Commentary

Methodological advances needed in analysis and interpretation of sham acupuncture validation studies

Jongbae J Park,1 Heejung Bang2

METHODS

Many challenges remain in the evaluation of the clinical efficacy of acupuncture: for example, how to define the particular form of acupuncture used, when there are so many varieties of practice that can be called ‘acupuncture’ and how to choose the control method appropriate to that particular form of acupuncture.1 Perhaps we are still some distance from closing the gap between the rigour of a study (internal validity) and its generalisability and real-world applicability (external validity). In this issue of Acupuncture in Medicine, two reports are presented on the validity of two sham acupuncture methods. In sincere appreciation of the efforts of Takakura et al3 and Tan et al4 in this difficult area, we would like to offer a few general points concerning the validation of acupuncture sham control and to comment on the two studies.

While it is established that the two main objectives of a validation study of a sham control procedure are to test whether it is indistinguishable from the real acupuncture, and to investigate its relative inertness, how much information needs to be disclosed during the informed consent process is crucial, but has not generally been agreed. Tan et al4 told the study participants: ‘this study will investigate whether people can tell the difference between a real acupuncture needle and a fake acupuncture needle with a blunt tip, using a small device’…’ This question seems to be inviting participants to look out for a fake needle with a blunt tip, and thus raises excessive suspicions about the intervention. The basis of informed consent is to ensure the ethical acceptability of the study, of which two major components are the participants’ safety and autonomy. We believe that, in a study that tests whether fake and real needles are indistinguishable, the scientific purpose of the study can be met in an ethical manner, still ensuring the safety of participants, by withholding explicit information that seems likely to bias the response, such as the words ‘fake needle’; we base our view on the assumption that distinguishing a real acupuncture needle from a fake one is significantly easier than doing the same for a real drug or tea, for example.

The real acupuncture in an experimental setting is expected to represent the acupuncture in the real-world that the study aims at testing. The distinctive sensations, which are an essential part of acupuncture, indicate that treatment has been received. If acupuncture study participants (or acupuncturists) ‘randomly guess’ whether they received (or applied) the real needle or not, then we should start to question whether the acupuncture was properly performed and doubt its external validity. Takakura et al’s previous study concluded that the newly developed no-touch control needles are effective for practitioner-blinded studies; acupuncturists incorrectly identified 66% of the new control needles. We computed the practitioner-blinding index (BI) (95% CI) using a validated statistical method by Bang et al. This BI is directly interpreted as the percentage of unblinding beyond chance and can capture differences in blinding between study groups. The index ranges from −1 to 1; BI=1 and 0 mean complete unblinding and random guessing, respectively. BI=−1 means complete opposite guessing that is, 100% participants guessed real intervention while they received sham treatment. The BIs of the real needles, no-touch and skin-touch control needles are 0.08 (0.26 to −0.10), −0.15 (0.03 to −0.35) and −0.41 (−0.25 to −0.57), respectively. With these results, we can interpret that 8%, 15% and 41% of those who applied the real needles, no-touch and skin-touch control needles guessed they penetrated with a real needle. Ironically, a larger percentage of those who used either the no-touch needle or the skin-touch needle guessed they used a real needle. This interpretation leaves us with an additional question: was the real needleling performed correctly, and whether it represents acupuncture needling in the real-world. While random guessing in the real group receiving the statin drug is not an issue, it may be an issue in the real acupuncture group.

Besides the confusion between validity and credibility, we often find the analysis of discrimination data misinterpreted and misleading conclusions drawn. Tan et al4 reported that discrimination accuracy between the real and sham needles was not statistically significant from chance level for the points in the lower limb, yet it was significant for the points in the upper limb. The researchers concluded that the sham device tested is more likely to successfully blind participants in differentiating in the lower limb than the upper limb. While random guessing in the real group receiving the statin drug is not an issue, it may be an issue in the real acupuncture group.

