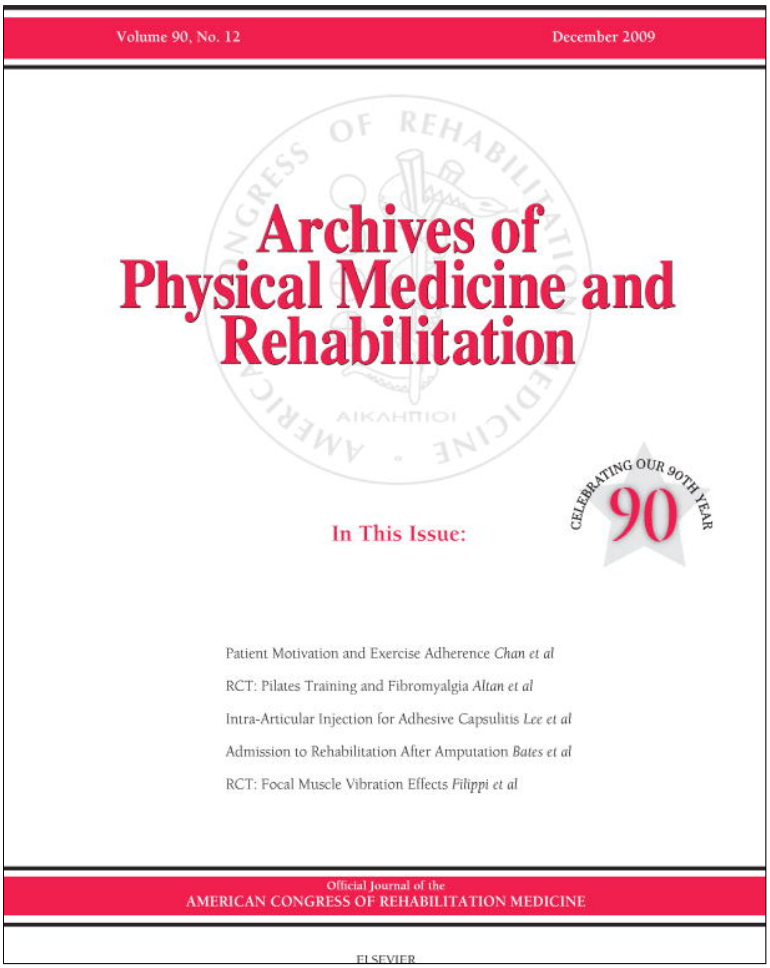

ORIGINAL ARTICLE

Discrimination of Real and Sham Acupuncture Needles Using the Park Sham Device: A Preliminary Study

Chee-Wee Tan, PhD, Lauren Christie, BSc(Hons), Véronique St-Georges, BSc(Hons), Nicola Telford, BSc(Hons)

ABSTRACT. Tan C-W, Christie L, St-Georges V, Telford N. Discrimination of real and sham acupuncture needles using the Park sham device: a preliminary study. Arch Phys Med Rehabil 2009;90:2141-5.

Objective: To evaluate the blinding effectiveness of the Park sham acupuncture device using participants' ability to discriminate between the real and sham acupuncture needles.

Design: The design was a yes-no experiment. Judgments were made on whether the real or sham acupuncture needle was administered.

Setting: University laboratory.

Participants: Healthy, acupuncture-naive university students and staff (N=20; median age, 22y; range, 18–48y) recruited through convenience sampling.

Interventions: Participants made yes-no judgments on whether the real or sham needle was administered to 8 acupoints (4 traditional and 4 nontraditional) along the Pericardium meridian (Pericardium 3 to Pericardium 6) on the dominant forearm.

Main Outcome Measures: The accuracy index, d', of participants' ability to discriminate between the real and sham needles (discriminability) was computed for the traditional alone, the nontraditional alone, and a combination of both types of acupoints.

Results: The participants' d' between the real and sham needles was not statistically significant from d' equal to 0 for the combined traditional and nontraditional acupoints comparison and the nontraditional acupoints alone comparison (combined, $t_{19}=1.20$, $P=.25$; nontraditional, $t_{19}=.16$, $P=.87$). However, the participants' d' was statistically significant from d' equal to 0 for the traditional acupoints comparison ($t_{19}=2.096$, $P=.049$).

