acupuncture IN MEDICINE

Original paper

Acupuncture in Medicine
1–8
DOI:10.1136/acupmed-2018-011671

Acupuncture in sham device controlled trials may not be as effective as acupuncture in the real world: a preliminary network meta-analysis of studies of acupuncture for hot flashes in menopausal women

Tae-Hun Kim[1], Myeong Soo Lee[2,3], Terje Alraek[4,5] and Stephen Birch[4]

Abstract

Background: Randomised controlled trials of acupuncture performed using sham interventions to control for the placebo effect have mostly used two types of sham techniques: techniques with minimal insertion of acupuncture needles with no additional stimulation (shallow needling control) and techniques with sham acupuncture devices that do not penetrate the skin (sham device control). To achieve successful blinding, sham device controlled acupuncture trials also use the acupuncture base unit in the verum acupuncture group, but in the shallow needling control trials this is not necessary for the verum acupuncture treatment.

Objective: In this study, we analysed the estimated comparative effectiveness of these two verum acupuncture modalities in studies of acupuncture for menopausal hot flashes that used two types of sham control treatments.

Methods: We conducted a network meta-analysis that included randomised controlled trials of acupuncture for hot flashes. Electronic databases, including Medline, Embase, Cochrane Library and AMED, were searched through March 2017. Data were extracted using a predefined data extraction tool by two independent reviewers. The risk of bias was assessed using the Cochrane risk of bias tool for randomised controlled trials. A five-node network meta-analysis was conducted based on the frequentist framework.

Results: Eight studies were included in this review. From the network meta-analysis, we found that verum acupuncture in the shallow needling controlled trials was more effective than verum acupuncture in the sham device controlled trials (SMD −7.27, 95% CI −9.11 to −5.43). Significant heterogeneity and inconsistency were not observed among the included studies or the comparisons.

Conclusions: From this preliminary analysis, we found that different types of verum acupuncture may have different effect sizes with respect to the severity of menopausal hot flashes.

Keywords

acupuncture, sham device, shallow needling, network meta-analysis, verum acupuncture

Accepted 29 June 2018

[1]Korean Medicine Clinical Trial Center, Korean Medicine Hospital, Kyung Hee University, Seoul, Republic of Korea
[2]Clinical Medicine Division, Korea Institute of Oriental Medicine, Daejeon, Republic of Korea
[3]Allied Health Sciences, School of Health and Social Care, London South Bank University, London, United Kingdom
[4]Department of Health Sciences, Kristiania University College, Oslo, Norway
[5]Faculty of Medicine, Department of Community Medicine, National Research Centre in Complementary and Alternative Medicine, UiT The Arctic University of Norway, Tromso, Norway

Corresponding author:
Myeong Soo Lee, Clinical Medicine Division, Korea Institute of Oriental Medicine, Daejeon 34054, Republic of Korea.
Email: drmslee@gmail.com

Introduction

Determining the appropriate control intervention is most important in all types of clinical trial. A placebo is a treatment that is physiologically inert and mimics the verum intervention so that patients remain blinded to their group assignment. Because of these characteristics, the placebo is generally accepted to be the 'gold standard' control intervention in pharmaceutical studies since it has the advantage of reducing the potential risk of bias that is introduced by not blinding patients and researchers. On this basis, different sham interventions that attempt to control for the placebo effect in acupuncture efficacy trials have been conducted for decades.[1]

In acupuncture trials, sham treatment techniques primarily consist of two types: minimal insertion of acupuncture needles (shallow needling control), and non-penetrating needles with a sham acupuncture device (sham device control).[2] Ongoing criticism of acupuncture sham controls has been raised in acupuncture studies[3] because sham acupuncture cannot be physiologically inert, regardless of whether the skin is penetrated or not. All types of sham acupuncture technique stimulate the skin, thereby causing afferent nerve activation that might induce many reactions including the limbic touch response in the brain.[4] This effect is supported by clinical research data, in which sham acupuncture has shown a clear effect beyond that of a placebo.[5] Moreover, the effect of acupuncture is much greater than the effect of other types of placebo intervention.[6] As a result of these innate limitations of sham acupuncture controls, the effect of acupuncture is often underestimated because of the less significant difference between acupuncture and these sham treatments, which has led to the wrong conclusion of 'no effect' of acupuncture being drawn in 'sham' controlled studies.[7] Thus, clinical practice guidelines[8] lack relevant information, about acupuncture, i.e., true effect size.

