PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Acupuncture treatment for knee osteoarthritis with sensitive points: protocol for a multi-center randomized controlled trial
AUTHORS	Tang, Li; Jia, Pengli; Zhao, Ling; Kang, Deying; Luo, Yanan; Liu, Jiali; Li, Ling; Zheng, Hui; Li, Ying; Li, Ning; Guyatt, Gordon; Sun, Xin

VERSION 1 – REVIEW

REVIEWER	JAVIER MATA Son Llàtzer University Hospital, Palma de Mallorca, Spain
REVIEW RETURNED	17-May-2018

GENERAL COMMENTS	This is a three arm parallel multicenter RCT study protocol, which is intent to test if acupuncture at highly sensitized points would achieve statistically better treatment outcomes than acupuncture at low/non-sensitized points and no acupuncture (i.e. waiting list), respectively. The research subject is interesting: acupuncture at more sensitized points would achieve better treatment effects on knee OA. No existing studies specifically examined this kind of therapeutic approach in knee OA. Acupuncture points represent an important issue regarding effects of acupuncture. The overall level of the paper is good. This paper has a potential to be accepted, but important points have to be clarified or fixed before it is published. General recommendations I would recommend the authors to read: • McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, Wang S, Schnitzer T. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage. 2015 May;23(5):747-60. doi: 10.1016/j.joca.2015.03.005. Review. • MacPherson H, Altman DG, Hammerschlag R, et al. (2010) Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): Extending the CONSORT Statement. PLoS Med (6): e1000261. doi:10.1371/journal.pmed.1000261 Researchers should enhance the adoption of the CONSORT and STRICTA statement to improve the reporting quality of RCT acupuncture and to ensure the truth and reliability of the conclusions. Lu LM, He J, Zeng JC, et al. Impact evaluation of CONSORT and STRICTA guidelines on reporting quality for randomized controlled trials of acupuncture conducted in China. Chin J Integr Med. 2017 Jan;23(1):10-17. doi: 10.1007/s11655-016-2451-z. Epub 2016 Mar 21. Ma B, Chen ZM, Xu JK, et al. Do the CONSORT and STRICTA Checklists Improve the Reporting Quality of Acupuncture and Moxibustion Randomized Controlled Trials Published in Chinese

	Journals? A Systematic Review and Analysis of Trends. PLoS One. 2016 Jan 25;11(1):e0147244. doi: 10.1371/journal.pone.0147244. eCollection 2016. Review. Trial Registration it is not up to date. Please update (last update October 3, 2017. Recruitment status: not yet recruitment) the manuscript said: "Patient enrolment started late October 2017". Strengths and limitations of this study In Strengths and limitations of this study the authors should consider and discuss whether it is a limitation neither a placebo nor sham acupuncture will be employed as an active control. Currently, sham or placebo acupuncture is used to assess the efficacy of the specific component of the acupuncture while reducing any possible influence from clinical contexts and other treatment-related processes. Another limitation is the absence of stratification. The objective of stratified randomization is to ensure balance of the treatment groups with respect to the various combinations of the prognostic variables. Simple randomization will not ensure that these groups are balanced within these strata so permuted blocks are used within each stratum are used to achieve balance. Specifying subgroups before the trial is conducted does not mitigate this bias; mitigation would require stratification according to the subgroup variable before randomization, so that patient characteristics would be balanced in the two groups within each subgroup stratum. A trial's credibility is weakened if the groups are not matched for important baseline characteristics Why do you use a waiting list instead of placebo or sham acupuncture? The authors seemed to have included participants with any level of pain. However, it is recommended to include people with pain intensity of at least 4 out 10 on the VAS. Please consider reviewing as this is an important limitation. Please refer to OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis DOI: http://dx.doi.org/10.1016/j.joca.2015.03.005 Introduction Pilot randomized trial of 36 patients. (NCT03008668): in the study design the target number of participants that you need for the study is 664 participants, there are only two arms but it doesn´t say nothing about pilot study. It is not up to date(the date of first posted and last update posted is the same January 2 2017). Could you explain that? METHOD AND ANALYSIS Please consider my comment about including a reference to the study is conforming to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis and STRICTA (Standards for Reporting Interventions in Clinical Trials of Acupuncture) guidelines for acupuncture studies. INCLUSION CRITERIA Consider my comment above about including people with pain intensity of at least 4 out of 10 (VAS for pain) "refractory knee pain" in the manuscript, Selection of acupoints Why do you choose only five points in the treatment? Explain that. The specific point locations used in treatments where standardised should be described in terms of an accepted nomenclature (e.g.

GB21) or in terms of anatomical location where there is no accepted name. (Not in capital letters). (eg. Yanglingquan (GB34) or GB34 (Yanglingquan)).
Interventions
Do you use a manual stimulation during 30 minutes? The manuscript: “The acupuncture stimulation lasts for 30 min”. Explain that.
Table 1.
The table should include: Schedule of Enrolment, Interventions, And Assessments (SPIRIT)
Though various presentation formats exist, key information to convey includes the timing of each visit, starting from initial eligibility screening through to study close-out; time periods during which trial interventions will be administered; and the procedures and assessments performed at each visit.
Figure. Example template of recommended content for the schedule of enrolment, interventions, and assessments. *
STUDY PERIOD
Enrolment Allocation Post-allocation Close-out
TIMEPOINT** -t1 0 t1 t2 t3 t4 etc. tx
ENROLMENT:
Eligibility screen X
Informed consent X
[List other procedures] X
Allocation X
INTERVENTIONS:
[Intervention A]

