against a particular procedure, the reader must be able to assess the quality of that review based on whether it has followed the appropriate systematic process.

The Cochrane Collaboration is an international effort to prepare, maintain, continuously update and disseminate systematic reviews of randomised controlled trials of the effects of health care. The Cochrane Collaboration has created a standard format for systematic reviews which is being adopted worldwide.
EVALUATING THE QUALITY OF SYSTEMATIC REVIEWS

This section outlines what qualities make up a good systematic review. This is summarised in Table 1. Each of the listed qualities must be satisfied in order to determine the overall value of a particular review. I have also used an example of a recently published systematic review to determine whether this particular review has achieved these qualities. This article searched for recently published studies on adverse events after cervical spine manipulation. You can obtain this article from the website of the Medical Journal of Australia (www.mja.com.au) for your own comparison and critique.

Table 1: Checklist for Interpreting Review Articles

- Are the results valid?
- Did the review explicitly address a sensible clinical question?
- Was the search for relevant studies detailed and exhaustive?
- Were the primary studies of high methodological quality?
- Were assessments of studies reproducible?
- What are the results?
- What are the overall results of the review?
- How precise were the results?
- How can I apply the results to patient care?
- How can I best interpret the results to apply them to the care of patients in my practice?
- Were all clinically important outcomes considered?
- Are the benefits worth the costs and potential risks?

Table 2: Checklist of Data Sources for a Systematic Review (Adapted from Greenhalgh)

Importantly, the ‘Search Strategy’ undertaken by the reviewer, clearly outlining where the studies were sourced from should be included. If another reviewer then carries out the same search the same literature should be identified.

The Ernst article searched 5 databases, including MEDLINE, EMBASE, the Cochrane Library, AMED (Allied and complementary medicine database) and CISCOM (Centralised Information Service for Complementary Medicine). This search strategy also included consultation with nine “experts”, whose credentials were not given, and the searching of the references of located articles. This search strategy is comprehensive and would expect to locate all relevant articles for the chosen objective. Other databases that could have been utilised include the CINAHL and Mantis databases. However, I expect that the databases utilised would cover all journals that these databases also list.

HOW WERE STUDIES INCLUDED?

In a systematic review, you can equate the studies found in the search strategy to subjects in a research study. Specific inclusion and exclusion criteria should then be applied to these studies. For example, study designs that will be included, populations to be studied, interventions, etc. Two independent reviewers should then assess each study for their inclusion or exclusion for the review. The reviewers should be blinded, information such as the identification of the author and the publishing journal should be removed so that the reviewers are not influenced by these when applying the inclusion criteria.

The Ernst article’s study selection criteria included all case reports containing original data relating to serious
adverse events associated with spinal manipulation. There was only one reviewer who chose articles for inclusion, potentially exposing this review to bias. The reviewer may have had a predetermined view on adverse events associated with spinal manipulation. If a study identified by the search did not agree with this view then it might have been excluded. For example, an article is identified in the literature search that reveals an adverse event after a medical practitioner performed a spinal manipulation. This article may be excluded if the reviewer believes that all adverse events from spinal manipulation are caused by chiropractors. The use of two reviewers would limit the possibility of this type of bias occurring.

**ASSESSING METHODOLOGICAL QUALITY OF INCLUDED STUDIES**

One of the tasks of a systematic reviewer is to draw up a list of criteria against which each study should be judged. Each study may then be assigned a score based on the strength of each aspect of its methodology. However, care should be taken in developing such scores since there is no gold standard for the ‘true’ methodological quality of a study and composite quality scores are often neither valid nor reliable in practice.

This section of the review can be very useful from the reader’s perspective. Authors will commonly draw up a Table that directly compares each methodological aspect of the studies, so that a quick glance can identify the strengths and weaknesses of the primary studies.

The Ernst study does not include a quality assessment of each of the articles identified. Case studies are considered the weakest in the hierarchy of epidemiological study types. They are subject to considerable bias and in particular a cause and effect relationship is not possible to determine from such a study. The author of this review assumes that each of the adverse events outlined in the case studies is cause and effect, and in fact states that “the acute onset of symptoms after the manipulation made a causal relationship likely”.

**HOW WAS THE DATA COMBINED?**

A common way to review the results in a narrative review is to count the number of studies supporting either side of an issue and choose the view receiving the most votes. This is clearly unsound because it ignores research design, sample size and effect size, amongst other differences between the studies included in the review.

A meta-analysis is a statistical pooling technique which combines the results of all the studies included in the review. The meta-analysis may be performed after the systematic review procedure has been undertaken. The meta-analysis allows for a number of smaller studies to be combined so that essentially one big study is undertaken. This area of statistical pooling is not a perfect science and caution must be observed when interpreting reviews using this process. In the low back pain therapy literature, for example, a meta-analysis has not been performed to date because of the number of low quality trials that have been performed.

Part of the reviewer’s job is also to tabulate relevant information on the inclusion criteria, sample size, baseline patient characteristics, withdrawal rate and results of primary and secondary end points of all the studies included. This information in table form is particularly useful for the reader because at a glance one can get an overview of all studies conducted and their various pros and cons. These tables can save the reader considerable time by avoiding the ploughing through of the methods sections of each paper.