Disputes about studies of acupuncture with sham controls have mainly focused on problems related to the control interventions—that is, sham acupuncture techniques. However, other questions need to be answered to evaluate the effect of acupuncture in these studies, for example how well the verum acupuncture is delivered in these studies and whether it reflects real-world practice. In general, sham device controlled acupuncture trials use a base unit that is an essential aspect to support the sham device attached to the skin; similarly, the verum acupuncture group also uses this approach to blind participants.[9,10] This base unit inhibits or reduces the ability to manipulate the needle, thereby diminishing the effectiveness of the needling technique.[11] However, in shallow needling controlled trials, verum acupuncture does not use the base unit, and hence normal unconstrained needling techniques that are much closer to clinical practice in the real world are applied. We hypothesised that the effectiveness of verum acupuncture that uses the base unit will be less than that of verum acupuncture that does not use this base unit. In order to test this hypothesis, we assessed whether differences occurred in the estimates of the effect of two verum acupuncture modalities in studies with two different types of sham control (shallow needling control or sham device control). To examine this, we selected a specific condition, menopausal hot flashes, because clinical evidence has been evaluated previously in several systematic reviews (SRs) and many clinical studies have adopted different types of sham control intervention in which we are interested.[12–14]

The objective of this study was to evaluate whether the effect of verum acupuncture in sham acupuncture device controlled trials is different from its effect in shallow needling controlled trials. Although previous SRs have studied the effects of acupuncture versus sham acupuncture on hot flashes and menopausal symptoms in patients with breast cancer,[12–14] studies directly comparing these two interventions have not been performed; therefore, we assessed indirect evidence through a network meta-analysis. To simplify the analysis, we included randomised controlled trials of acupuncture for hot flashes in otherwise healthy, menopausal women.

Methods

This study is a network meta-analysis, guided by recommendations from the Cochrane collaboration.[15] We conducted a network meta-analysis that included randomised controlled trials of acupuncture for menopausal hot flashes on the frequentist framework. The population of this study was perimenopausal or postmenopausal women with hot flash symptoms. Women with hot flashes who were patients with or survivors of breast cancer were excluded. Interventions were verum acupuncture in the sham device controlled trials (A1) and verum acupuncture in the shallow needling controlled trials (A2). Sham acupuncture devices (S1) and shallow needling (S2) were included as comparators in this study. The Park sham needle[9] and Streitberger needle[10] are well known S1 interventions that consist of non-penetrating stimulation of the skin and require the use of a sham base unit, as per the verum acupuncture group, A1. For S2, skin penetration is applied without active manipulation of the acupuncture needles. A waiting list control (W) was selected as a common comparator for the network meta-analysis. Only manual acupuncture was included in this study, and we excluded studies that used electro-acupuncture. Outcomes were severity scores of hot flashes. For the units of analysis in the outcome assessment, we considered one time point (the first results after treatment sessions) only. In general, acupuncture studies usually show clinical heterogeneity in the selection of acupuncture needling locations, total number of treatment sessions and stimulation methods. This was a preliminary study, and therefore we

assumed that transitivity of all the comparisons was satisfied—that is, the relative relationships were consistent, if any study met the inclusion and exclusion criteria of this review.

Electronic database searches of Medical Literature Analysis and Retrieval System Online (Medline) via PubMed, Excerpta Medica Database (Embase), Cochrane Central Register of Controlled Trials (CENTRAL) and Allied and Complementary Medicine Database (AMED) were conducted in March 2017 by one author (MSL). A search strategy was adopted for each database, and two text words, 'acupuncture' and 'hot flashes', were used for each database search. Study population, ethnicity, acupuncture and control group interventions, treatment duration, main outcomes, outcome assessment time points and study result data were extracted using a predefined data extraction form by two authors (TH-K and MSL).

Risk of bias was assessed according to the recommendations from the Cochrane collaboration, and the overall risk of bias was decided based on a rating method for quality of evidence assessment from the GRADE guidelines[16]: no serious limitations (low risk of bias in all domains), serious limitations (high risk of bias in one domain) or very serious limitations (high risk of bias in more than one domain). Individual and overall risk of bias was evaluated by two authors (TH-K and MSL) and discussed if there were different opinions.

