

Validating a New Non-penetrating Sham Acupuncture Device: Two Randomised Controlled Trials

Jongbae Park, Adrian White, Clare Stevinson, Edzard Ernst, Martin James

Jongbae Park
research fellow

Adrian R White
senior lecturer

Clare Stevinson
research fellow

Edzard Ernst
professor
Complementary
Medicine, Peninsula
Medical School,
Universities of Exeter
& Plymouth, Exeter,
UK

Martin A James
consultant physician
Dept. of Medicine for
the Elderly, Royal Devon
& Exeter Hospital, UK

Correspondence:
Jongbae Park

Jongbae.Park@pms.ac.uk

Summary

For clinical trials of acupuncture, it would be desirable to have a sham procedure that is indistinguishable from the real treatment, yet inactive. A sham needle has been designed which telescopes instead of penetrating the skin. The Park Sham Device involves an improved method of supporting the sham needle and requires validation. The objective of these studies was to test whether the sham procedure using the new device was 1) indistinguishable from the same procedure using real needles in acupuncture naïve subjects, and 2) inactive, where the specific needle sensation (de qi) is taken as a surrogate measure of activity. The studies were designed as subject and assessor blind, randomised controlled trials. Study 1) included 58 patients enrolled in a clinical trial of acupuncture for acute stroke. Study 2) included 63 healthy, acupuncture naïve, adult volunteers. The interventions used were real or sham acupuncture using the Park Sham Device. Study 1) was set in a district general hospital, and study 2) in a university laboratory. The outcome measure in study 1) was the form of treatment that patients believed they had received. In study 2) the outcome measure was experience of de qi, as judged by three acupuncture experts. No patient in either group (study 1) believed he or she had been treated with the sham needle. In 40 volunteers (study 2) for whom experts achieved consensus, the relative risk of experiencing de qi with real acupuncture to that with sham acupuncture was 15.38 (95% CI 2.26 to 104.86). The inter-rater reliability of all 13 experts (study 2), calculated from their judgements on 10 subjects selected by randomisation, was 0.52 (95% CI 0.19 to 0.61). In conclusion, the results suggest that the procedure using the new device is indistinguishable from the same procedure using real needles in acupuncture naïve subjects, and is inactive, where the specific needle sensation (de qi) is taken as a surrogate measure of activity. It is therefore a valid control for acupuncture trials. The findings also lend support to the existence of de qi, a major concept underlying traditional Chinese acupuncture.

Keywords

Acupuncture, sham acupuncture, non-penetrating sham control, randomised controlled trial.

Introduction

In clinical trials of the efficacy of manual treatments like surgery or acupuncture, control groups are commonly given 'sham' procedures. Sham procedures, in order to be true placebos, must be 1) indistinguishable from the real treatment and 2) inactive. In acupuncture trials, various controls for the process of inserting a needle and stimulating it have been used. These include needling off point or at inappropriate points,^{1,2} pricking or scratching with blunt needle,^{3,4} needling with an adhesive plug at the tip,⁵ pressing a finger-nail,⁶ off point and sham electrostimulation,⁷ sham TENS,⁸ plastic guide

tube,⁹ guide-tube with blunted cocktail stick,¹⁰ sham electrostimulation,¹¹ and blunted needle with a piece of cube shaped elastic foam to position the needle.¹² None of these seems to be ideal.

A new sham acupuncture needle has been developed.¹³ It appears similar to the real needle, but is blunt and free to slide within the handle so that, when pressed, it telescopes into the handle instead of penetrating the skin. It was originally applied through an adhesive dressing for support,^{13,14} but we developed a device (Park Sham Device; PSD) which includes a sham needle and an adhesive tube.¹⁵ We believe that this adhesive tube holds the sham needle more firmly in place,

