ResearchOnline@ND

University of Notre Dame Australia
ResearchOnline@ND

Health Sciences Papers and Journal Articles

School of Health Sciences

2009

Interpretive bias in acupuncture research? A case study

Neil E. O'Connell

Benedict M. Wand
University of Notre Dame Australia, Benedict.Wand@nd.edu.au

Ben Goldacre

Follow this and additional works at: http://researchonline.nd.edu.au/health_article

Part of the Alternative and Complementary Medicine Commons

This article was originally published as:
O'Connell, N. E., Wand, B. M., & Goldacre, B. (2009). Interpretive bias in acupuncture research? A case study. *Evaluation and the Health Professions, 32* (4), 393-409.
http://doi.org/10.1177/0163278709353394

This article is posted on ResearchOnline@ND at
http://researchonline.nd.edu.au/health_article/23. For more information,
please contact researchonline@nd.edu.au.

Running Head: Interpretive bias in acupuncture research

Interpretive bias in acupuncture research? A case study.

Neil E O'Connell
Lecturer
Centre for Research in Rehabilitation,
School of Health Sciences and Social Care
Brunel University,
West London, United Kingdom.

Benedict M Wand
Associate Professor
School of Health Sciences,
The University of Notre Dame Australia,
Fremantle, WA, Australia.

Ben Goldacre
Research Fellow
Dept of Psychological Medicine
King's College London
United Kingdom

Correspondence to neil.oconnell@brunel.ac.uk

Abstract

Acupuncture is one of the most widely used and broadly researched of the complementary and alternative therapies, but high quality trials generally show no benefit over sham acupuncture. Many would view this result as evidence of ineffectiveness for this intervention.

This discussion paper focuses on the report of one large multi-centre randomised controlled trial of acupuncture for chronic low back pain in the lay and academic press, the ensuing discussion, and its impact on both clinical practice and service provision.

We suggest that interpretive bias has affected reporting, leading to questionable conclusions and advocacy in favour of this form of care which may exceed the evidence. We also suggest that a lack of understanding of research into the placebo effect may have contributed to confusion in the interpretation of these trials.

Keywords

Acupuncture, clinical trials, interpretive bias, back pain.

Main Body

Introduction

Acupuncture is one of the most commonly used of the complementary and alternative medicine (CAM) therapies and has gained a level of acceptance in conventional medicine that other CAM therapies have failed to achieve. Acupuncture has also been the subject of an extensive research effort to assess its clinical efficacy for a variety of conditions. A review of all acupuncture related Cochrane reviews concludes that acupuncture fails to demonstrate efficacy for a wide range of conditions but does demonstrate some efficacy in nausea and headache (Ernst, 2008a). A broader systematic review of systematic reviews of acupuncture (Derry et al., 2006) concluded that when controlling for methodological and interpretive limitations, no robust evidence could be found that acupuncture is effective for any indication. Most recently a review of acupuncture for the treatment of pain concluded that the analgesic effects of acupuncture are small, clinically insignificant and cannot be clearly distinguished from bias (Madsen et al., 2009). Despite these results acupuncture for low back pain has recently received government approval in Germany (Haake et al., 2007), has been endorsed in UK clinical guidelines (NICE 2009) and there have been calls for its inclusion in the European guidelines for the management of chronic low back pain (Yuan et al., 2008).

The clear and consistent message from high quality research is that acupuncture offers little, if anything, beyond a placebo response, yet a

frequent interpretation of this outcome is that acupuncture is a useful intervention. The reasons why an intervention with such an unimpressive evidence base might be promoted rigorously represent an important public health issue. While the explanations are undoubtedly numerous, one key phenomenon seems to be interpretive bias.

Kaptchuk presents the concept of interpretive bias as an umbrella term for a number of mechanisms through which the results of research may be distorted at the stage of interpreting data, rather than collecting it. It can affect both the protagonists and the prospective audience of research, at any stage from data analysis to the appraisal of the results (for review see Kaptchuk, 2003). Interpretive bias is a common phenomenon throughout medicine and has been demonstrated, for example, in the discussion sections of industry funded meta-analyses, when contrasted with comparable independent meta-analyses (Jørgensen et al., 2006).

This paper uses the example of one recent large randomised controlled trial on acupuncture to illustrate the issue of interpretive bias, and describe how it may affect both the reporting and political impact of clinical trials of acupuncture.

The GERAC back pain trial

The recent GERAC trial of acupuncture for chronic low back pain (CLBP) (Haake et al., 2007) received significant international media attention (e.g.