Direct effect estimates (treatment contrast) were calculated for each pairwise comparison via a random effects meta-analysis. Because measurement tools for hot flashes were reported to be different across the studies, standard mean differences (SMDs) were selected for effect estimates. In this study, we used change-from-baseline outcome scores that were calculated for studies with final scores only. We conducted a five-node network meta-analysis (A1 vs A2 vs S1 vs S2 vs W) to calculate the indirect and mixed effect estimates based on the frequentist framework, which is based on the contrast-based model.[17 18] First, a multivariate network meta-analysis was conducted with the same heterogeneity variance under the consistency model. Then, we evaluated the inconsistency in the network meta-analysis either at the global level of the whole network or the local level of a specific comparison. For global inconsistency, we conducted a network meta-analysis under the inconsistency model (design-by-treatment interaction model). For the assessment of local inconsistency, the node splitting method was adopted in this study. We used Review Manager 5.3.5 (Nordic Cochrane Centre; http://community.cochrane.org/tools/review-production-tools/revman-5/revman-5-download) to calculate the direct effect estimates and STATA 14.2 (StataCorp LLC, Texas, USA) for the network meta-analysis, which was performed according to routine analytic procedures suggested by Chaimani et al.[19]

Data availability

All data generated or analysed in this study are reported in the article and online supplements.

Results

From the electronic database search, a total of 186 articles were located and initially included by screening titles and abstracts. Thirty-four hard copies were reviewed, and eight studies were included in the final review (Figure 1).[20–27] The reasons for the exclusion of individual studies are presented in Supplemental material 1.

Among the included studies, four studies compared A1 versus S1,[22 23 26 27] and three studies compared A2 versus S2.[20 24 25] One study included three groups comparing A1 versus S1 versus W,[27] and one study included two groups comparing A2 versus W.[21] The detailed acupuncture treatments and outcomes differed across the individual studies according to the assessment duration and tools used. The study characteristics and detailed acupuncture treatment information are shown in Supplemental material 2 and 3.