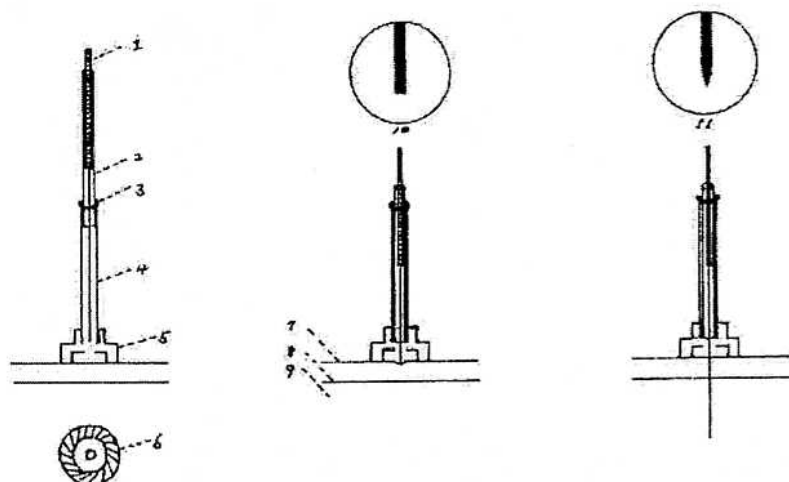


Figure 1a Park Sham Device with sham needle and real needle. 1. Needle handle, 2. Guide tube, 3. Guide O-Ring, 4. Park tube, 5. Flange, 6. Double sided tape, 7. Skin, 8. Dermis, 9. Muscle, 10. Dull tip of sham needle, 11. Sharp tip of real needle

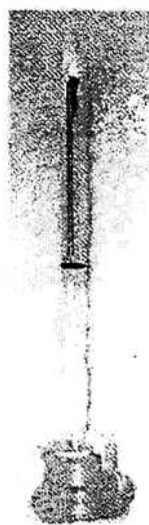


Figure 1b Park Sham Device

and maintains the sterility of a genuine acupuncture needle. (Figure 1) The original sham needle had a copper handle, so we developed a version with a stainless steel handle to match the needles we intended to use in a clinical trial. This paper describes studies that have investigated whether the procedure using the PSD is 1) indistinguishable from the same procedure using real needles in acupuncture naïve subjects, and 2) inactive, where the specific needle sensation (de qi) is taken as a surrogate measure of activity. This needle sensation is believed to be an essential

element of acupuncture treatment representing the needle's effect in controlling 'the flow of energy (qi) in the body'.¹⁶ According to the principles of Chinese medicine, energy flow is intimately related to health.

The needle sensation (de qi) is an unusual experience which may be difficult to describe but which acupuncture practitioners claim to recognise. Few researchers have explored the concept, but in a seminal study, Vincent found that patients undergoing acupuncture described the sensation with the terms 'pulling', 'numb', 'heavy', 'dull' and 'aching', together with 'spreading', 'radiating' and 'pulsing' and general pain descriptors.¹⁷ However, it was not clearly established that these particular sensations were directly associated with the de qi phenomenon. We therefore planned to involve experienced acupuncturists to judge the presence or absence of de qi.

Methods

Park Sham Device

The PSD is illustrated in figure 1. Each PSD is composed of a needle (either real or sham) and a base unit comprising a guide tube and the Park tube with a foot-plate that attaches to the skin with double-sided adhesive tape. The real and the sham needles are the same size (0.35mm x 70mm), and manufactured in stainless steel (Dong Bang Acupuncture Inc., Korea).

Study 1) Is PSD indistinguishable from the same procedure using real needles in acupuncture naïve subjects?

Patients admitted to hospital with an acute stroke were recruited to a patient- and assessor-blinded randomised controlled trials (RCT) that compared the effect on rehabilitation of real acupuncture with that of the sham acupuncture. We report here the findings of the tests of blinding of the first 58 patients who were treated between November 1999 and April 2001. After 12 sessions of treatment, patients were asked by the assessor 'When you volunteered for the trial, you were informed that you had an equal chance of receiving acupuncture or sham (pretend) acupuncture. Which acupuncture do you think you received?' The choice of responses was 'acupuncture, sham acupuncture and don't know'. The study received approval from the North & East Devon Research Ethics Committee and informed consent was obtained from all patients.

Study 2) Is PSD inactive, where the specific needle sensation (de qi) is taken as a surrogate measure of activity?