Conclusions: The Park sham acupuncture device appears to be effective in blinding participants to real acupuncture intervention when it is applied to the nontraditional acupoints and when traditional and nontraditional acupoints are combined on the forearm along the pericardium meridian. However, the sham device does not appear to blind participants effectively when traditional acupoints alone are used for the same context.

Key Words: Acupuncture; Placebos; Rehabilitation; Signal detection, psychological; Validation studies at topic.

IN ORDER TO CONDUCT randomized controlled trials for examining the effects of acupuncture, a viable and credible sham acupuncture device is required. In recent years, blunt acupuncture needles that retract into the handles were invented as a possible solution for blinding research participants. The 2 most commonly used types of sham devices that can be used to administer the sham acupuncture needle are the Streitberger sham device[1] and the Park sham device.[2] Randomized controlled trials examining the effects of acupuncture have started to use these sham devices and needles.[3-5]

Studies examining the blinding effectiveness of the sham devices usually investigate similarities between the sensory intensities elicited by the sham and real needles that are felt by the participants.[1] If the sensory intensities to both needle types are not statistically different from each other, then it is indirectly inferred that the sham device is capable of blinding the participant. However, Takakura and Yajima[6] commented that the presence of a needle penetration sensation was not indicative of whether the participants believed a real needle was administered. This meant that participants' reports of sensation were not verifiable and therefore did not provide information about the blinding effectiveness of the sham needles.

Instead of sensory intensities, the proportion of correct and incorrect judgments made by participants on whether the real or the sham needle has been administered may be obtained. This is analogous to the situation in which a clinician makes a diagnostic judgment about whether a radiograph shows the presence (or absence) of pathology. This method generates 4 separate outcome measures representing the participant's accuracy in judgment (table 1). The 4 outcome measures are (1) the true positive rate (the proportion of times the participant correctly judges the real needle to be the real needle), (2) the false negative rate (the proportion of times the participant incorrectly judges the real needle to be the sham needle), (3) the false positive rate (the proportion of times the participant incorrectly judges the sham needle to be the real needle), and (4) the true negative rate (the proportion of times the participant correctly judges the sham needle to be the sham needle). The comparison of 4 separate outcomes between different studies contributes a degree of difficulty and effort in determining the participants' ability to discriminate between the real and sham needles. This level of analytical complexity may be simplified by combining both the true positive rate and the false positive rate into a single outcome measure, d'.[7] The true positive rate and false positive rate are analogous to the outcome measures of sensitivity and 1-specificity, respectively, in the study of diagnostic accuracy.[8] To obtain d', the true positive rate is subtracted from the false positive rate and then

List of Abbreviations

CI	confidence interval
d'	discriminability
PC	pericardium

From the School of Health Sciences, Queen Margaret University Edinburgh, Edinburgh (Tan); South Grange Care Centre, Barchester Healthcare, Monifieth (Christie), United Kingdom; Back in Motion Rehab Inc, Vancouver, BC, Canada (St-Georges); Department of Physiotherapy, Roodlands Hospital, National Health Service Lothian, East Lothian, United Kingdom (Telford).

Supported by Queen Margaret University Edinburgh.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Reprint requests to Chee-Wee Tan, PhD, Queen Margaret University Edinburgh, Queen Margaret University Dr, Edinburgh EH21 6UU, United Kingdom, e-mail: *ctan@qmu.ac.uk*.

0003-9993/09/9012-00504$36.00/0
doi:10.1016/j.apmr.2009.08.142

Table 1: Stimulus-Judgment Matrix and the Associated Outcomes

	Participants' Judgments	
Needle Administered	Real	Sham
Real	True positive	False negative
Sham	False positive	True negative

standardized into a z score. The z score is a measure by which the difference in numerical value of the outcome measure and the mean of the distribution is divided by the SD of the distribution. The use of the standardized score allows comparison between different studies. The interval of d' values ranges from 0 to infinity. When an ideal sham device offers complete blinding, d' is 0.[7] This means that the real needle and sham needle are completely indistinguishable by the participants.

Besides the type of outcome measure used, previous sham device evaluation studies exhibit 2 other limitations. First, previous studies have used only 1 acupoint for evaluating the sham devices.[1,9] Multiple acupoints, which may include both traditional and nontraditional acupoints, should ideally be used to simulate clinical practice better. Second, previous studies have also used between-groups designs to compare real and sham needle discrimination using the sham devices. Within-subject designs may be more appropriate for sham device evaluation because this allows each participant to compare directly and discriminate between the real and sham needles.