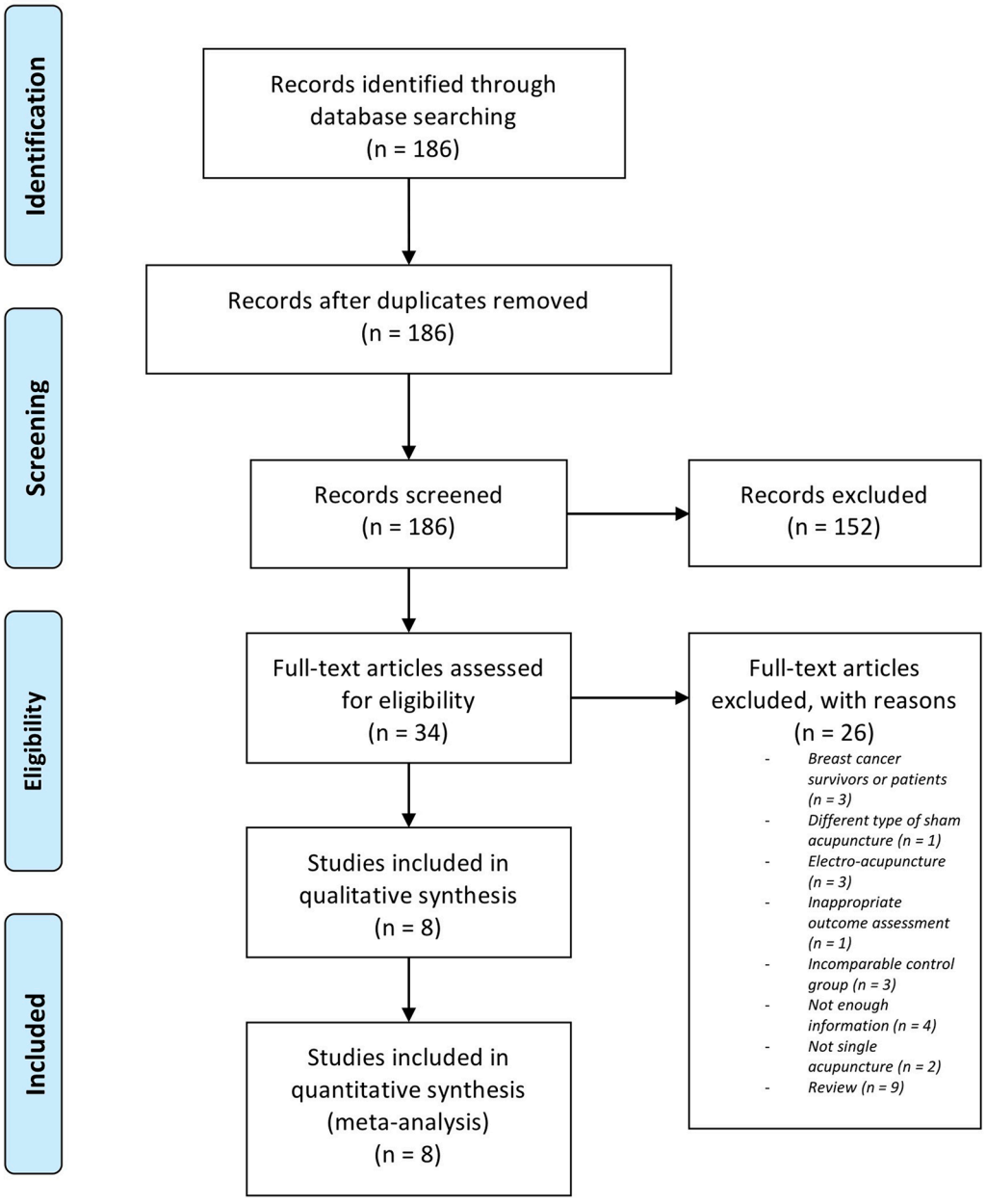
Risk of bias

All studies had a low risk of bias in the sequence generation, blinding of participants and selective outcome reporting domains. Regarding allocation concealment, blinding of outcome assessor and incomplete outcome data domains, many studies did not provide detailed information or implied a high risk of bias (Supplemental material). A serious risk of bias was not observed overall according to the GRADE approach.[16]

Comparative effectiveness of verum acupuncture in sham device controlled trials (A1) and in shallow needling controlled trials (A2)

A summary of the effect estimates of direct evidence and calculated mixed evidence is presented in Table 1. Studies comparing A1 and A2 directly were not found (Figure 2). From the network meta-analysis, we found that the comparative effectiveness of A2 was significantly better than that of A1 (SMD −7.27 (95% CI −9.11 to −5.43)) and this result did not seem to be changed by the heterogeneity in the network because predictive intervals did not include zero (Figure 3). A1 was found to be better than W only (SMD 1.33 (95% CI 0.12 to 2.54)) and significant differences were not observed between S1 and A1 (SMD 0.29 (95% CI −0.30 to 0.87)). A2 showed a better effect than S1 (SMD 7.56 (95% CI 5.71 to 9.40)) and W (SMD 8.60 (95% CI 7.21 to 9.99)). However, significant differences were not observed between S2 and A2 (SMD 1.05, 95% CI (−0.17, 2.26)).

Figure 1. Study flow diagram.



Inconsistency

Evidence of inconsistency was not observed at the global or local level. From the network meta-analysis based on the inconsistency model, testing for inconsistency did not show statistical significance (P=0.429). From the node splitting analysis, no significant inconsistency was found between the direct and indirect summary estimates from every split individual comparison (Table 2).

Discussion

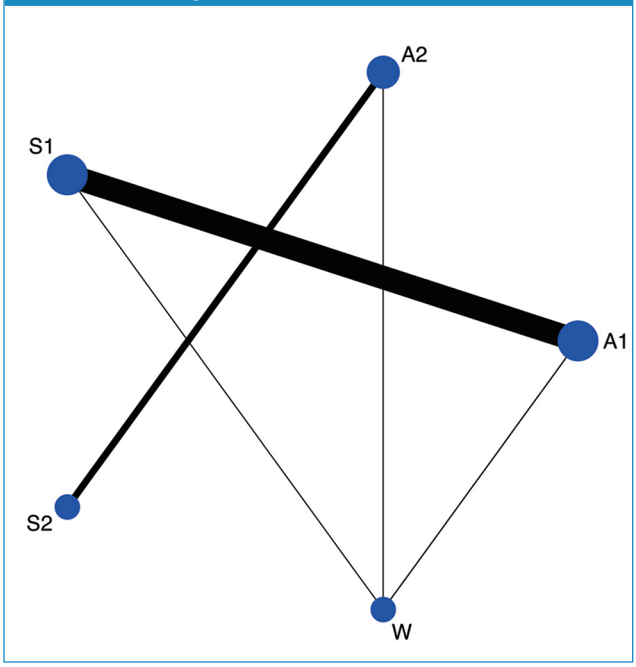
In this preliminary study, we assessed the comparative effect estimates of the two types of verum acupuncture (A1 vs A2) for menopausal hot flashes via a network meta-analysis. Eight studies were finally selected for inclusion out of 186 articles. The network meta-analysis did not present serious bias limitations when assessed using the GRADE methods. We found that verum acupuncture in the shallow needling controlled studies was more effective than verum