Subjects

Volunteers were recruited to a subject and assessor-blinded study through the local press and by contacting students through university staff. Healthy volunteers aged over 16 years, who had not previously received acupuncture and were lifelong English speakers, were invited to take part in the study. Volunteers who were pregnant, taking analgesic medication, or had a bleeding disorder, or who had current pain or local injury preventing access to the acupuncture point were excluded. Volunteers were requested not to take analgesics or consume alcohol within the 8 hours prior to the study. Volunteers were informed of the potential risks of acupuncture and provision of indemnity, and were then asked to give written consent.

A sample size of 30 per group was expected to have 90% power at 0.05 significance level (two-sided) to detect the difference between the two groups based on the assumption that de qi is felt by 35% of the sham-needle group and 75% of the real needle group (from clinical experience). The North & East Devon Local Research Ethics Committee approved the study protocol.

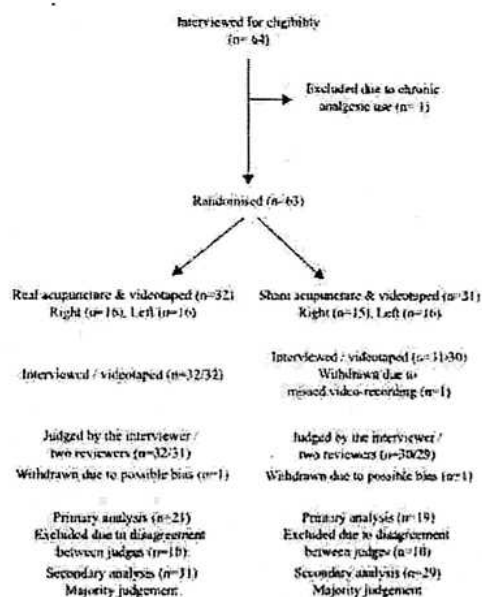


Figure 2 Flow diagram of the progress through study 2

Procedure

A flow diagram of the study is given in figure 2. Volunteers were told that the study aimed to investigate the feelings produced by two different types of acupuncture needle. They were then randomised to receive either real or sham acupuncture at the Hegu (LI4) point (between first and second metacarpals) in either the left or right hand. Random numbers were created using block randomisation, and sealed in opaque envelopes by a person not involved with the study. The envelopes were opened only after consented volunteers were introduced to the acupuncturist (JP), who was the sole person to know the intervention allocation. Volunteers lay supine on the clinic couch. The acupuncturist attached the base unit to the selected point and then tapped the top of the needle shaft: in the case of the real needle, it penetrates skin; in the case of sham needle, it simply presses on the skin and recoils into handle. The guide tube was then slid forward within the Park tube, revealing the handle, which was twisted and moved up and down for 30 seconds. Needles were inserted (real) or positioned (sham) vertical to the skin. After removing the needle and base unit, the acupuncturist covered the point with 3M

micropore tape (to maintain blinding) then left the room. The interviewer (AW, an acupuncture teacher and practitioner with 20 years' experience) entered, remaining blind to the subject's intervention, and interviewed the subject with structured, open questions. The questions asked routinely were: 'You have just been treated with a needle. How would you describe the sensations? What did it feel like? Did it feel like anything you have felt before? Where did you feel these sensations – at the place itself or anywhere else?' When volunteers had difficulty describing sensations, they were encouraged with the following options: 'You can use any words you like. Perhaps it would help if you told me what it felt Dr P was doing to you?' The interviewer recorded his judgement on whether or not the volunteer had experienced de qi or not according to his/her responses. Finally, the subject completed a questionnaire, modified from Vincent (Vincent 1989), consisting of 25 sensations to be scored on a 4-point scale from 0 (none) to 3 (strong). In addition, the questionnaire assessed the subject's knowledge of acupuncture, familiarity with the concept of de qi or 'needle sensation', and any difference between his/her expectations and actual experience of acupuncture. Throughout the needling and interview procedure, the subject's face and upper body were video-recorded.

Measures and analysis

The primary outcome measure was the consensus judgement of three experts of whether or not the volunteer had experienced de qi. If in doubt, a judgement of 'not de qi' was used. One of the experts was always the interviewer; the other two experts judged after watching the videotape of the interview. Experts were identified by three UK professional acupuncture societies (Acupuncture Association of Chartered Physiotherapists, British Acupuncture Council, British Medical Acupuncture Society). Sixteen individuals volunteered, from which 12 reviewers were selected: two from each of the three organisations and six more through randomisation. Each of the 12 reviewers saw 10 individual video recordings. Thus, two reviewers and the interviewer (AW) assessed each of the 60 subjects.