Based on the limitations identified for previous sham evaluation studies, the primary objective of this preliminary study was to evaluate the blinding potential of the Park sham device using a within-subjects study design and the outcome measure of d'. Therefore, it was hypothesized that the participants' d' was not statistically significantly larger than d' equal to 0 (unable to discriminate between the needle types). The secondary objective of this study was to examine whether the use of traditional points and nontraditional points would affect the blinding ability of the sham devices. It was hypothesized that the participants' d' between the needle types on either the traditional or nontraditional acupoints was not statistically significantly larger than d' equal to 0.

METHODS

Participants

A favorable ethical opinion for this study was obtained from the Queen Margaret University Research Ethics Committee. Students and staff of the University were recruited as participants using convenience sampling. The inclusion criteria for this study were (1) age of 18 years or more, (2) naive to acupuncture intervention, and (3) able to provide informed consent. The exclusion criteria were (1) the presence of medical conditions that caused anesthesia to the dominant upper limb, (2) any wounds or injury to the upper limb, (3) needle phobia, (4) consumption of potentially analgesic medications 24 hours before the study procedures, and (5) pregnancy. All participants provided written informed consent for this study. Participants were allowed to withdraw from the study at any juncture of the study without providing reasons for doing so.

Experimenters

A research assistant recruited and inducted participants into the study, and the first author (C.-W.T.) administered the needles to all participants.

Materials

The Park sham acupuncture device was used for administering the sham and real needles (fig 1).[2] The device consists of a ring-base unit and a special oversized tube (Park tube). The ring-base of the device is kept in place on the participant's skin using double-sided tape. The internal circumference of the ring-base fits tightly around the Park tube. The standard guide tube that is included with the real acupuncture needle product slides into place inside the Park tube. Precut guide tubes (55mm) were made and fitted into the Park tube for predetermining the penetration depth of the real needles. This achieved the penetration depth of approximately 10mm for the real needles. The same length of guide tubes was also used for the sham needles. Both types of needles were of the same dimensions (0.25mm $\times$ 40mm) and manufactured by Dong Bang Acupuncture, Inc.[a]

Acupuncture Points

Eight acupoints along the PC meridian were chosen for this study (fig 2). Both traditional and nontraditional acupoints were selected to represent the Traditional Chinese Medicine and Western Medical Acupuncture approaches to acupuncture treatment. The 4 traditional acupoints and their designated needle types were (1) PC3 (sham), (2) PC4 (real), (3) PC5 (real), and (4) PC6 (sham). The 4 nontraditional acupoints and their designated needle types were (1) $PC3_{1.5}$, 1.5 cun distal to PC3 along the meridian connecting PC3 and PC7 (real); (2) $PC3_{3.0}$, 3.0 cun distal to PC3 along the meridian connecting PC3 and PC7 (real); (3) $PC3_{4.5}$, 4.5 cun distal to PC3 along the meridian connecting PC3 and PC7 (sham); and (4) $PC3_{6.0}$, 6.0 cun distal to PC3 along the meridian connecting PC3 and PC7 (sham).

Procedures

The participants verbally indicated their arm dominance. Afterward, the participant was positioned in inclined lying on a plinth with the elbow of the dominant arm resting comfortably on a table in a slightly flexed position. C.-W.T. attached the Park sham devices over the 8 chosen acupoints on the dominant forearm.

The sequence of needle administrations for the 8 acupoints was randomized using an online randomization generator.[10]

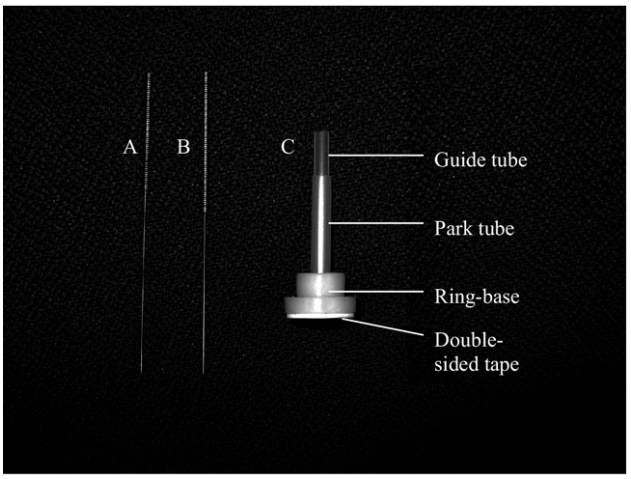

Fig 1. The photograph shows (A) a real acupuncture needle, (B) a sham acupuncture needle, and (C) the Park sham device.