Table 1. League table for direct effect estimates (upper triangle) and network meta-analysis effect estimates (lower triangle) of the compared interventions for hot flashes.

A1	–	0.29 (−0.29 to 0.87)	–	1.05 (0.12 to 1.98)
7.27 (5.41 to 9.13)	A2	–	0.87 (0.46 to 1.28)	8.61 (7.61 to 9.60)
−0.29 (−0.88 to 0.31)	−7.56 (−9.42 to −5.70)	S1	–	1.67 (0.64 to 2.70)
6.74 (4.76 to 8.71)	−0.53 (−9.42 to −5.70)	7.02 (5.04 to 9.01)	S2	–
−1.33 (–2.56 to −0.11)	−8.60 (−10.00 to −7.20)	−1.05 (−2.27 to 0.18)	−8.07 (−9.63 to −6.51)	W

Results are presented as the standard mean difference (SMD) and 95% CI. Comparison must be read from left to right (ie, treatment 1 vs treatment 2). A SMD less than zero indicates that treatment 1 is favoured in the network meta-analysis and treatment 2 is favoured in the direct comparison. A1: verum acupuncture in sham device controlled trials; A2: verum acupuncture in shallow needling controlled trials; S1: sham device control; S2: shallow needling control; W: waiting list.

Figure 2. Network graph of the effect of acupuncture on the severity of hot flashes. Each node represents an intervention, and the node size reflects the number of studies assigned to the intervention. The thickness of the lines between interventions is the inverse variance of the direct comparisons. A1: verum acupuncture in sham device controlled trials; A2: verum acupuncture in shallow needling controlled trials; S1: sham device control; S2: shallow needling control; W: waiting list.



acupuncture in the sham device controlled trials. Significant inconsistency was not observed.

This study presents several strengths. To the best of our knowledge, this is the first network meta-analysis comparing the effect of verum acupuncture modalities adopted in different controlled clinical trials. The differences in the effect estimates of acupuncture between different study designs have been explored.[28] A recent article presenting individual patient data meta-analyses from acupuncture studies in chronic pain suggested that the effect size of acupuncture changes when different control interventions are used. For chronic pain, acupuncture was more effective when a non-penetrating sham control was used than when penetrating needle control interventions were used (difference in effect size 0.45 (95% CI 0.12 to 0.78)), and this finding is similar to that of our study.[28] In that study, the authors assumed that the results were caused by the characteristics of control interventions, such as the certainty level of credibility in blinding and physiological inertness between penetrating needle control and sham device control, and they did not assess differences between verum acupuncture modalities within each study design. However, because the study results are derived from an indirect comparison, the difference in effect sizes between the verum acupuncture modalities might have affected the results.

In our study, we conducted a network meta-analysis to compare the effect estimates between verum acupuncture in the sham device controlled studies (A1) and verum acupuncture in the shallow needling controlled studies (A2) through an indirect comparison method using a common comparator: waiting list control. Accordingly, the difference between the two verum acupuncture modalities was clarified. Second, we tried to exclude potential clinical heterogeneity by limiting the population of this review to only perimenopausal and postmenopausal women with hot flashes. Although the previous study is sufficiently rigorous in its meta-analysis study methodology with individual patient data, it still has limitations of clinical heterogeneity, and various conditions are included in chronic pain, including headache, migraine, tension headache, knee arthritis and chronic back pain.[28]

Further potential limitations were seen. First, these study results were not derived from a direct comparison between A1 and A2. A network meta-analysis, in principle, is conducted on the basis of consistency and transitivity assumptions.[29] For transitivity, we did not consider that important effect modifiers, such as treatment duration, traditional acupuncture point selection and styles, outcome assessment duration and severity of hot flashes, were different between

Figure 3. Interval plot with prediction of the severity of hot flashes with acupuncture. Network estimates from all pairwise comparisons are suggested in the graph. The black horizontal lines represent the 95% CIs, and the red horizontal lines represent the 95% predictive intervals (95%Pri). A1: verum acupuncture in sham device controlled trials; A2: verum acupuncture in shallow needling controlled trials; S1: sham device control; S2: shallow needling control; W: waiting list.