In addition, in order to assess inter-reviewer reliability, a set of 10 recordings was randomly selected and sent to all reviewers. We calculated the approximate inter-rater reliability (IRR) (including approximate 95% confidence intervals) through ANOVA using a two-way random effects model.^{16,19} This method is based on every subject's videotape being judged by each reviewer and the assumption that reviewers were randomly selected (at least without investigator bias) from a population of experts.

The primary analysis, as pre-determined by the protocol, was performed using only those subjects on whom a consensus was achieved by all three assessors. Subsequently a further, exploratory analysis was performed on all subjects using the majority judgement. Data were analysed by calculating relative risk and its 95% confidence interval using SPSS 9.0.

Results

In study 1, 58 patients (M/F=30/28), aged from 38 to 87 years old were treated during the study period. No patients in either group believed they had been treated with the sham needle (Table 1).

Table 1 Responses of stroke patients in sham-controlled trial who were asked 'Did you receive real or sham acupuncture?' (Study 1)

	Needle	Real	Sham	Total
Answer				
Real		11	9	20
Sham		0	0	0
Don't know		18	20	38
Total		29	29	58

In study 2, 63 volunteers were recruited for the trial between November 2000 and February 2001, of whom three were withdrawn for the following reasons: video recording unobtainable for one volunteer due to a technical problem; possible bias identified with two volunteers who had previously attended a lecture on acupuncture research given by the second author (AW). Therefore 60 volunteers were included in the analysis (mean age 37.1 (SD 16.7) years, male/female ratio 14/46). Baseline variables were similar in the two groups (Table 2).

All three assessors agreed on the presence or absence of de qi in 40 (66.67%) of the 60 volunteers. The results in this group (Table 3)

Table 2 Baseline characteristics of volunteers (Study 2)

Characteristic	n	Treatment group		Total
		Real acupuncture (n = 31)	Sham acupuncture (n = 29)	
Age: Mean (SD), [Range]	60	35.5 (15.6), [18-77]	38.7 (16.9), [18-75]	37.1 (16.7), [18-77]
Male/Female	60	8/23	6/23	14/46
Volunteers Included				
Knowledge of acupuncture *	60	1.10±0.30	1.14±0.44	-
Comparison of expectation vs experience †	60	2.16±2.73	2.38±1.97	-
Familiarity with De qi or 'needle sensation ‡	60	None	None	None
Subjects in Consensus group				
Age, mean (SD), [Range]	40	N=21 35.7 (14.7), [18-77]	N=19 38.9 (17.6), [18-68]	37.2(16.0), [18-77]
Male/Female	40	5/16	4/15	9/31

* 3 point-scale, 1 = slight, 2 = moderate, 3 = extensive.

† 6 point-scale, 0 = not at all what the subject expected, 5 = Exactly what the subject expected.

‡ Yes/No binomial scale.

Table 3 Contingency table of 40 subjects in primary analysis

Judgement Needle type	De qi	Non-De qi	Total
Real	17	4	21
Sham	1	18	19
Total	18	22	40

Table 4 Contingency table of 60 subjects in secondary analysis

Judgement Needle type	De qi	Non-De qi	Total
Real	22	9	31
Sham	6	23	29
Total	28	32	60

show that the relative risk of experiencing de qi with real acupuncture to that with PSD was 15.38 (95% CI 2.26 to 104.86). The analysis involving all 60 subjects, based on the majority judgement of the three assessors (Table 4), indicates that the relative risk of experiencing de qi with real acupuncture to that with PSD was 3.43 (95% CI 1.62 to 7.24).

Inter-reviewer reliability (IRR), based on the judgements on 10 randomly selected subjects, of all 13 reviewers was 0.52 (95% CI 0.19 to 0.61). In the 6 subsets each judged by 3 assessors, the IRRs were 0.59 (0.39 to 0.92), 0.49 (0.28 to 0.69), 0.63 (0.44 to 0.78), 0.76 (0.61 to 0.89), 0.21 (0.02 to 0.38) and 0.75 (0.60 to 0.87), respectively.