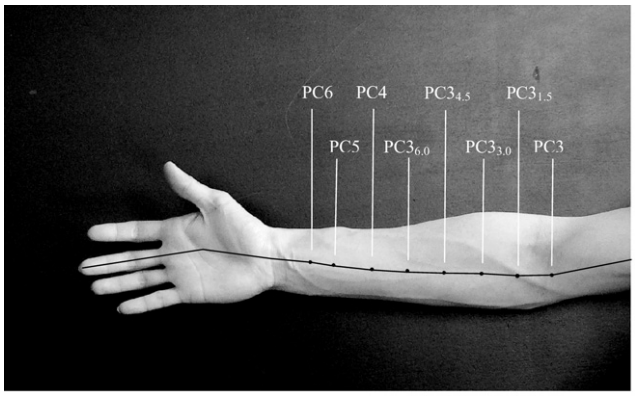

Fig 2. The PC meridian acupoints on the forearm. The figure shows 4 traditional acupoints (PC3, PC4, PC5, PC6) and 4 nontraditional acupoints (PC3$_{1.5}$, PC3$_{3.0}$, PC3$_{4.5}$, PC3$_{6.0}$). The subscript after the acupoint nomenclature denotes the number of cun along the meridian distal to the PC3 point. The line connecting the acupoints represents the course of the PC meridian.

The precut guide tubes were inserted into the Park tube. C.-W.T. then opened the bubble pack (containing either the sham or real needle) in full view of the participant and took out the needle. C.-W.T. held the guide tube firmly and slightly depressed it down onto the skin before needle insertion. This procedure simulated the clinical practice of using the needle guide tube to stretch the skin over the acupoints before needle insertion. The needle was carefully placed into the guide tube. A quick, gentle tap was given to the proximal end of the needle handle for insertion or simulating needle insertion. No needle handles were twirled during or after insertion. C.-W.T. asked the participants, "Do you think the real acupuncture needle has been administered?" The participants answered yes or no. The research assistant recorded the participant's judgment on a score sheet. This denoted the end of 1 administration. This is called a yes-no experiment[7,11] and was performed for all 8 acupoints. Each participant received 8 administrations, providing a total of 8 judgments. The entire procedure for each participant lasted approximately 30 minutes.

Analysis

The participants' judgments were classified into true positives, false positives, false negatives, and true negatives. A d' was computed for each participant by subtracting the z score for the false positive rate from the z score for the true positive rate. Separate d' values were computed for the traditional acupoints, nontraditional acupoints, and a combination of both acupoint types using the same data set. One sample t tests at α equal to .05 were performed to compare the d' generated for the traditional acupoints alone, nontraditional acupoints alone, and both acupoint types with d' equal to 0. The proportion of correct judgments for each acupuncture point was also generated to obtain a breakdown of participants' judgments.

RESULTS

Participants

Twenty healthy volunteers (14 women and 6 men) took part in the experiment with no participant dropouts. The participants' median age was 22 years (range, 18–48y).

Discriminability of Participants

The judgments for all participants were categorized into the 4 judgment categories. Table 2 shows the stimulus-judgment matrix of all judgments and the matrices of participants' judgments separated by the type of acupoints (both acupoint types, traditional alone, or nontraditional alone).

The mean ± SD d' (95% CI) of participants for both acupoint types was .43±1.59 (–.32 to 1.17). The mean ± SD d' values (95% CI) for the traditional acupoints and nontraditional points were .63±1.34 (.001 to 1.257) and .05±1.35 (–.58 to .68), respectively.

The d' for both acupoint types and nontraditional acupoints alone were not statistically significantly larger than d' equal to 0 (both, $t_{19}=1.20$, $P=.25$; nontraditional, $t_{19}=.16$, $P=.87$). The d' for traditional acupoints alone was statistically significantly larger than d' equal to 0 ($t_{19}=2.096$, $P=.049$).

Judgment Mapping of Acupoints

A further breakdown of the participants' judgments was performed to provide a more detailed analysis. The correct judgments to the type of needle administered were tallied for each acupoint and plotted on a graph for exploration of any patterns within the data. Figure 3 shows the proportion of judgments that the participants correctly identified for each acupoint. Out of the 8 acupoints, 3 acupoints had proportion of correct judgments at or less than 50%. All of these 3 acupoints were nontraditional points.

