Journal homepage:
www.jcimjournal.com/jim
www.elsevier.com/locate/issn/20954964
Available also online at www.sciencedirect.com.

• Commentary

The methodology flaws in Hinman's acupuncture clinical trial, Part I: Design and results interpretation

Arthur Yin Fan
McLean Center for Complementary and Alternative Medicine, PLC., Vienna, VA 22182, USA

Keywords: acupuncture; arthralgia; chronic knee pain; Hinman; Zelen design; clinical trial; statistics; flaws

Citation: Fan AY. The methodology flaws in Hinman's acupuncture clinical trial, Part I: Design and results interpretation. *J Integr Med*. 2015; 13(2): 65–68.

In the October 2014 edition of *JAMA*, Dr. Hinman and her colleagues published an acupuncture clinical trial entitled "Acupuncture for Chronic Knee Pain: A Randomized Clinical Trial" and concluded that "in patients older than 50 years with moderate or severe chronic knee pain, neither laser nor needle acupuncture conferred benefit over sham for pain or function. Our findings do not support acupuncture for these patients"[1].

The author strongly disagrees with this conclusion, as there were serious flaws in the trial design, the statistical analysis of the data and in the interpretation of the results of this study. This commentary is prepared in two parts. In Part I, the study design, acupuncture dose and interpretation of results will be investigated. In Part II, the Zelen design and dilution effect will be discussed.

1 There is a major mistake in the primary testing factor in this RCT: the laser acupuncture should be the primary testing factor, not the needle acupuncture

In a vigorous randomized controlled trial (RCT), there is a primary testing factor or objective. An example of a major or primary testing factor could be a new therapy of unknown efficacy. The major testing factor should be compared with a non-intervention group, or what is commonly known as a control group (or Naïve group), negative control (sham intervention group) and, if possible, with a positive control group, in which this positive control has already been tested as an effective therapy[2,3]. In Hinman *et al's* study[1], there are four groups: control group, conventional or needle acupuncture group, laser acupuncture group and sham laser acupuncture group. Obviously, in this RCT, the major testing factor is laser acupuncture; acupuncture served as a positive control. This is supported by evidence found outside of the article in question. Dr. Lao *et al*[4] and Dr. Li[5] pointed out that the authors registered this trial as testing laser acupuncture, instead of conventional acupuncture[6,7]. There has also been strong evidence that acupuncture is an effective therapy for chronic knee pain[2,3]. Thus, acupuncture should have served as the positive control in this trial.

If acupuncture was the major testing factor as the article title indicated[1], it should have had the proper control groups. In this trial with a Zelen design, the patients in the control (non-intervention) group did not have informed consent, but the patients in the acupuncture group did have informed consent. This means that there were differences in informed consent among the groups (*i.e.*, not matched), so there was no comparability between these groups as the patients were not blinded. Moreover, the sham laser acupuncture group actually could not be treated as a valid negative control for the acupuncture

http://dx.doi.org/10.1016/S2095-4964(15)60170-4
Received February 20, 2015; **accepted** March 5, 2015.
Correspondence: Arthur Yin Fan, CMD, PhD, LAc; Tel: +1-703-499-4428; E-mail: ArthurFan@ChineseMedicineDoctor.US

treatment, because these two interventions do not have comparability in both characteristics and form (*i.e.*, not matched). Also, there was no blinding method performed between these two groups, because both the patients and the administrators who performed the interventions knew the difference between the groups, such as needling acupuncture and sham laser acupuncture; in contrast, there was only blinding between laser acupuncture and sham laser acupuncture (*i.e.*, matched, to some extent). The laser acupuncture instead of the needle acupuncture was the primary or major testing factor; if acupuncture was the primary testing factor, as Hinman emphasized, the quality of this trial would be very poor, because the trial does not follow the scientific principle for an RCT[8–10], as there was neither any blinding used nor correct control groups for such an acupuncture clinical trial.

It is improper to test two different testing factors in one RCT[8–10], such as both laser acupuncture and acupuncture, as was reported in this study design[1]. One could suggest that Hinman intended to confuse readers in this article; she claimed the trial was to test both acupuncture and laser acupuncture. However, in the title of the article, she clearly stated "acupuncture for chronic knee pain: a randomized clinical trial", which suggests that acupuncture was the primary testing factor — not laser acupuncture.

2 The interpretation of the results was misleading

Let us assume that the data and statistic work are all correct in the original article.