Fleming, 2007, Johnson, 2007, Lawrence, 2007, BBC, 2007, Reuters, 2007, Williams, 2007). This well designed study was a three arm trial comparing genuine “verum” acupuncture against sham acupuncture and conventional therapy. The verum acupuncture group received acupuncture to points according to the principles of traditional Chinese medicine (TCM). The sham acupuncture group received a course of superficial needling at non-acupuncture points without needle manipulation and without the elicitation of strong sensation. The conventional therapy group received standard treatment according to German clinical guidelines including medication and various forms of physical therapy. The trial found that both true and sham acupuncture led to significantly greater improvements than did conventional therapy, but the outcomes from sham and true acupuncture were not significantly different.

We and others have previously criticised the GERAC back pain trial as a failure of interpretation. (Goldacre, 2007; Ernst, 2008b; Wand & O’Connell, 2008). In our view the most appropriate interpretation of the equivalence in outcome between the verum and sham conditions is that acupuncture was not effective beyond placebo. The study’s authors, however, conclude in the original article (Haake et al., 2007) and in subsequent commentary (Endres et al., 2008a, Endres et al., 2008b) that acupuncture is an effective and useful treatment option for low back pain above and beyond placebo. In defence of this position they have offered a number of arguments, some of which are being increasingly used by acupuncture researchers when interpreting the outcomes of clinical trials. We review each of these below.

"*The lack of significant differences between the results for the two forms of acupuncture should not be taken as proof that either, or both, were solely due to a placebo effect. If one were to do so one would have to conclude that the results of the conventional treatment were also attributable solely to placebo because acupuncture was significantly more effective than conventional treatment*" (Endres et al. 2008a).

The train of logic here is hard to follow, but seems to rest on a belief that the placebo effect is a unitary phenomenon, and that other factors may not come to bear on clinical outcomes. In reality factors such as natural history, statistical regression and the Hawthorne effect will have contributed to the improvements seen in all three groups.

Furthermore, there is a large literature on the variations in the magnitude of effect between different forms of placebo. The examples are too numerous to review in detail here, (see Moerman, 2002 for review): salt water injections have been shown to be more effective for pain relief than inert pills; four placebo pills a day have been shown to be more effective for clearing gastric ulcers than two; and so on. Of particular relevance, Kaptchuk et al. (2006a) recently compared two different placebo treatments for arm pain, one of which was a sugar pill, and one of which was a 'ritual', a fake treatment modelled on acupuncture: the more elaborate placebo ritual had a greater benefit than the simple pill. The same team have also demonstrated that the effect of placebo

acupuncture can be enhanced incrementally through manipulation of the therapeutic interaction (Kaptchuk et al., 2008).

Placebo research emphasises the importance of positive expectation in mediating the placebo response (Vase et al 2002) and the power of positive expectations in acupuncture research is well established. In a post hoc analysis of four controlled acupuncture trials, Linde et al. (2007) demonstrated that in a large and varied cohort, prior expectations of acupuncture were generally very positive in the majority of subjects. The results showed that patients' prior expectations of treatment were significantly associated with clinical outcomes in both the sham and verum acupuncture groups. Similarly Bausell et al. (2005) demonstrated that positive expectations of acupuncture were more strongly related to outcome than the treatment condition (verum or sham) for pain relief following dental surgery. Kalauokalani et al. (2001) also demonstrated a strong influence of prior expectations of treatment in a comparative trial of acupuncture and massage for low back pain. Together these results suggest that expectations of acupuncture are likely to have impacted on the results of the study, tending to produce a positive result for verum and "sham", or "placebo", acupuncture. The GERAC authors acknowledge the importance of expectation in mediating the outcome of acupuncture treatment. While the exact procedure is not entirely clear from the original manuscript (Haake et al 2007) or the pre-trial publications (Molsberger et al 2006, Haake et al. 2003) it appears that they stratified for patient's expectations of acupuncture within their randomisation process. While this ensures balance between the verum and sham acupuncture groups

it does not account for the probable low expectations of conventional treatment. The baseline data appear to confirm that all groups did indeed have very high expectations that acupuncture would be beneficial (7.6-7.7/10) and we would suggest that this is likely to have disadvantaged conventional treatment.

There are additional concerns around the choice of conventional medicine control, and the kind of patients recruited, which may have had a significant impact on outcome. The therapies included in the conventional treatment condition have themselves been shown to demonstrate poor efficacy for chronic low back pain, if any at all (Assendelft et al., 2004, Clarke et al., 2007, French et al., 2006, Guzmán et al., 2002, Hayden et al., 2005), and some are not recommended in recent European guidelines for the management of chronic LBP (Airaksinen et al., 2006). Furthermore, the study compared a novel treatment (acupuncture) to a form of care (conventional treatment) that had previously failed that cohort of patients over a period of several years. The apparent superiority of acupuncture over conventional care may well have arisen from the comparison of a failed treatment - with a potential nocebo effect - against a novel treatment known to be associated with particularly strong placebo effects (Kaptchuk et al., 2006a). It is therefore entirely feasible that a convincing placebo treatment might outperform these therapies.