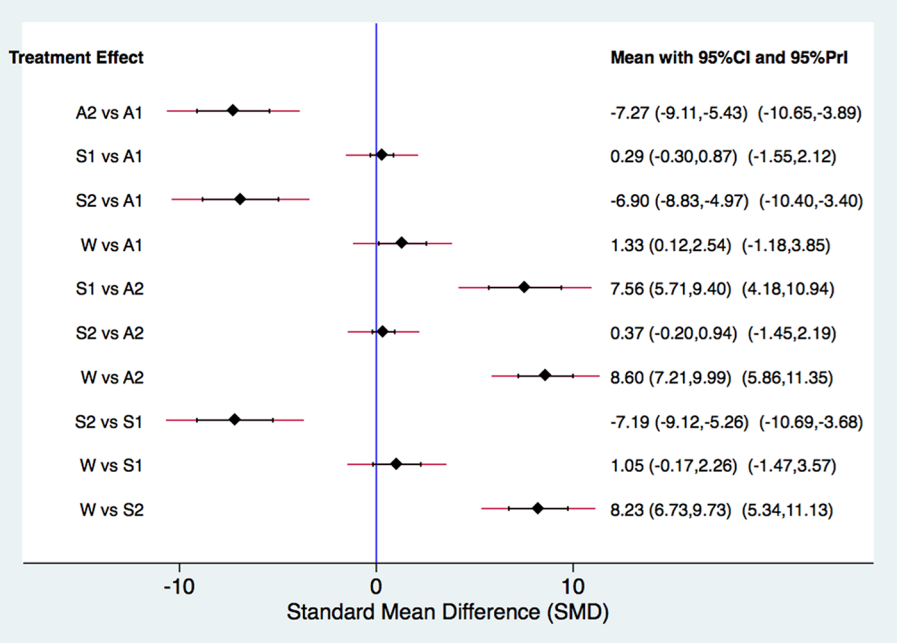


Table 2. Results of testing inconsistency at the local level through the node splitting method.

	Direct comparison		Indirect comparison		Difference		
Side	Coefficient	SE	Coefficient	SE	Coefficient	SE	P values
A1 S1	0.2883341	0.3035104	14.83142	100.119	−14.54309	100.1195	0.89
A1 W	1.101483	0.717156	2.335835	1.415787	−1.234352	1.55115	0.43
A2 S2	0.5328116	0.3499449	15.07478	115.4379	−14.54197	115.4381	0.90
A2 W	8.604417	0.7141588	0.7736375	55.48766	7.830779	55.49262	0.89
S1 W	1.285056	0.722441	0.0609693	1.406863	1.224087	1.550472	0.43

A1: verum acupuncture in the sham device controlled trial; A2: verum acupuncture in the shallow needling controlled trial; SE, standard error; S1: sham device control; S2: shallow needling control; W: waiting list.

the studies, which might have resulted in considerable clinical heterogeneity among the included studies. Although significant inconsistency was not observed at the global and local levels, transitivity of the included comparisons could not be ensured. In addition, the estimated effect size of each intervention seems to be overestimated in comparison with previous results from the Cochrane review, which suggested a smaller effect estimate in a similar population.[30] Several factors might be related to this; only a small number of studies were included in the analysis and some small studies with small sample size had a great influence on the estimation of effects, which might be a limitation of the frequentist methodology in network meta-analysis. In this sense, our study results are only preliminary and should be interpreted carefully. It seems appropriate that the result of this network meta-analysis needs to be used for comparing the relative differences between interventions of indirect comparison and the effect estimates obtained should not be accepted as is.

Second, publication bias might have occurred. Several studies were excluded from our study because we could not find available data for the analysis (Supplemental material1).