2.1 ***There is no significant difference between the laser acupuncture group and the control group, or between the laser acupuncture group and the sham laser acupuncture group***

From this statement, two conclusions are possible. Either laser acupuncture is not effective, or the strength or concentration of the laser acupuncture treatment has not elicited a response. One can also theorize that perhaps there were other variables at play. This interpretation only looks at the results from the laser acupuncture, non-intervention and sham laser acupuncture groups. This interpretation of the results cannot speak to the efficacy of needle acupuncture in the treatment of knee pain. Again, one must also consider that there could have been other factors that influenced the study results. One must also verify that the laser acupuncture used was as effective as possible (*i.e.,* timing or duration of the treatment or stimulation).

2.2 ***There is no significant difference in the results between the laser acupuncture group and the acupuncture group, and at the same time, there was no significant difference between the laser acupuncture group and the control group***

It is possible that the treatments in the acupuncture and/or the laser acupuncture groups were not used at an optimal dose (*i.e.*, the stimulation or duration of the laser or needle was not enough to induce an effect, such as would be recommended in other trials on the same interventions), the sample sizes were too small, or the results found were due to other methodological flaws. Based on the design and the statistics, there is no reason to compare the acupuncture group with the sham laser acupuncture group (a tested or positive control *vs.* a negative control); further, these two styles of interventions are not comparable. Should the researchers want to look at the effectiveness of conventional, needle acupuncture, a separate design or study would be needed in order to get valid results. Acupuncture and laser acupuncture are two separate interventions that would need two separate control groups. Comparing the primary intervention, acupuncture, to results from the negative control group of a different intervention (sham laser acupuncture) may cloud statistical results.

In fact, in the original article, there is a significant difference between the acupuncture group and the control group with regards to the Overall Pain score and Western Ontario and McMaster Universities Osteoarthritis Index; both were $P<0.05$ at week 12 of the study. This means that acupuncture did decrease pain intensity and improve function in patients with chronic knee pain. However, the author did not interpret or report this crucial information. Furthermore, the concept of laser acupuncture should not be mixed with acupuncture. Laser acupuncture does not involve a needle, or instrument piercing of the skin; therefore, the authors' conclusions about the ineffectiveness of laser acupuncture should not be applied to acupuncture.

One could conclude that Dr. Hinman misleads readers by testing acupuncture as a major intervention in this RCT. There was no significance between the positive control and the naïve control (*i.e.*, acupuncture and control groups). Therefore, we can only conclude that the positive control, acupuncture was under-dosed or the study was otherwise flawed. That the positive control shows significance is a basic sign of the success of a clinical trial. From this perspective, Hinman's trial was a failed clinical trial for laser acupuncture. As it would be unethical to publish an astonishing article, with a group of almost scrapped data and confusing logic, that misleads the readers, including the general public, medical society and policy makers, the researchers should have re-adjusted or re-designed their study instead of publishing it.

3 The "under-dosed" acupuncture treatments diluted the potential real effectiveness of acupuncture

In the article[1], the acupuncture treatments were "20 minute treatments that were delivered once or twice weekly for 12 weeks, with 8 to 12 sessions in total permitted". This

means that in the acupuncture group, the patients received 0, 1 or 2 acupuncture treatments per week, in which a total of 8–12 sessions were allowed. The fact that each patient received a different number of treatments plays a factor in the results. Further, the treatment time was very short. Often, acupuncture treatments last longer than 20 min per session on average. Finally, there are other important details about the treatment, such as depth of needle insertion and the *deqi* technique, that were not mentioned in the publication. Needling depth and needling technique (*i.e.*, manipulation) are very important in acupuncture and those practicing in the acupuncture field would agree that these details would also influence the effectiveness of treatment. This acupuncture dose (by which is meant number of sessions) is significantly less than the doses in published acupuncture RCTs on chronic knee pain (or other chronic pain)[2,3]. In licensed acupuncturists daily practices, practitioners use 45-minute acupuncture treatments, at least twice per week for 12 weeks. Practitioners also use a deeper stimulation and *deqi* technique. The acupuncture treatments in this paper were, at most, sub-optimal.

4 Laser acupuncture and acupuncture would be effective in Hinman's RCT, if the statistics were re-analyzed after re-adjusting the data