"To minimise potential nocebo effects in those assigned to the conventional treatment group participants were informed before randomisation that they would receive 10 acupuncture sessions after the completion of the study regardless of assignment and were offered the alternative of participating in an observational study examining the long-term effects of acupuncture" (Endres et al. 2008a)

To promise acupuncture at the end of the trial implies that this is the most likely active therapy: this may reinforce favourable expectations and increase the negative expectations associated with conventional care, as discussed above. The GERAC authors suggest that this approach would ensure that patients with strong positive expectations of acupuncture and negative expectations of conventional therapy could opt out (Endres et al., 2008a, Haake et al., 2007). It is arguable whether such steps would have effectively managed this problem; in fact they may have had the opposite effect, making subjects with negative expectations of conventional therapy *more* likely to participate, and to continue to participate in the conventional care arm of the trial, reassured that they would receive acupuncture afterwards. This would enhance a placebo response for sham and verum acupuncture in comparison with conventional medicine, rather than control for it.

Comparison with a companion GERAC trial further supports this interpretation. In the GERAC migraine trial it appears that no offer of acupuncture at the completion of the trial was made to subjects (Diener et al., 2006). In this trial 13% of recruited subjects withdrew their consent on being

informed of group assignment and 85% of those withdrew from the group allocated to conventional care (Diener et al., 2006). It is worth noting how dramatic the differences in drop-out rates between each group were: 308 subjects were randomised to standard treatment, and 106 (34%) withdrew consent after learning of their assignment; 652 subjects were randomised to receive acupuncture, and only 19 subjects (3%) withdrew. In the low back pain study, prior to randomisation, only 3.5% of eligible subjects withdrew their consent, and it seems no subject withdrew their consent after learning of their group allocation (Haake et al., 2007). Assuming that the proportion of recruited patients with low expectation of conventional treatment was reasonably similar between these two studies, it would appear that a higher proportion of individuals with a negative expectation of conventional therapy are likely to have been included in the back pain trial by the adoption of these strategies.

Also relevant is the finding that even in trials where patients do not have a strong view on the efficacy of each treatment, inadequate blinding has been associated with significant overestimation of treatment efficacy (Schultz et al., 1995, Wood et al., 2008).

"Physiotherapists in this study worked with patients for much longer periods than acupuncturists" (Endres et al., 2008a)

Endres et al. appear to suggest that this factor may have controlled for some of the confounding influences of treatment expectation and other non-specific

treatment effects. It is likely that treatment expectation is influenced more by the context and meaning of the treatment to the patient rather than the absolute time spent with the patient. We are not aware of any data indicating that the simple duration of face to face contact between therapist and patient has a strong influence on placebo outcomes. As already noted, acupuncture was a novel treatment for all of the study participants and expectations of it were high. 8 minutes of the ritual of careful application of acupuncture treatment followed by 20 minutes rest with the needles in situ, carrying with it the hope and expectation of benefit may well elicit a more powerful placebo response than 30 minutes of treatment that the patient has previously been exposed to without lasting benefit. There is indirect evidence that the novelty of acupuncture may be of importance. Cherkin et al. (2009) found that acupuncture (real or simulated non-penetrating acupuncture) performed better to usual care for chronic back pain in a study where all subjects were acupuncture-naïve. The same research group had previously demonstrated no superiority of acupuncture over an education programme for the same condition, but in that study a prior experience of acupuncture did not exclude participants (Cherkin et al. 2001). In this trial where the playing field was arguably more level, therapeutic massage was found to be superior to acupuncture.

"*Even placebo effects are associated with real biochemical processes in the brain, which blur the distinction between a real treatment with specific action mechanisms and a placebo treatment without such specific mechanisms.*" (Endres et al., 2008a)

Here the authors seem to question the relevance of distinguishing placebo responses from specific responses to a treatment. Placebo treatments do produce real effects, and that is why it is desirable to control for these effects in experimental design: to determine whether the treatment itself is causing a beneficial effect. Many everyday experiences can cause changes in brain biochemistry, this is a “trivially true” observation; and we also agree that treatment ritual appears to be associated with demonstrable physiological changes (Wager 2005), and significant therapeutic benefits (Vase et al., 2002). This is not evidence of efficacy for acupuncture: it is evidence of efficacy for the placebo effect.