There was no noticeable adverse event other than minor bleeding or discomfort around the point of needle insertion in a few subjects.

Volunteers who received sham acupuncture found it met their expectations (Table 2) to the same degree as did those who received real acupuncture, providing further evidence that the two are indistinguishable.

Discussion

These data suggest that a new sham acupuncture device, the PSD, is both indistinguishable from real acupuncture and inactive, thus meeting important criteria for use in sham-controlled trials.

Since not a single patient recognised that they had received treatment with PSD, it seems robustly indistinguishable. However, this finding may be limited by the facts that these patients may have had abnormal sensation because of their recent stroke and were possibly disorientated from hospital admission. In addition, patients involved in clinical trials may be more likely than healthy volunteers to believe that they have received genuine treatment. Experience in other clinical settings will be required for further validation of the device.

The PSD's lack of activity is supported by the finding that healthy volunteers were more likely to experience de qi, used here as a surrogate for clinical effectiveness, with a real acupuncture needle than with the PSD. This result was confirmed when using the (weaker) majority opinion of the judges. However, one limitation of

these findings is that the responses of the volunteers will be influenced by the questions asked. The volunteers were representative of English speaking populations, and the result of this study can be generalised to other similar groups. Experience at other acupuncture points and with other practitioners will be necessary in order to confirm that de qi is not generated in other settings. It will be noted that one out of 18 volunteers in the PSD group was judged to have experienced de qi, and 6 out of 28 using the majority data. This may be due theoretically to measurement error, to observer error or to the fact that the PSD on the acupuncture point may produce a sensation which is described in similar terms to de qi yet is somehow different. In any of these cases, PSD still represents a useful sham procedure.

The inter-rater reliability (IRR) among 13 reviewers was 0.52 (95% CI 0.19 to 0.61), which is considered reasonable in a medical context.²⁰ This is lower than anticipated at the outset of the study, and is likely to be due to differences in the education and practice of the reviewers. The acupuncture experts reached complete consensus in only two-thirds of cases. Of the remaining 20 cases, the two judges who used video disagreed in 12 cases, while in 8 cases they agreed with each other but disagreed with the original interviewer. It is notable that there was a low IRR between two particular video judges. This level of disagreement might have several explanations. Sixteen experts from three professional organisations may not be representative of the whole population of acupuncturists, and the judgement may contain the complex combination of judgement itself and judges' variation in knowledge and experience. Error may have been introduced by the limitation of the technology i.e. video. In addition, this study was conducted at only one acupuncture point and on only one occasion.

The sham needle has been tested previously,¹⁷ but we believe that the methods are open to criticism. Volunteers in that study were asked if they could feel 'a dull sensation' which may not accurately reflect de qi. In addition, the authors tested whether the needles were distinguishable by asking volunteers whether they believed the needle had 'penetrated', which is only part of the acupuncture experience. In a clinical study

using the sham needle,¹⁸ patients were asked four credibility questions but the sample size was insufficient to test equivalence of the procedures.

Important improvements in PSD (Figure 1) compared with the earlier design include the avoidance of potential contamination of the needle, and the smaller 'footprint' on the skin. However, further investigation of this device will be required to establish its performance in the hands of different clinicians, in different acupuncture points particularly with a different anatomical configuration (e.g. in more curved skin, or in hairy regions), and with sham needles made to match different manufacturers' styles.

We conclude that the newly developed PSD is valid as a control procedure for trials of the efficacy of acupuncture, or, more specifically, the needle penetration aspect of treatment. Our second conclusion arising from this work is perhaps more profound. Naïve volunteers cannot distinguish between real needles and the sham device, yet are much more likely to experience de qi with the former than the latter. This finding gives some support to the existence of the phenomenon of de qi, one of the major concepts underlying traditional Chinese acupuncture. We believe that the question of whether or not de qi can be scientifically demonstrated to exist deserves further rigorous study.