Third, the different effect sizes of two forms of verum acupuncture might have originated from the different

placebo effects of two sham control modalities. It has been suggested that control group interventions can affect the effect size in acupuncture trials.[28] To evaluate which factor is more important in determining the effect size of verum acupuncture, fundamental studies assessing physiological changes should be conducted on these two types of verum acupuncture needling in the future.

Finally, core databases were used only for searching and we did not include various other electronic databases such as Clinical Trials.gov, Google Scholar and Turning Research into Practice (TRIP), which might be a potential limitation of the literature search in this study.[30] However, all the studies identified in the previously published SRs were located in this study, and hence our search strategy might have sufficient methodological rigour to be reliable.

From this review, we found that verum acupuncture in shallow needling controlled trials (A2) is more effective at reducing the severity of menopausal hot flashes than verum acupuncture in sham device controlled trials (A1), and these results are consistent with previous research.[28] Interestingly, shallow needling sham acupuncture (S2) was also found to be more effective than sham device acupuncture (S1). Currently, whether skin penetration has more powerful therapeutic effects on hot flashes than pressure of the skin with the sham device has not been established and must be evaluated in the future. In addition, our results failed to answer one question. In the comparison of A1 and S2, a significantly larger effect in favour of S2 was observed, which is inconsistent with the general idea that verum acupuncture will be more effective than sham acupuncture. These issues require further investigation and would be best confirmed through direct comparative clinical studies in the future.

We found that different types of verum acupuncture might have different effect sizes, which is similar to the different types of sham acupuncture.[28] We assumed that verum acupuncture in the sham device controlled acupuncture study might differ from acupuncture in real-world practice. The base unit, which is necessary to retain a sham acupuncture device on the patient's skin, is also used for verum acupuncture. Using these tools, researchers can achieve successful blinding of participants; however, they might reduce or lose appropriate stimulation, which is of great importance in real-world acupuncture. In a previous report, we noted that manual manipulation of needles may be adversely affected by this base unit, which allows acupuncture needles to be manipulated only with limited depth, stimulation and direction.[11] In addition, using the base unit might introduce negative contextual effects (nocebo) that outweigh the specific effects of needling because it is not usually used in normal acupuncture practice. Accordingly, whether differences in physiological and therapeutic effects are caused by acupuncture in real-world practice and verum acupuncture in sham device controlled trials must be determined.

In conclusion, verum acupuncture used in shallow needling controlled trials (A2) and sham device controlled trials (A1) may have different effect sizes. Future clinical studies should perform a direct comparison between different types of verum acupuncture to confirm this result.

Acknowledgements

We would like to express our sincere gratitude to Nancy E Avis who provided us with data for analysis.

Contributors

T-HK, MSL, TA and SB devised the study design and collected the data. T-HK and MSL conducted the systematic review and network meta-analysis. T-HK wrote the first draft of this manuscript and MSL, TA and SB revised the manuscript. All authors approved the final version of the manuscript accepted for publication.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: MSL was supported by grants from Korea Institute of Oriental Medicine (K17292 and K18122). T-HK was supported by a grant from Kyung Hee University in 2015.

Patient consent

Not required.

Provenance and peer review

Not commissioned; externally peer reviewed.

Supplemental material

Supplemental material for this article is available online.

ORCID iDs

Tae-Hun Kim https://orcid.org/0000-0002-8448-3219
Myeong Soo Lee https://orcid.org/0000-0001-6651-7641

References

1. Lundeberg T, Lund I, Sing A, et al. Is placebo acupuncture what it is intended to be? *Evid Based Complement Alternat Med* 2011; 2011: 1–5.
2. Linde K, Niemann K, Schneider A, et al. How large are the nonspecific effects of acupuncture? A meta-analysis of randomized controlled trials. *BMC Med* 2010; 8: 75.
3. Birch S, Alraek T, Kim KH, et al. Placebo-controlled trials in acupuncture: problems and solutions. *Evidence-based research methods for Chinese medicine*. Singapore: Springer, 2016: 55–64.
4. Campbell A. The limbic system and emotion in relation to acupuncture. *Acupunct Med* 1999; 17: 124–130.
5. Zhang CS, Tan HY, Zhang GS, et al. Placebo devices as effective control methods in acupuncture clinical trials: a systematic review. *PLoS One* 2015; 10:e0140825.
6. Hróbjartsson A and Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010:CD003974.