The Ernst study includes a table which outlines all of the major details of each of the 31 case reports that were identified by the search strategy. This includes patient characteristics and clinical indication for spinal manipulation, the type of therapist, the adverse event, the diagnostic studies utilised to determine the nature of the adverse event and the outcome for the patient. The data is not combined in any meaningful way, which in this case is appropriate considering the diversity in treatment methods and adverse events.

**HOW SENSITIVE ARE THE RESULTS TO THE WAY THE REVIEW HAS BEEN DONE?**

It has been demonstrated that a completely untenable therapy can have perceived benefit by changing certain aspects in a systematic review. Counsel and colleagues demonstrated an entirely spurious relationship between the result of shaking a dice and the outcome of an acute stroke. They reported a series of artificial dice rolling experiments in which red, white, and green dice represented different therapies for acute stroke. Overall, the ‘trials’ conducted with the dice showed no significant benefit from the 3 therapies. However, the simulation of a number of perfectly plausible events in the process of a meta-analysis, such as exclusions of some of the studies on the grounds of ‘methodological quality’, led to an apparently highly significant benefit of ‘dice therapy’ in acute stroke.

If these simulated results pertained to a genuine health care controversy, how would it be possible for the reader to detect these subtle biases? The reader needs to conduct a ‘Sensitivity Analysis’, to examine the ‘what ifs’.

The reader needs to ask themselves some of the following questions. What if the authors of the systematic review had changed their inclusion criteria? What if their methodological quality scores had been assigned differently? What if trials of lower methodological quality
had been included (or excluded)? What if drop out patients in the studies were treated differently, ie assumed that they all got better, or worse?

If the reader finds that relatively small changes to the data in various ways makes little or no difference to the review’s overall results, you can assume that the review’s conclusions are relatively robust. If, however, the key findings disappear when any of the ‘what ifs’ changes, the conclusions should be expressed far more cautiously and you should hesitate before changing your practice in the light of them. The discerning reader takes a sensitivity analysis into account before accepting the conclusions on face value.

**BIAS IN SYSTEMATIC REVIEWS**

Bias is a systematic error that distorts the review. Bias can be estimated by looking at the method and conduct of the review. There are a number of instances where bias can result in the formulation of a systematic review (Table 3).

| Bias results from:               |
|---------------------------------|
| Failure to locate all relevant studies. |
| Failure to determine inclusion criteria. |
| Inclusion of poor quality studies. |
| Failure to correctly extract data. |
| Selective inclusion of favourable results. |
| Excluding studies with missing data. |

Table 3. Bias in Systematic Reviews

Failure to locate all relevant studies can occur with an inadequate search strategy. If all appropriate databases are not utilised then important studies may be missed. Failure to locate all relevant studies can also be due to publication bias. Publication bias occurs if the reason a study is not published is related to its outcome. For example, a clinical trial is undertaken assessing the value of an intervention and the study does not show a favourable outcome. The investigators decide not to publish the trial and the results are never released to the wider community.

In the Ernst article publication bias is particularly evident, as it is highly likely that adverse events due to spinal manipulation are under-reported. Such a review that Ernst has performed cannot expect to find all adverse events due to spinal manipulation. The collection of more comprehensive data would require a large cohort study, following many individuals who are exposed to spinal manipulation and determining if there are adverse events.

**ARE THE CONCLUSIONS OF THE REVIEW VALID?**

The conclusion of the Ernst article is that serious complications of cervical spine manipulations appear to occur regularly. As outlined above, this conclusion is not supported by the data presented in the review, and would seem to be more a reflection of the author’s opinion than by the review that has been undertaken. The second conclusion is that the incidence of adverse events due to spinal manipulation is unknown and should be established as a matter of urgency through adequately designed investigations. This conclusion is supported by the review and ironically suggests that case reports are not an appropriate means of making conclusions.

**CONCLUSION**

Systematic reviews offer the busy clinician an easy and effective means of interpreting a wide range of current literature. The discerning reader must exercise caution so as not to accept the conclusions of the review on face value, and a checklist of critical appraisal should be applied to each review to determine if the results are valid.

**REFERENCES**

1. Sackett DL, Strauss SE, Richardson WS, et al. Evidence-Based Medicine: How to Practice and Teach EBM. Churchill Livingstone, Edinburgh, Scotland 2000., 2nd Ed.
2. Assendelft WJJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. JAMA 1995; 274:1942-1948.
3. Chalmers I, Altman DG, eds. Systematic reviews. BMJ Publishing Group, London 1995.
4. Pauling L. How to live longer and feel better. Freeman, New York, 1986.
5. Ernst E. Manipulation of the cervical spine: a systematic review of case reports of serious adverse events, MJA 1995–2001; 176(8):376-380.
6. Greenhalgh T. How to read a paper: papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997; 315(7109):672-675.
7. Emerson JD, Burdick E, Hoaglin DC, Mosteller F, Chalmers TC. An empirical study of the possible relation of treatment differences to quality scores in controlled randomized clinical trials. Controlled Clin Trials 1990; 11:339-352.
8. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials: current issues and future directions. Int J Health Technol Assess 1996; 12:195-208.
9. Assendelft WJJ, Koes BW, van der Heijden GJMG, Bouter LM. The effectiveness of chiropractic for treatment of low back pain: an update and attempt at statistical pooling. J Manip Physiol Ther 1996; 19(8):499-507.
10. Counsell CE, Clarke MJ, Slattery J, Sandercock PAG. The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? BMJ 1994; 309:1677-1681.