In a separate discussion paper, one GERAC author appears more amenable to the possibility that the clinical effects of acupuncture may be mediated by placebo mechanisms (Endres, 2008). Following a discussion of the neurophysiologic correlates of placebo he wonders “whether such biological mechanisms that explain placebo-induced acupuncture (sic) should not properly be classified as specific effects.” Similarly Lundeberg et al. (2007) suggest that engagement of higher brain self-appraisal and reward systems may underpin the effects of acupuncture. Such mechanisms are implicated in any therapeutic interaction and cannot be claimed as a feature specific to acupuncture; again they are a benefit of (and argument for) the placebo effect.

There are wider issues around the choice of sham control in this specific study. The authors state in the pre-trial protocol for the GERAC back trial that for “ethical and logistical reasons” it was decided that the sham would not be a purely suggestive placebo and instead represents “minimal acupuncture” (Molsberger et al., 2006), which significantly muddies the waters. If the authors believed *a priori* that the sham group may be a plausible therapy, then why was this condition chosen and why call it a “sham”? Also it is unclear what ethical concerns preclude the use of an inert placebo (such as non-invasive needling) to test a currently unproven intervention.

The placebo control problem in acupuncture research.

The problem of what constitutes an appropriate placebo control for acupuncture is a source of discussion and is still not fully resolved (Ernst, 2008a). Devices are now available and validated that allow for the delivery of non-penetrating placebo acupuncture and for therapist and patient blinding. Critics of these approaches suggest that since they involve the device touching the skin they may not be truly inert (Langevin et al., 2006, Lund & Lundeberg, 2006, Lundeberg et al., 2007). It has been argued that trials using invasive or non-invasive shams fail to demonstrate the effects of “true” acupuncture since they themselves are active treatments (Langevin et al., 2006, Lund & Lundeberg, 2006, Lundeberg et al., 2007). By this argument any tactile sensory input might be considered an active treatment. Kaptchuk (2006b) has strongly argued that simply scratching or touching the skin is unlikely to be a plausible physiological treatment beyond ritual, stating that

“acupuncture should not hide behind the excuse that anything you do with a needle, including waving it, is a form of acupuncture”. Madsen et al. (2009) have demonstrated that the type of placebo acupuncture employed in studies has no significant effect on efficacy in pain trials whether or not the needling is shallow, incompatible with the principles of TCM, or even entirely non-penetrating.

Similarly a large multi-centre trial of acupuncture for chronic back pain was published with strong similarities to the GERAC trial (Cherkin et al., 2009). Individually tailored acupuncture with little restriction on the clinicians’ choices was compared to standardized acupuncture (a prescriptive approach considered effective by experts in the field), to “simulated acupuncture” in which subjects were stimulated by pressing a non-penetrating toothpick against the skin in such a way as to mimic the sensation of acupuncture, and to a “usual care” group who received no study-related care beyond what they or their physicians chose routinely. Again none of these acupuncture groups demonstrated superiority. Notably, the type of sham used in the current study did not allow for blinding of the treating acupuncturists, yet despite this potential for bias simulated and real acupuncture still demonstrated equivocal effects. All acupuncture groups performed better than the usual care group, but this group was arguably disadvantaged to a greater level than the conventional therapy group in the GERAC trial. Many in this group may have received no significant care and for this reason and those already discussed they are unlikely to have been the beneficiaries of a placebo effect. These results again confirm that acupuncture treatment guided by the principles of

TCM offers no additional benefit, and that acupuncture is no more effective than sham treatment. Thus it does not matter where the needles are placed, how deep they are inserted, or whether or not they are inserted at all. These findings are unsupportive of the concept of specific peripheral needling effects.

Discussions around the placebo issue within the acupuncture literature are informative in themselves, and provide further examples of what we have described as interpretive bias. In the published account of a recent discussion on sham and placebo controls by a group of acupuncture researchers (Langevin et al., 2006) it is acknowledged that verum acupuncture rarely out performs any type of sham condition. Numerous explanations for this phenomenon are posited, but the emphasis of discussion is firmly focused on how sham conditions may elicit active therapeutic effects rather than on what is arguably the most parsimonious explanation - that acupuncture does not have an active effect beyond placebo.

In a recent editorial members of the British Medical Acupuncture Society (White and Cummings, 2009) contend that “it is unfortunate that placebo control acupuncture remains problematic.” They suggest that the use of non-penetrating sham needles “pose a challenge” as they may “condition responses to a greater extent than a placebo pill.” This rather confusing argument appears to suggest that since an acupuncture-specific placebo condition induces a particularly strong placebo effect it disadvantages acupuncture in clinical trials. We find ourselves in disagreement with this

suggestion and would argue that the strong placebo effect elicited by these sham needling methods qualifies them as an ideal candidate to control for the influence of expectancy in acupuncture trials.