The original raw data could not be provided by the authors, so it is difficult to re-conduct an analysis of variance test to compare the differences among the groups; a *student-t* test was used for the purposes of this article instead. The data of the overall pain from Table 2 in the original paper[1] were re-adjusted and re-present in Table 1 (the analysis for other parameters are omitted to save space). Some data was switched to the control group (this included the 21 participants who did not receive any treatment although they were originally allocated to the intervention groups). Thus, the new sample sizes are 90, 58, 54 and 54 in the control, laser acupuncture, sham laser acupuncture and acupuncture groups, respectively. Assuming that these 21 cases have the same mean (4.4) and standard deviation (2.4) as those in the control group, as they did not receive any intervention (*i.e.*, laser acupuncture, sham laser acupuncture or acupuncture), this would equate them to those in the control group who also did not receive any intervention. (**Dr. Fan notes:** in all groups, patients were treated with non-steroidal anti-inflammatory drugs without reporting dosage and frequency in the original study[1].) The new mean(s) calculation needs to consider eliminating the influence of the patients who did not get the intervention, from the original mean. For example, in the acupuncture group with 64 patients (used in the original article), the total Overall Pain score was: 3.3×64=211.2; the total Overall Pain score in the ten drop-out patients was: 4.4×10=44; then in the acupuncture group with 54 patients, the total Overall Pain score should be: 211.2−44=167.2. The new mean will be: 167.2/54=3.09. We can get new means and standard deviations for the laser acupuncture, sham laser acupuncture and acupuncture groups for re-calculating the *P* value (Table 1).

Although the doses (treatment quality and quantity) used were sub-optimal, as compared to other studies[2,3,11], both laser acupuncture and acupuncture were actually effective in decreasing Overall Pain score at week 12 (both $P<0.01$; Table 1). The new results show that acupuncture is more effective than laser acupuncture, and that laser acupuncture is more effective than both the sham laser acupuncture and the control. However, sham laser acupuncture also seems to be effective ($P<0.01$); the reason why it is effective needs further investigation. Compared with the sham laser acupuncture (the negative control), or the acupuncture (the positive control), there are no statistical differences either between the laser acupuncture and the sham acupuncture, or between the laser acupuncture and the acupuncture. This phenomenon suggests that the original data need to be re-grouped and re-analyzed. If the authors realized that the quality of the current reported trial[1] is poor and cannot be fixed, the RCT should be re-designed and re-conducted.

5 Acknowledgements

Dr. Arthur Yin Fan would like to thank Ms. Sarah Faggert for editing support.

Table 1 Comparison of four groups at week 12, with the pain score, and *P* value

Group	Served as	Sample size (*n*)	Mean	Standard deviation	Compared with the control	
					t value*	*P* value
Control	Naive control	90	4.40	2.4	n/a	
Laser acupuncture	Major testing factor	58	3.28	2.2	2.90	<0.01
Sham laser acupuncture	Negative control	54	3.33	2.2	2.68	<0.01
Acupuncture	Positive control	54	3.09	2.2	3.34	<0.01

*Analysis of variance could not be conducted due to original data could not be got from the author[1].

6 Competing interests

The author is an independent researcher, and declares that he has no competing interests.

REFERENCES

1 Hinman RS, McCrory P, Pirotta M, Relf I, Forbes A, Crossley KM, Williamson E, Kyriakides M, Novy K, Metcalf BR, Harris A, Reddy P, Conaghan PG, Bennell KL. Acupuncture for chronic knee pain: a randomized clinical trial. *JAMA*. 2014; 312(13): 1313–1322.

2 Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. *Ann Intern Med*. 2004; 141(12): 901–910.

3 Vickers AJ, Linde K. Acupuncture for chronic pain. *JAMA*. 2014; 311(9): 955–956.

4 Lao L, Yeung WF. Treating chronic knee pain with acupuncture. *JAMA*. 2015; 313(6): 627–628.

5 Li YM. Treating chronic knee pain with acupuncture. *JAMA*. 2015; 313(6): 628.

6 Hinman RS, McCrory P, Pirotta M, Relf I, Crossley KM, Reddy P, Forbes A, Harris A, Metcalf BR, Kyriakides M, Novy K, Bennell KL. Efficacy of acupuncture for chronic knee pain: protocol for a randomized controlled trial using a Zelen design. *BMC Complement Altern Med*. 2012; 12: 161.

7 Australia New Zealand Clinical Trials Registry. *Acupuncture for chronic knee pain trial [registry identifier: ACTRN12609001001280]*. [2015-02-22]. https://anzctr.org.au/Trial/Registration/TrialReview.aspx?id=308156.

8 Consort Transparent Reporting of Trials. *Consort 2010: Trial design*. [2015-03-09]. http://www.consort-statement.org/checklists/view/32-consort/72-trial-design.

9 National Health Institutes. Quality Assessment of Controlled Intervention Studies. Systematic evidence reviews & clinical practice guidelines. [2015-03-04]. http://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/rct.

10 West A, Spring B. Randomized controlled trials. [2015-03-09]. http://www.ebbp.org/course_outlines/randomized_controlled_trials/

11 Baxter GD, Tumilty S. Treating chronic knee pain with acupuncture. *JAMA*. 2015; 313(6): 626–627.