1Asian Medicine & Acupuncture Research, Physical Medicine and Rehabilitation, University of North Carolina at Chapel Hill, North Carolina, USA
2Division of Biostatistics and Epidemiology, Weill Medical College of Cornell University, New York, USA
Correspondence to Jongbae J Park, Asian Medicine & Acupuncture Research, Physical Medicine and Rehabilitation, University of North Carolina at Chapel Hill, CB #7200, Chapel Hill, North Carolina 27599-7200, USA: jongbae_park@med.unc.edu

5 This BI is directly interpreted as the percentage of unblinding beyond chance and can capture differences in blinding between study groups. The index ranges from −1 to 1; BI=1 and 0 mean complete unblinding and random guessing, respectively. BI=−1 means complete opposite guessing that is, 100% participants guessed real intervention while they received sham treatment. The BIs of the real needles, no-touch and skin-touch control needles are 0.08 (0.26 to −0.10), −0.15 (0.03 to −0.35) and −0.41 (−0.25 to −0.57), respectively. With these results, we can interpret that 8%, 15% and 41% of those who applied the real needles, no-touch and skin-touch control needles guessed they penetrated with a real needle. This interpretation leaves us with an additional question: was the real needleling performed correctly, and whether it represents acupuncture needling in the real-world. While random guessing in the real group receiving the statin drug is not an issue, it may be an issue in the real acupuncture group.

Besides the confusion between validity and credibility, we often find the analysis of discrimination data misinterpreted and misleading conclusions drawn. Tan et al4 reported that discrimination accuracy between the real and sham needles was not statistically significant from chance level for the points in the lower limb, yet it was significant for the points in the upper limb. The researchers concluded that the sham device tested is more likely to successfully blind participants in differentiating in the lower limb than the upper limb. While random guessing in the real group receiving the statin drug is not an issue, it may be an issue in the real acupuncture group.
We would also like to urge the authors to be very cautious with the notion that the participants appeared able to differentiate between real and sham acupuncture needles for the TE acupuncture points in the upper limbs, since the mean P(C) for the upper limb acupuncture points alone and the lower limb acupuncture points alone were 0.63 (95% CI 0.51 to 0.74) and 0.50 (95% CI 0.38 to 0.62) respectively. These figures are very close to overlapping P(C)=0.5, indicating that the real and sham needles are indistinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-
tinguishable by the participant as the real and sham needles are indis-

Takakura et al interpret the results of their previous study\(^3\) as ‘acupuncture-experienced volunteers made equal numbers of correct and incorrect judgments about the skin-touch placebo needle.’ However, the Bs (95% CI) of the real needles and skin-touch needles are 0.57 (0.54 to 0.2), and −0.12 (0.06 to −0.31), respectively. These figures indicate that 37% of real acupuncture recipients believed that they received acupuncture, and 12% of sham needle recipients so believed. While this finding might not be statistically significant with the sample size used, without appropriate sample size estimation, it is hard to know how valuable the interpretation and conclusions are.

The same device was tested again by the same investigators in 2011, and the practitioner Bs (95% CI) of the real needles, no-touch and skin-touch control needles are 0.25 (0.46 to 0.04), 0.58 (0.75 to 0.04) and 0.34 (0.54 to 0.13), respectively.\(^3\) This finding can be interpreted to show that of the sample population, 25% of those who applied the real needle guessed correctly, and 58% and 34% of those who used no-touch needles and skin-touch needles guessed they had applied fake needles. The authors, however, drew the conclusion that the no-touch control needles may be used a blind control for the acupuncture control or to test the physiological effect of the skin-touch needles. We find this logic difficult to follow.

For future reference, we feel obliged to share further methodological concerns about Takakura et al’s study\(^4\): (1) sample bias was created by using those who were familiar with receiving acupuncture from the acupuncture school; (2) the same volunteers served as their own control, which makes it impossible to provide independent data and (3) the study was designed and carried out by the developer, and funded by the joint-owner of the intellectual property of the device.

In summary, defining acupuncture and developing a relevant control within an experimental setting remains challenging. The diverse forms of acupuncture practice, rooted in different histories and cultures, add a great burden in acupuncture research. Further methodological progress is needed in certain areas of validation studies of sham control procedures, including the amount of disclosure to study participants, the variables that need to be collected, analytic design and interpretation strategy.

Acknowledgements The authors thank Selena Beckman-Harned for her assistance in preparing this manuscript. JJP and Selena Beckman-Harned acknowledge the reception of grant support from the Jaseng Medical Foundation, Korea.

Competing interest JJP developed and supplies the Park Sham Device.

Provenance and peer review Commissioned; internally peer reviewed.

Contributors JJP and HB jointly contributed to the draft of this commentary.

Accepted 1 August 2011

doi:10.1136/acupmed-2011-010057

REFERENCES
Methodological advances needed in analysis and interpretation of sham acupuncture validation studies

Jongbae J Park and Heejung Bang

doi: 10.1136/acupmed-2011-010057

Updated information and services can be found at:
http://aim.bmj.com/content/29/3/168

These include:

**References**

This article cites 5 articles, 4 of which you can access for free at:
http://aim.bmj.com/content/29/3/168#BIBL

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/