The consequences of interpretive bias

We feel that the arguments and evidence offered by the trial authors specifically and the acupuncture community more widely do not adequately justify their conclusions, and that this collective reluctance to meaningfully engage with the possibility that acupuncture is a placebo treatment is a good example of interpretive bias. This is not merely an academic argument. Evidence in medical literature is used to make decisions about individual patients' treatment, ultimately impacting on patients' health. Interpretive bias of this kind can have serious consequences: it is important, firstly, that the reported conclusions of a piece of research are an accurate representation of the findings, as they will often be all that is read. There are 5,000 medical journals published every month (Greenhalgh, 1997). To pick an example from just one discipline, it has been estimated that each month over 7,000 items relevant to primary care – studies, letters and editorials – are published, material which would take physicians trained in epidemiology over 600 hours to read in full and interpret (Alper et al., 2004).

A recent review of drug meta-analyses (Jørgensen et al., 2006) presents a useful parallel from mainstream medicine. This paper demonstrated that industry supported meta-analyses found similar effect sizes to comparable

independent studies, but differed in interpretation and discussion. Industry sponsored meta-analyses uniformly recommended use of the experimental drug, and gave more favourable conclusions than the equivalent Cochrane reviews, despite the comparable effect sizes. It appears that a similar bias may be present in the acupuncture literature; an earlier systematic review of acupuncture systematic reviews suggested a trend for reviewers with affiliation to a department of complementary medicine to conclude in favour of acupuncture (Derry et al., 2006).

As well as individual treatment decisions, evidence from research is also used to inform judgements about the deployment of limited healthcare resources, which ultimately impacts on individual patients' health. The GERAC trials specifically arose from the recommendations of the German Board of Physicians and Insurance Companies that research should be performed to provide an evidence base for deciding whether to include acupuncture as an insured health benefit. The conditions on which acupuncture would be considered useful are quite clear (Molsberger et al., 2006): acupuncture had to be (1) more effective than placebo or sham acupuncture and (2) tested against conventional therapy.

These trials clearly did not demonstrate efficacy over placebo. The authors' new interpretation that an improvement with a sham intervention represents a successful outcome for acupuncture has, however, had direct political consequences. Based on this representation of the trial's results, the German Federal Joint Committee of Physicians and Health Insurance Plans made

acupuncture for low back pain and knee pain an insured benefit (Haake et al., 2007, Park et al., 2008). Similarly in the United Kingdom, the National Institute for Health and Clinical Excellence (NICE) recently presented new clinical guidelines for the treatment of persistent low back pain that endorse the use of acupuncture, based in part on the findings of the GERAC back pain trial (NICE 2009). This endorsement comes despite the panel finding that the 2 trials identified in their review that employed a sham control group (Brinkhaus et al. 2006, Haake et al., 2007) demonstrated no superiority of verum acupuncture. Back pain is an extremely common presentation in primary care, and there is reason to believe that such decisions may lead to a significant expansion in service provision for an alternative therapy which has not shown itself to be any more effective than a sham control.

Although a comprehensive review of the entire canon of acupuncture research is beyond the scope of this article, recent examples suggest that interpretive bias may also extend beyond the reporting of individual clinical trials. Recently a Cochrane review has concluded by recommending acupuncture for migraine despite finding no difference between verum and sham acupuncture (Linde et al. 2009a). Another Cochrane review by the same group (Linde et al., 2009b) that studied acupuncture for tension headache found a significant effect of acupuncture over sham acupuncture, but the positive results of the meta-analysis were strongly influenced by one large trial (Endres et al., 2007). Notably the authors of that trial had themselves concluded that their data did not clearly demonstrate the superiority of verum acupuncture over sham therapy (Endres et al. 2007).

Finally, a recent systematic review (Yuan et al. 2008) of acupuncture for low back pain concluded that while there was only “moderate” evidence that acupuncture is better than no treatment, there was “strong” evidence that it is no better than sham treatment. Despite this the authors conclude that “acupuncture should be advocated in the European Guidelines for the treatment of low back pain.” This position is defended as the effects seen are similar to other therapies, such as manipulation, which are currently recommended within the guidelines. It seems that despite strong evidence that acupuncture is no more effective than placebo, non-inferiority to other modalities of limited efficacy is enough to justify its endorsement.

This seems to be a retrogressive step. Ideally evidence based practice in healthcare would “raise the bar” uniformly, rather than lower it, and there is a need throughout medicine for evidence based reappraisal of treatments which are currently used but of unproven efficacy, or with proven lack of efficacy. This process of rational disinvestment in failed treatments is often neglected in favour of testing new ones, but such a cultural blind spot is hard to justify in health economic terms.