DISCUSSION

Summary of Study Findings

This preliminary study found that the participants appeared to be able to discriminate between the sham and real needles when traditional acupoints alone were used. However, the participants' discriminability between the needle types was not statistically significant from d' equal to 0 when both needle types and the nontraditional acupoints alone were used.

Discriminability of Participants

The d' of .43 for both acupoint types in this study is 2.3 times larger than that calculated from the data of Park et al[12] ($d'=.19$). There are 2 probable reasons that our d' was larger than that in the study by Park.[12] First, we used a within-subject design, and the study by Park[12] used a between-groups design. This meant that the participants in our study might have used sensations felt from previous needle administrations to inform judgments of subsequent needle administrations. This process would potentially improve the participant's d' to the type of needle administered.

Second, it may be possible that our method of needle administration provided more visual or tactual cues compared with the needle administration method used by Park et al.[12] The cues may consist of a gestalt of sensations that inform the participant consciously or unconsciously in making the judg-

Table 2: Stimulus-Judgment Matrix for Traditional and Nontraditional Acupoints

Needle Administered	Participants' Judgments (Proportions)					
	Nontraditional		Traditional		Both	
	Real	Sham	Real	Sham	Real	Sham
Real	.28	.72	.75	.25	.63	.37
Sham	.26	.74	.38	.62	.45	.55

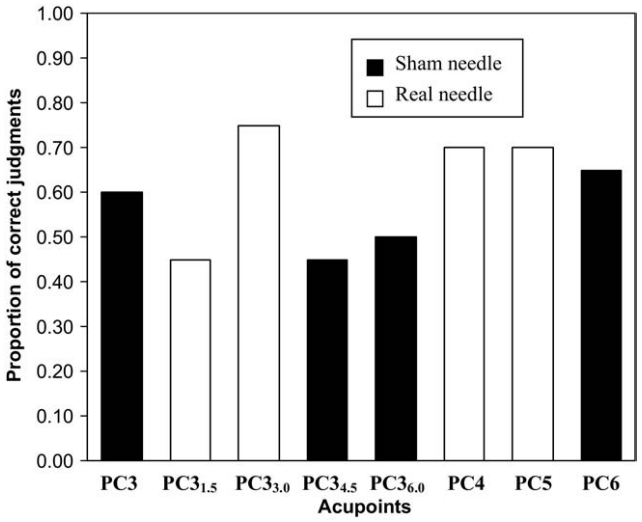

Fig 3. Judgment mapping for PC3 to PC6 acupoints. The bars represent the proportion of correct judgments made by the participants for the acupoints. The nontraditional acupoints generally have a lower percentage of correct judgments than the traditional acupoints.

ments. It is also possible that some participants may have perceived subtle behavioral cues unconsciously displayed by the acupuncturist despite the standardized protocol.

It was interesting to note that d' differs between traditional and nontraditional acupoints. A further breakdown of the correct judgments obtained for each acupoint showed that only 1 nontraditional acupoint generated a high number of correct judgments (fig 3). Participants provided 75% correct judgments for acupoint PC3$_{3.0}$ compared with PC3$_{1.5}$, PC3$_{4.5}$, and PC3$_{6.0}$, which had 45%, 45%, and 50% correct judgments. It is unclear whether this indicates sensory and physiologic differences between traditional and nontraditional acupoints. It is known that sensation thresholds differ between different bodily regions.[13,14] The acupoints used in this study traverse the C5 to C7 dermatomal and C6 to T1 myotomal regions. Therefore, it is possible that tactile thresholds for different regions influenced the differences in d' observed between the acupoint types. However, this study did not examine the cutaneous or muscular sensation thresholds. Therefore, the correlation between the sensation thresholds and the d' observed between the acupoint types cannot be established.

Implications for Acupuncture Research

Based on the results, the Park sham device appeared to blind the participants for the nontraditional acupoints or for both acupoint types used on the forearm along the PC meridian. Because most acupuncture studies use either both traditional and nontraditional acupoints or traditional acupoints alone, this suggests that the sham device may not provide blinding for the traditional acupoints alone approach on the forearm. The results of this study should be interpreted with caution when extrapolated to acupoints located on other body regions because other bodily regions may be dissimilar in terms of structural, sensory, and perceptual characteristics. Another factor that may contribute to participants' d' is the visual proximity of the body region receiving the intervention, which was not explored in this study.

