

Whittington Hospital Journal Club

Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain Thomas, K. J. Et al *BMJ* 2006 free pdf available from:  
<http://www.bmj.com/content/333/7569/623.abstract>

14 September 2010, hosted by the Physiotherapy Department  
14 staff present  
Session facilitated by Senga Steel and Richard Peacock

We used the CASP Randomised Controlled Trial checklist to structure our discussion  
<http://www.sph.nhs.uk/what-we-do/public-health-workforce/resources/critical-appraisals-skills-programme>

### **1. Did the study ask a clearly-focused question?**

Yes.

The paper aimed to determine whether a short course of traditional acupuncture improves longer term outcomes in patients with persistent non-specific low back pain. Acupuncture was compared with usual care.

Patients were recruited from primary care with non-specific back pain of 4 – 52 weeks duration.

The acupuncture was provided by 6 experienced, registered acupuncturists based in 2 non-NHS clinics in York.

Treatment consisted of up to 10 individualised sessions.

Patients in the usual care group received NHS treatment as assessed by their general practitioner, which could include drugs, recommended exercises, physiotherapy and massage.

Treatment was for 12 weeks.

The primary outcome measure was the bodily pain dimension of the SF36 survey measured at baseline, 3 months, 12 months and 24 months. Secondary outcomes included the Oswestry pain disability index, the McGill present pain index and the remaining 7 dimensions of SF36.

There was also an economic evaluation of the data, which is separately published and is available from:

<http://www.bmj.com/content/333/7569/626.full.pdf+html?sid=3121a6e1-68f9-46c9-b480-2f9de5932ea8>

### **2. Was this a randomised controlled trial (RCT) and was it appropriately so?**

The paper is described as a pragmatic, open, randomised controlled trial. This means that both patients and practitioners were aware of the arm they were randomised to. For more on blinding see section 4 below.

Previous research in this area has tended to compare acupuncture with sham acupuncture but the authors argue that sham procedures are problematic. They attempted to address the possible confounding effects of patients' prior beliefs about the efficacy of acupuncture by asking 2 questions prior to randomisation: whether patients expected their back pain to improve and whether they believed acupuncture might help. This was then compared with SF36 scores at 24 months.

### **3. Were participants appropriately allocated to intervention and control groups?**

The randomisation method is described. The researchers decided to adopt a 2:1 randomisation pattern, so that there were approximately twice as many patients allocated to the acupuncture arm as to usual care. Randomisation in these proportions is a recognised technique although not common in research we have previously critiqued. The authors argue that it allowed them to test for effects between acupuncturists. Readers of the paper need to bear this in mind when looking at the results!

GPs from 16 general practices identified patients suitable for management in primary care. They were then assessed for eligibility by the study researcher. There were a number of exclusion criteria including those currently receiving acupuncture, possible spinal disease, motor weakness, prolapsed central disc, bleeding disorders, past spinal surgery or those pending litigation. We discussed the exclusion characteristics in the journal club and whether they effected the generalisability of the study.

Allocation to individual acupuncture practitioners was made by convenience to patients and availability of appointments. A table is provided describing baseline characteristics of the 2 groups. The authors feel the groups are well balanced in every respect but one: there were 11 patients (7%) assigned to the acupuncture group who were permanently unable to work owing to low back pain and none in the usual care group.

### **4. Were participants, staff and study personnel “blind” study group?**

As mentioned in 2 above, neither practitioners nor patients were blinded in this study. We discussed whether this mattered and whether it introduced the potential for bias in the study. The questions concerning prior belief asked at baseline were a way to control for this. We also discussed the possibility of unconscious bias on the part of the authors as no null hypothesis was stated: this is considered further below. The authors were from the University of Sheffield, the Foundation for Traditional Chinese Medicine in York and general practice.

The statistical analysis of the primary outcome was repeated by a statistician who was blinded to the treatment allocation.

**5. Were all of the participants who entered the trial accounted for at its conclusion?**

We had concerns in this area. Firstly there seemed to be considerable blurring between the arms. For example 40% of those in the acupuncture group received drugs, 35% had recommendations for exercise and over 20% had physiotherapy or manipulation. In addition, around 10% of the patients in **both** arms had acupuncture outside of the study. We felt that this made it very difficult to determine what might explain any difference in outcome between the groups. There is some loss to follow-up as the study progresses: 76% of patients are followed up over 24 months. The paper states that analysis occurred on an intention to treat basis: this is usually considered a strength in a RCT. However, the presentation of results do not always appear to apply intention to treat. A total of 241 patients were randomised but in the table presenting differences in outcome score, the unadjusted SF36 analysis for 12 months is based on 215 patients and for 24 months on 182 patients ie just for patients completing follow-up. It would have helped to have had a better explanation as to why we are not presented with intention to treat results. The authors state that they detected no differences between the characteristics of missing cases in the two arms, but that “the effect of the missing data is unknown”: we were not clear if the authors were suggesting missing data might have favoured the acupuncture arm, but were concerned that this undermining of intention to treat might represent possible bias.

**6. Were the participants in all groups followed up and data collected in the same way?**

As far as we can tell, yes. Follow-up was principally carried out by post. Were responses were not received main outcome data was sought by telephone.

**7. Did the study have enough participants to minimise the play of chance?**

The study was originally planned with a 1:1 randomisation and 100 patients in each group would have been necessary to detect a 10 point difference in the SF36 score at 12 months with 90% power, assuming a 10 – 15% drop-out rate. The 10 point difference was considered a clinically significant based on pilot data.

When it was decided to run the study with a 2:1 randomisation a total of 240 patients were required to maintain the original power. This seemed odd to us, as the usual care group would be reduced to 80 and we were not sure how this smaller group could retain the same power.

In total 241 patients were randomised, so adequate power was achieved as long as intention to treat was applied. In the event we are not given intention to treat results and so the study may have been underpowered.

**8. How are the results presented and what is the main result?**

There was no significant difference between the 2 groups in SF36 scores at 12 months (adjusted mean difference of 6 points in favour of acupuncture) and a small significant difference at 24 months (9 point difference in favour of acupuncture) based on patients completing follow-up. The authors acknowledge that the 10 point difference in scores was not achieved at either 12 or 24 months but argue that a 5 point difference does actually represent a clinically worthwhile benefit and between 5 and 9 points is a “moderate effect”. We were not comfortable with this, or with some of the authors’ use of language. They describe the result at 12 months as weak evidence (we would have preferred no evidence) and at 24 months as stronger evidence (at best we considered this as questionable).

None of the other pain-related outcome measures reported achieve statistical significance and there is no evidence of functional improvement.

There were some improvements in patient perceived benefits for acupuncture, for example satisfaction with treatment and concerns about back pain. The authors discuss the fact that acupuncture is more than just needling and includes elements of the therapeutic relationship and attention from the practitioner.

Those in the acupuncture group who believed acupuncture would help with their back pain did not have significantly better results than those randomised to usual care.

**9. How precise are these results?**

P values and confidence intervals are provided for results. Overall effectiveness of the intervention is, however, unclear.

**10. Were all important outcomes considered so the results can be applied?**

This is an interesting area of research but the journal club felt the findings should be interpreted with caution in light of some of the methodological problems we considered. At best there was a modest benefit at 24 months. We felt there was blurring between the treatment arms, problems over the powering of the study, inconsistent use of intention to treat and confusion over the difference in outcome that should be considered significant. In general we felt this was a missed opportunity and agreed with one of the responses to the article in the BMJ that suggested the authors “do it again and do it better”.

The authors discuss the limitations of the study and make a number of suggestions for future research, both qualitative and quantitative.