CEIPE THE CONSORTIUM OF EVIDENCE-INFORMED PRACTICE EDUCATORS

The Savvy Practitioner

A bulletin for practitioners and teachers of evidence-based practice.

Target audience this issue:

- ✓ *Classroom faculty*
- ✓ *Clinicians*
- ✓ *Faculty in general*
- ✓ *EIP core instructors*

David Newell, PhD
Anglo-European College of Chiropractic

Editor: Ron LeFebvre, DC
Executive Director CEIPE
University of Western States

6/11/16
Issue 4

Critical Appraisal of Meta-Analysis: What does heterogeneity mean?

The great advantage of a meta-analysis is that in addition to a narrative summary of the present state of knowledge concerning a particular clinical question it can also provide a numerical estimate of the overall effect of a particular treatment or intervention. This is based on pooling the numerical findings from a number of studies that have asked the same or similar research questions.

In reviews such as this, the size of the effects from each of the randomized controlled trials (RCTs) is essentially "averaged" by combining the effects found in each of the single studies into one overall effect. This gives an estimate of the effect of the treatment as if it were generated from a large RCT that contained all the subjects from the individual RCTs. So, for example, if there were 10 RCTs each with 50 subjects, a meta-analysis can give the estimated effect size as if an RCT of 500 subjects had been carried out. This can increase the statistical power by combining small RCTs into one large RCT, thus reducing the possibility of finding no effect when indeed there really is an effect (Type II error).

This sounds like a great solution to the multitude of small RCTs out there that maybe lack the power to find an effect amongst all that statistical noise. BUT there are a number of caveats that need addressing before such a meta-analysis can be relied upon. These include the efficiency of the search strategy, the rigour of appraisal of the quality of included studies, and a number of others that you can find on critical appraisal instruments like the one you can access by clicking here.

One of these issues is the subject of this Savvy Practitioner and is called *heterogeneity*. The explanation goes something like this.

Despite the fact that a number of RCTs collected in a systematic review may report an effect size (for example a reduction in pain) and it may be easy to pool these together mathematically the question remains, should we? For example, I could be interested in the average pain reduction as associated with chiropractic care for low back pain. I collect several studies that have the following characteristics:

Help students recognize when not to trust the results from a meta-analysis.

Want to know more? Want handouts for your students?

Consult the Educator's Exchange!

Click through the following webpages: **EIP Resources > Pre-appraised Literature > Systematic Reviews** and download more materials on meta-analyses.

Lost your link to the Educator's Exchange?

Try http://bit.ly/CEIPE.

You will need your password and user name.

Don't have access to the Educator's Exchange?

To sign up for this closed website, just contact rlefebvre@uws.edu and you will be sent an invitation to set a username and password.

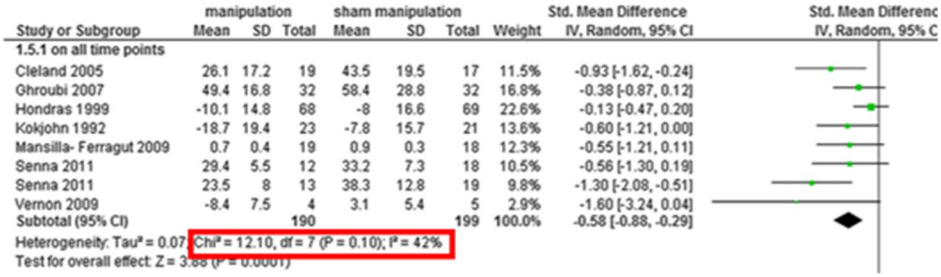
1. *Study sample is males with very chronic LBP in London*
2. *Study recruited females with very mild upper thoracic pain in Nairobi*
3. *Subjects were children with non-specific LBP visiting private chiropractic practices in southern USA*
4. *Hospital patients in Singapore post back fusion surgery undergoing chiropractic care and intense rehabilitation*

What do you notice? Of course you would expect these studies to give very different results in terms of reduction in pain and the context of their settings and subjects. Would you be happy if I simply add the effects found by these studies together and then report this as a pooled result representing the effect of chiropractic care on back pain?

I'm hoping you would say no! Why is this so?

Well obviously the results of these studies, i.e. the effects that they report, are not really comparable. They have too much *clinical heterogeneity*. Quite different sets of subjects have been investigated along with a range of different settings and severity of conditions. If a meta-analysis had pooled these studies together to produce one effect, then you might have less confidence that it really represents the truth because each of the studies is so different that it makes no sense just to add the results together. But even in this example, there may be other potential threats to pooling that include *methodological* and *statistical heterogeneity!*

Methodological heterogeneity refers to differences in how the studies were conducted, potential errors, and the types of care rendered. *Statistical heterogeneity* refers to the amount of variation in the results of smaller studies included in the review and whether it is beyond what would be expected by chance. Too much variation in any of these factors may threaten the validity of the pooling. Heterogeneity can be tested by a Chi ^{2} statistic, where a p value< 0.1 indicates significant heterogeneity and the pooling should not have probably been carried out. This type of heterogeneity should be reported in a meta-analysis normally within the forest plot (see below) or in the text. Another common measure of heterogeneity is something called I^2, a percentage that ranges from 0% to 100% with anything above 75% indicates SEVERE HETEROGENIETY; i.e. the studies were very different from each other and either there should be no pooling of data or the overall pooled result should be interpreted with extreme caution. In the case below the heterogeneity figures are acceptable (Chi^2 p=0.1, I^2 =42%) and the studies appear to have been similar enough to pool the effects.



From: Scholten-Peeters, G.G. et al., 2013. Is manipulative therapy more effective than sham manipulation in adults? a systematic review and meta-analysis. Chiropractic & Manual Therapies, 21(1), pp.1–1.