We would also like to suggest that these many acupuncture studies demonstrating no benefit over sham could be presented more positively as evidence for the efficacy of the placebo effect, rather than of acupuncture. Many have argued passionately for the wider use of placebo treatments in mainstream medicine, on grounds of pragmatism, and this is not in itself an unreasonable position. However a discussion on this subject requires that

researchers are cautious and clear about whether a response is due to nonspecific placebo effects, or specific to the treatment tested, and it is important to perform and interpret controlled trials with this distinction in mind for many reasons. Patients may reasonably wish to know if specific or non-specific mechanisms underpin the effectiveness of any treatment they are receiving. Where the evidence shows a treatment to perform as a placebo, then there may also be ethical issues for clinicians around openly declaring this to patients, given the recent rise in importance of concepts such as informed consent, respect for patient autonomy, and working with patients to discuss evidence and choose treatment options collaboratively. Purchasers of health services may also find it useful to know if they are paying for theatre, and make an informed and pragmatic decision on the costs, benefits and ethics of such an activity at an organisational level.

Finally it seems reasonable to suggest that if placebo treatments are to be considered routinely they must at least be free of adverse effects. Cherkin et al., (2009) found a significantly higher incidence of minor adverse affects (pain, dizziness, back spasms) with penetrative acupuncture versus simulated acupuncture that was not offset by any significant therapeutic benefit. In this study simulated acupuncture was associated with no adverse affects. In the largest safety study to date of over 229,000 patients who received acupuncture, 8.6% of acupuncture patients reported at least one adverse effect including bleeding and haematoma (6.1%) (Witt et al. 2009). While serious adverse effects such as pneumothorax and nerve injury were extremely rare it is arguable that any risk of iatrogenic injury might discourage

the use of real acupuncture as an ethical placebo intervention, and that perhaps only the equally effective non-penetrating simulated acupuncture should be considered.

Conclusion

There has been a wealth of clinical research into acupuncture and the results are not compelling in their support for the therapy, showing little or no benefit over sham acupuncture. Despite this the acupuncture literature repeatedly fails to address one of the more plausible explanations for these findings: that acupuncture may be ineffective beyond placebo. We have given an account of this process around one recent individual study, permitting an exploration of the issues in detail. This specific trial was not unusual in its interpretation, but was high-profile, covered extensively in the media, had a significant political impact on healthcare provision, and offers an example of how, with interpretive bias, even gold standard research methodologies can lead to questionable conclusions.

References

Airaksinen O., Brox J.I., Cedraschi C., Hildebrandt J., Klaber-Moffett J., Kovacs F., et al. (2006) Chapter 4. European guidelines for the management of chronic nonspecific low back pain.; COST B13 Working Group on Guidelines for Chronic Low Back Pain. *European Spine Journal.* Suppl 2, S192-300

Alper B.S., Hand J.A., Elliott S.G., Kinkade S., Hauan M.J., Onion D.K., Sklar B.M. (2004) How much effort is needed to keep up with the literature

relevant for primary care? *Journal of the Medical Library Association*. 92(4), 429–437.]

Assendelft, W.J.J., Morton, S.C., Yu, E.I., Suttorp, M.J., & Shekelle, P.G. (2004) Spinal manipulative therapy for low back pain. *Cochrane Database of Systematic Reviews*: CD000447.

Bausell, R.B., Lao, L., Bergman, S., Lee, W., & Berman, B.M. (2005). Is acupuncture analgesia an expectancy effect? preliminary evidence based on participants' perceived assignments in two placebo-controlled trials. *Evaluation and the Health Professions*, 28(1), 9-26.

BBC News (2007, 25 September). Needles 'are best for back pain'. Retrieved from http://news.bbc.co.uk/

Brinkhaus, B., Witt, C.M., Jena, S., Linde, K., Streng, A., Wagenpfeil, S. et al. (2006) Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Archives of Internal Medicine,* 166(4), 450-7.

Cherkin D.C., Eisenberg D., Sherman K.J., Barlow W., Kaptchuk T.J., Street J. et al., (2001) Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Archives of Internal Medicine*,161(8),1081-8

Cherkin D.C., Sherman K.J., Avins A.L., Erro J.H., Ichikawa L., Barlow W.E., Delaney K et al., (2009) A randomized trial comparing acupuncture, simulated acupuncture and usual care for chronic low back pain. *Archives of Internal Medicine,* 169(9),858-866

Clarke, J.A., van Tulder, M.W., Blomberg, S.E.I., de Vet, H.C.W., van der Heijden, G.J.M.G., Bronfort, G., et al. (2007) Traction for low-back pain with or with out sciatica. *Cochrane Database of Systematic Reviews*: CD003010.