Study Limitations

Each acupoint in this study was designated to be administered either a real or a sham needle. The data collected from this design were insufficient to generate comparative d' values for each individual acupoint. The initial decision to adopt this design was based on reducing the complexity of randomization.

Standard clinical acupuncture practice usually consists of needles administered at multiple acupoints on multiple meridians for different bodily regions. Mapping of d' to the type of needle (real or sham) on the various acupoints and meridians is needed to inform future studies that use sham devices.

The d' outcome measure used in this study is generated using participants' judgments. This outcome measure is descriptive, and its associated theoretical framework does not assume any underlying mechanisms contributing to the perceptual performance of the participants.[15] Further work is required in examining the different dimensions of input (physiologic and psychologic) that may influence a participant's judgment accuracy for the type of needle administered. This information will inform future studies in optimizing the blinding ability of sham acupuncture devices.

CONCLUSIONS

This preliminary study found that the Park sham acupuncture device appears not to be effective in blinding participants for the traditional acupoints alone condition on the forearm along the PC meridian. However, the sham device appears to be effective for blinding participants on the nontraditional acupoints alone condition and when both acupoint types are used. Further studies are needed to elucidate the factors that will help optimize the sham device's blinding effectiveness.

Acknowledgments. We thank Derek Santos, PhD, Cathy Bulley, PhD, and the reviewers for their constructive criticisms for drafts of this article.

References

1. Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. Lancet 1998;352:364-5.
2. Park JB, White A, Lee HJ, Ernst E. Development of a new sham needle. Acupunct Med 1999;17:110-2.
3. Goldman RH, Stason WB, Park SK, et al. Acupuncture for treatment of persistent arm pain due to repetitive use: a randomized controlled clinical trial. Clin J Pain 2008;24:211-8.
4. Jubb RW, Tukmachi ES, Jones PW, Dempsey E, Waterhouse L, Brailsford S. A blinded randomised trial of acupuncture (manual and electroacupuncture) compared with a non-penetrating sham for the symptoms of osteoarthritis of the knee. Acupunct Med 2008;26:69-78.
5. Kennedy S, Baxter GD, Kerr DP, Bradbury I, Park J, McDonough SM. Acupuncture for acute non-specific low back pain: a pilot randomised non-penetrating sham controlled trial. Complement Ther Med 2008;16:139-46.
6. Takakura N, Yajima H. A double-blind placebo needle for acupuncture research. BMC Complement Altern Med 2007;7:31.
7. Macmillan NA, Creelman CD. Detection theory: a user's guide. 2nd ed. Mahwah: Lawrence Erlbaum Associates; 2005.
8. Griner PF, Mayewski RJ, Mushlin AI, Greenland P. Selection and interpretation of diagnostic tests and procedures: principles and applications. Ann Intern Med 1981;94:557-92.
9. Takakura N, Yajima H. A placebo acupuncture needle with potential for double blinding: a validation study. Acupunct Med 2008;26:224-30.
10. Dallal GE. Randomization.com. Available at: http://www.randomization.com. Accessed February 17, 2008.

11. Gescheider GA. Psychophysics: the fundamentals. 3rd ed. Mahwah: Lawrence Erlbaum Associates; 1997.
12. Park JB, White A, Stevinson C, Ernst E, James M. Validating a new non-penetrating sham acupuncture device: two randomised controlled trials. Acupunct Med 2002;20:168-74.
13. Kosek E, Ekholm J, Nordemar R. A comparison of pressure pain thresholds in different tissues and body regions: long-term reliability of pressure algometry in healthy volunteers. Scand J Rehabil Med 1993;25:117-24.
14. Kosek E, Ekholm J, Hansson P. Pressure pain thresholds in different tissues in one body region: the influence of skin sensitivity in pressure algometry. Scand J Rehabil Med 1999;31:89-95.
15. Pastore RE, Crawley EJ, Berens MS, Kelly MA. "Nonparametric" A' and other modern misconceptions about signal detection theory. Psychon Bull Rev 2003;10:556-69.

Supplier

a. Dong Bang Acupuncture, Inc, Room No.606-2, TechnoPark Bldg A, #150 Yatap-dong, Bungdang-gu, Sungnam-si, Kyunggi-do, 463-760, Korea.