7. Kim TH, Lee MS and Alraek T. Acupuncture for the management of menopausal and perimenopausal symptoms: Current clinical evidence and perspectives for future research. *Maturitas* 2017; 100: 82–83.
8. Birch S, Lee MS, Robinson N, et al. The U.K. NICE 2014 Guidelines for Osteoarthritis of the Knee: lessons learned in a narrative review addressing inadvertent limitations and bias. *J Altern Complement Med* 2017; 23: 242–246.
9. Park J, White A, Stevinson C, et al. Validating a new non-penetrating sham acupuncture device: two randomised controlled trials. *Acupunct Med* 2002; 20: 168–174.
10. Streitberger K and Kleinhenz J. Introducing a placebo needle into acupuncture research. *Lancet* 1998; 352: 364–365.
11. Kim TH, Kang JW and Lee MS. What is lost in the acupuncture trial when using a sham intervention? *Acupunct Med* 2017; 35: 384–386.
12. Carlos Lopes-Júnior L, Cruz LAPda, Leopoldo VC, et al. Effectiveness of traditional Chinese acupuncture versus sham acupuncture: a systematic review. *Rev Lat Am Enfermagem* 2016; 24: 24.
13. Chien TJ, Hsu CH, Liu CY, et al. Effect of acupuncture on hot flush and menopause symptoms in breast cancer- a systematic review and meta-analysis. *PLoS One* 2017; 12: e0180918.
14. Salehi A, Marzban M and Zadeh AR. Acupuncture for treating hot flashes in breast cancer patients: an updated meta-analysis. *Support Care Cancer* 2016; 24: 4895–4899.
15. A network meta-analysis toolkit - performing NMA in STATA. 2018 http://methods.cochrane.org/cmi/network-meta-analysis-toolkit (accessed 30 Apr 2018).
16. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). *J Clin Epidemiol* 2011; 64: 407–415.
17. Salanti G, Higgins JP, Ades AE, et al. Evaluation of networks of randomized trials. *Stat Methods Med Res* 2008; 17: 279–301.
18. White IR. Network meta-analysis. *Stata J* 2015; 15: 951–85.
19. Chaimani A, Higgins JP, Mavridis D, et al. Graphical tools for network meta-analysis in STATA. *PLoS One* 2013; 8:e76654.
20. Avis NE, Legault C, Coeytaux RR, et al. A randomized, controlled pilot study of acupuncture treatment for menopausal hot flashes. *Menopause* 2008; 15: 1070–1078.
21. Avis NE, Coeytaux RR, Isom S, et al. Acupuncture in Menopause (AIM) study: a pragmatic, randomized controlled trial. *Menopause* 2016; 23: 626–637.
22. Ee C, Xue C, Chondros P, et al. Acupuncture for menopausal hot flashes: a randomized trial. *Ann Intern Med* 2016; 164: 146–154.
23. Huang MI, Nir Y, Chen B, et al. A randomized controlled pilot study of acupuncture for postmenopausal hot flashes: effect on nocturnal hot flashes and sleep quality. *Fertil Steril* 2006; 86: 700–710.
24. Kim D-I, Roh J-J, Choi M-S, et al. A clinical trial to assess the efficacy of acupuncture on hot flashes in postmenopausal women. *Korean J Oriental Med* 2007; 28: 74–85.
25. Kim DI, Jeong JC, Kim KH, et al. Acupuncture for hot flushes in perimenopausal and postmenopausal women: a randomised, sham-controlled trial. *Acupunct Med* 2011; 29: 249–256.
26. Nir Y, Huang MI, Schnyer R, et al. Acupuncture for postmenopausal hot flashes. *Maturitas* 2007; 56: 383–395.
27. Painovich JM, Shufelt CL, Azziz R, et al. A pilot randomized, single-blind, placebo-controlled trial of traditional acupuncture for vasomotor symptoms and mechanistic pathways of menopause. *Menopause* 2012; 19: 54–61.
28. MacPherson H, Vertosick E, Lewith G, et al. Influence of control group on effect size in trials of acupuncture for chronic pain: a secondary analysis of an individual patient data meta-analysis. *PLoS One* 2014; 9:e93739.
29. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015; 162: 777–784.
30. Dodin S, Blanchet C, Marc I, et al. Acupuncture for menopausal hot flushes. *Cochrane Database Syst Rev* 2013; 7: CD007410.