Derry, C.J., Derry, S., McQuay, H.J., & Moore, R.A. (2006). Systematic review of systematic reviews of acupuncture published 1996-2005. *Clinical Medicine, Journal of the Royal College of Physicians of London, 6*(4), 381-386.

Diener, H., Kronfeld, K., Boewing, G., Lungenhausen, M., Maier, C., Molsberger, A., et al. (2006). Efficacy of acupuncture for the prophylaxis of migraine: A multicentre randomised controlled clinical trial. *Lancet Neurology*, 5(4), 310-316.

Endres, H.G., Böwing, G., Diener, H.C., Lange, S., Maier, C., Molsberger, A., Zenz, M., Vickers, A.J., Tegenthoff, M. (2007) Acupuncture for tension-type headache: a multicentre, sham-controlled, patient-and observer-blinded, randomised trial. *Journal of Headache & Pain*, 8(5), 306-14

Endres H.G. (2008) Acupuncture: Specific and Non-Specific Effects. *Forschende Komplementärmedizin.* 15:6-8

Endres, H.G., Molsberger, A., & Haake, M. (2008a) Sham acupuncture is not a placebo - reply. *Archives of Internal Medicine, 168*(9), 1012.

Endres, H.G., Molsberger, A., & Haake, M. (2008b) Acupuncture ineffective, attention effective? - reply. Archives of Internal Medicine, 168(5), 551-552.

Ernst, E. (2008a) Acupuncture: What does the most reliable evidence tell us? *Journal of Pain and Symptom Management,* Article in Press doi:10.1016/j.jpainsymman.2008.04.009

Ernst, E. (2008b). Acupuncture ineffective, attention effective? Archives of Internal Medicine, 168(5), 551.

Fleming N (2007, Sept 26) Acupuncture 'best therapy for back pain'. Daily Telegraph, United Kingdon. Retrieved from http://www.telegraph.co.uk/

French, S.D., Cameron, M., Walker, B.F., Reggars, J.W., & Esterman, A.J. (2006) A Cochrane review of superficial heat or cold for low back pain. *Spine,* 31(9), 998-1006.

Goldacre, B. (2007) Medicalisation - don't take it lying down. Retrieved October 10th 2008 from http://www.badscience.net/2007/09/542/

Greenhalgh, T. (1997) How to read a paper. London: BMJ Books

http://news.google.co.uk/archivesearch?q=back+pain+acupuncture&cid=862115041430753 8&um=1&sa=N&start=10. Retrieved January 20th 2009

Guzmán, J., Esmail, R., Karjalainen, K., Malmivaara, A., Irvin E., & Bombardier C. (2002) Multidisciplinary bio-psycho-social rehabilitation for chronic low-back pain. *Cochrane Database of Systematic Reviews*: CD000963.

Haake, M., Müller, H.H., Schade-Brittinger, C., Prinz, H., Basler, H.D., Streitberger, et al. (2003) The German multicenter, randomized, partially

blinded, prospective trial of acupuncture for chronic low-back pain: a preliminary report on the rationale and design of the trial. Journal of Alternative and Complementary Medicine. (5),763-70.

Haake, M., Müller, H., Schade-Brittinger, C., Basler, H. D., Schäfer, H., Maier, C., et al. (2007). German acupuncture trials (GERAC) for chronic low back pain: Randomized, multicenter, blinded, parallel-group trial with 3 groups. *Archives of Internal Medicine, 167*(17), 1892-1898.

Hayden, J.A., van Tulder, M.W., Malmivaara, A., & Koes, B.W. (2005) Exercise therapy for treatment of non-specific low back pain. *Cochrane Database of Systematic Reviews*: CD000335.

Johnson C.K. (2007, September 25) Study: Acupuncture Works for Back Pain. The Washington Post. Retrieved from http://www.washingtonpost.com/

Jørgensen AW, Hilden J, & Gøtzsche PC (2006) Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *British Medical Journal*

Kalauokalani, D., Cherkin, D. C., Sherman, K. J., Koepsell, T. D., & Deyo, R. A. (2001). Lessons from a trial of acupuncture and massage for low back pain: Patient expectations and treatment effects. Spine, 26(13), 1418-1424.

Kaptchuk, T. J. (2003). Effect of interpretive bias on research evidence. *British Medical Journal, 326*(7404), 1453-1455.

Kaptchuk, T.J., Stason, W.B., Davis, R.B., Legedza, A.R., Schnyer, R.N., Kerr, C.E., et al. (2006a). Sham device v inert pill: randomised controlled trial of two placebo treatments. *British Medical Journal,* 332(7538), 391-7

Kaptchuk, T. J. (2006b) A plastic eagle feather is a placebo. Rapid response to, Kaptchuk, T. J., Stason, W. B., Davis, R. B., Legedza, A. T. R., Schnyer, R. N., Kerr, C. E., et al. (2006). Sham device versus inert pill: Randomised controlled trial of two placebo treatments. *British Medical Journal, 332*(7538), 391-394.