Acknowledgements

The authors would like to thank the following: the volunteers for participating; Anna Hari for helpful comments on the design; Drs Anthony Hemsley, John Chambers and Vanessa Stedman for assessing the patients; Dr James Goodwin, St. Loye's School, and Guy Faulkner for helping advertise and conduct the study; Messrs Mark Bovey, Peter Mole, David F Mayor, Peter White, Drs Anthony Campbell, Richard James, Michael Cummings, Hywel Watkin, Ms Rachel Peckham, Ms Nadia Ellis, and Mrs Val Hopwood for reviewing the video; and Mr Michael Potter, Drs Ceri Salisbury, Jacqueline Filshie, and Frankie Reid for volunteering to review; Dongbang Acupuncture Inc., Korea for donating Park Sham needles. The first author (JP) acknowledges with gratitude the support of The British Chevening Scholarship and Dr Youngkwon Kim.

Reference list

1. Gaw AC, Chang LW, Shaw L-C. Efficacy of acupuncture on osteoarthritic pain. A controlled, double-blind study. *N Engl J Med* 1975;293(8):375-8.
2. Godfrey CM, Morgan P. A controlled trial of the theory of acupuncture in musculoskeletal pain. *J Rheumatol* 1978;5(2):121-4.
3. Moore ME, Berk SN. Acupuncture for chronic shoulder pain. An experimental study with attention to the role of placebo and hypnotic susceptibility. *Ann Intern Med* 1976;84(4):381-4.
4. Jensen LB, Melsen B, Jensen SB. Effect of acupuncture on headache measured by reduction in number of attacks and use of drugs. *Scand J Dent Res* 1979;87(5):373-80.
5. Gallacchi G, Muller W, Plattner GR, Schnorrenberger CC. [Acupuncture and laser treatment in cervical and lumbar syndrome]. *Schweiz Med Wochenschr* 1981;111(37):1360-6.
6. Junnila SY. Long-term treatment of chronic pain with acupuncture. Part I. *Acupunct Electrother Res* 1987;12(1):23-36.
7. Henry P, Baille H, Dartigues JF, Jogeix M. Headache and acupuncture. In: Pfaffenrath V, Lundberg PO, Sjaastad O, editors. *Updating in Headache*. Berlin: Springer Verlag; 1985.
8. Petric J, Hazleman B. Credibility of placebo transcutaneous nerve stimulation and acupuncture. *Clin Exp Rheumatol* 1985;3(2):151-3.
9. Lao L, Bergman S, Hamilton GR, Langenberg P, Berman B. Evaluation of acupuncture for pain control after oral surgery: a placebo-controlled trial. *Arch Otolaryngol Head Neck Surg* 1999;125(5):567-72.
10. White AR, Resch KL, Chai JC, Norris CD, Modi SK, Patel JN *et al*. Acupuncture for episodic tension-type headache: a multicentre randomised controlled trial. *Cephalalgia* 2000;20(7):632-7.
11. Gosman-Hedstrom G, Claesson L, Klingenstein U, Carlsson J, Olausson B, Frizell M *et al*. Effects of acupuncture treatment on daily life activities and quality of life: a controlled, prospective, and randomised study of acute stroke patients. *Stroke* 1998;29(10):2100-8.
12. Karst M, Rollnik JD, Fink M, Reinhard M, Piepenbrock S. Pressure pain threshold and needle acupuncture in chronic tension-type headache--a double-blind placebo-controlled study. *Pain* 2000;88(2):199-203.
13. Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. *Lancet* 1998;352(9125):364-5.
14. Kleinhenz J, Streitberger K, Windeler J, Gussbacher A, Mavridis G, Martin E. Randomised clinical trial comparing the effects of acupuncture and a newly designed placebo needle in rotator cuff tendinitis. *Pain* 1999;83(2):235-41.
15. Park J, White A, Lee H, Ernst E. Development of a new sham needle. *Acupunct Med* 1999;17(2):110-2.
16. The Academy of Traditional Chinese Medicine. *An Outline of Chinese Acupuncture*. Peking: Foreign Languages Press, 1975.
17. Vincent CA, Richardson PH, Black JJ, Pither CE. The significance of needle placement site in acupuncture. *J Psychosom Res* 1989;33(4):489-96.
18. Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* 1979;86(2):420-8.
19. Fleiss JL, Shrout PE. Approximate interval estimation for a certain intraclass correlation coefficient. *Psychometrika* 1978;43(3):259-62.
20. Cicchetti DV, Sparrow SA. Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic* 1981;86(2):127-37.