Kaptchuk, T.J., Kelley, J.M., Conboy, L.A., Davis, R.B., Kerr, C.E., Jacobson, E.E., et al. (2008) Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome. British Medical Journal, 336(7651), 999-1003

Langevin, H. M., Hammerschlag, R., Lao, L., Napadow, V., Schnyer, R. N., Sherman, K. J., et al. (2006). Controversies in acupuncture research: Selection of controls and outcome measures in acupuncture clinical trials. *Journal of Alternative and Complementary Medicine, 12*(10), 943-953.

Lawrence J. (2007, September 25) Acupuncture is best way to treat back pain, study finds. The Independent. Retrieved from http://www.independent.co.uk/

Linde, K., Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., et al. (2007). The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain. *Pain,* 128(3), 264-271.

Linde, K., Allais, G., Brinkhaus, B., Manheimer, E., Vickers, A., White, A.R. (2009a) Acupuncture for migraine prophylaxis. *Cochrane Database of Systematic Reviews*, 1 : CD001218

Linde, K., Allais, G., Brinkhaus, B., Manheimer, E., Vickers, A., White A.R. (2009b) Acupuncture for tension-type headache. *Cochrane Database of Systematic Reviews*, 1 : CD007587

Lund, I., & Lundeberg, T. (2006). Are minimal, superficial or sham acupuncture procedures acceptable as inert placebo controls? *Acupuncture in Medicine, 24*(1), 13-15.

Lundeberg, T., Lund, I., Näslund, J. (2007) Acupuncture- self-appraisal and the reward system. *Acupuncture in Medicine* 25(3), 87-99.

Madsen, M.V., Gøtzsche PC., Hróbjartsson, A. (2009) Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. *British Medical Journal* 338, a3115

Moerman, D.E., (2002) Meaning, Medicine and the 'Placebo Effect' New York: Cambridge University Press

Molsberger, A. F., Streitberger, K., Kraemer, J., Brittinger, C. S., Witte, S., Boewing, G., et al. (2006). Designing an acupuncture study: II. the nationwide, randomized, controlled german acupuncture trials on low-back pain and gonarthrosis. *Journal of Alternative and Complementary Medicine, 12*(8), 733-742.

NICE (2009) Low back pain: the early management of patients with persistent (longer than 6 weeks) non-specific low back pain. NICE clinical guideline 88. Retrieved from www.nice.org.uk on 27/5/09.

Park, J., Linde, K., Manheimer, E., Molsberger, A., Sherman, K., Smith, C., et al. (2008). The status and future of acupuncture clinical research. *Journal of Alternative and Complementary Medicine, 14*(7), 871-881.

Reuters (2007 September 25) Acupuncture helps back pain, don't ask how. Retrieved from http://uk.reuters.com/

Schultz, K.F., Chalmers, I., Hayes, R.J., Altman, D.G. (1995) Empirical evidence of bias: Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association*, 273, 408–12

Vase, A., Riley III, J. L., & Price, D. D. (2002) A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain,* 99, 443–452

Wager T. D. (2005) The Neural Bases of Placebo Effects in Anticipation and Pain. *Seminars in Pain Medicine*, 3, 22-30

Wand B.M., O'Connell N.E. (2007) German Acupuncture Trials for Chronic Low Back Pain. *Archives of Internal Medicine*, 168(9),1011

White A., Cummings M. (2009) Does acupuncture relieve pain? Interpreting the effects of sham acupuncture holds the answer. *British Medical Journal*, 338, a2760

Williams C (2007, September 24) Fake or Not, Acupuncture Helps Back Pain: Even Sham Version of Technique Beats Conventional Therapies. ABC News. Retrieved from http://abcnews.go.com/

Witt C.M.,, Pach D., Brinkhaus B., Wruck K., Tag B., Mank S., Willich S.N. (2009) Safety of acupuncture: results of a prospective observational study with 229,230 patients and introduction of a medical information and consent form. *Forschende Komplementärmedizin und Klassische Naturheilkunde*, 16(2), 91-7

Wood L., Egger M., Gluud L.L., Schulz K.F., Jüni P., Altman D.G., et al. (2008) Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *British Medical Journal* 336(7644), 601-5

Yuan J., Purepong N., Kerr D.P., Park J., Bradbury I., McDonough S. (2008) Effectiveness of Acupuncture for Low Back Pain. A Systematic Review. *Spine*, 33(23), E887-900