[Intervention B] X X
[List other study groups]

ASSESSMENTS:
[List baseline variables] X X
[List outcome variables] X X etc. X
[List other data variables] X X X X etc. X

*Recommended content can be displayed using various schematic formats. See SPIRIT 2013 Explanation and Elaboration for examples from protocols.
**List specific time points in this row.
Primary outcome
The current WOMAC survey is comprised of 24 items divided into three subscales: Pain (5 items) with a score range of 0–20, stiffness (2 items) score range of 0–8, and physical function (17 items) score range of 0–68. The physical functioning questions cover everyday activities. These scales will be used separately and will not be summed. Patients will respond orally to the five levels with the following criteria: “none” = 0, “a bit” = 1, “quite a bit” = 2, “a lot” = 3, and “very much” = 4. If two or more questions are left unanswered, the scale will be declared invalid. If the patient does not respond to one question, a mean will be taken from the results of the other questions. The range will be 0–98.Could you explain the Chinese version of WOMAC ?, in the manuscript you said that « Each of the 24 items will be graded on a visual analog scale ranging from 0 to 10, with higher scores reflecting more pain, stiffness and poorer physical function » What is the subscale and the toatl range ? It is important becouse you use this in the sample size.
Sample size
A change of more than 12 units by 14 weeks is an almost 40% improvement from baseline (WOMAC function not in the WOMAC

	total score) page 906. Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. Ann Intern Med. 2004 Dec 21;141(12):901-10. Discussion Page 16 line 45. Could you explain better this idea I don´t understand: Were this hypothesis proved, the findings would have profound impact on the theory and practice of acupuncture on KOA. Page 16 line 46. Replace Frist by First You don´t say nothing about data sharing, (to facilitate reproducibility and data reuse). Figure 1. Flow diagram The flow diagram should present a complete timeline of the study visits, enrolment process, interventions, and assessments performed on participants. E.g. References List the names and initials of all authors if there are 3 or fewer; otherwise list the first 3 and add 'et al.' Journals from BMJ use a slightly modified version of Vancouver referencing style. (http://www.citethisforme.com). References must be complete, including initial(s) of author(s) cited, title of paper, journal, year of publication, and volume and page numbers. Revise all the references. 1. Jeanette E, Victoria H, Stephen B, et al. Acupuncture for osteoarthritis of the knee: A systematic review. Arthritis & Rheumatism2013; 4: 819-25. I cannot find this reference. Similar is: Ezzo J, Hadhazy V, Birch S,et al. Acupuncture for osteoarthritis of the knee: a systematic review. Arthritis Rheum. 2001;44(4):819-25. 2. Lack of year of publication, and volume and page numbers. 3. BMC Complement Altern Med. 2010;10:32. doi: 10.1186/1472-6882-10-32. 4. Arthritis Rheum. 2009;61(12):1704-11. doi: 10.1002/art.24925. 5. Osteoarthritis Cartilage. 2011;19(11):1314-22. doi: 10.1016/j.joca.2011.08.004. Epub 2011 Aug 16. 6. …

REVIEWER	Barbara Shay University of Manitoba, Canada
REVIEW RETURNED	30-May-2018

GENERAL COMMENTS	In the introduction, there are a few instances where the sentence tense does not agree, it should be in the present and sometimes it is in the past tense. Also the report of the incidence of KOA is somewhat confusing between actual incidence and those comparing to "person-years". can you convert the person years to incidence mathematically, then it would be easier for the reader to compare? the three arms include two treatments and one control. can you justify why there is no placebo arm included. I assume this is becasue there is enough literature to substantiate that placebo has an effect but actual needling has the placebo plus a physiological effect and therefore does not need to have a group, but I think it needs to be said. on page 10 lines 7-16. are the acupoints listed the same as the tenderpoints or ashi points. Or will the tenderpoints be needled even if they are not the acupoints? could you clarify how often or even if the needles are manually stimulated during the 30 minute treatment?

	It is curious that even once the study is complete, you will offer the control group, the non-study standard acupuncture treatment. If the high sensitive points are superior, why wouln't you establish the treatment points in the same way for the control group's crossover treatment. could you point out the STRICTA criteria and how the study follows this criteria? Why are you allowing the control group to undergo non-acupucnture treatments? What if these are quite successful and there is less difference between the two acupucnture treatment arms and the control group. I think you should either allow the acupucnture groups the same non-acupucnture treatments or discontinue any treatments even for the control group. After all you are allowing NSAIDS to all participants so why not the non-pharmacological treatment of their pain. If you think there is a difference then just justify it on evidence-informed grounds. Having said all of this, the study is well thought out and seems feasible. I will be interested in the results. Thank you.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

1. I would recommend the authors to read:

- **McAlindon TE, Driban JB, Henrotin Y, Hunter DJ, Jiang GL, Skou ST, Wang S, Schnitzer T. OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis. Osteoarthritis Cartilage. 2015 May;23(5):747-60. doi: 10.1016/j.joca.2015.03.005. Review.**
- **MacPherson H, Altman DG, Hammerschlag R, et al. (2010) Revised Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): Extending the CONSORT Statement. PLoS Med (6): e1000261.doi:10.1371/journal.pmed.1000261.**

[Response] Thank you so much for your suggestion. We have read the recommended articles, cited these two references.

2. Researchers should enhance the adoption of the CONSORT and STRICTA statement to improve the reporting quality of RCT acupuncture and to ensure the truth and reliability of the conclusions.

[Response] Thank you for your suggestion. We will follow the CONSORT and STRICTA are followed when reporting the protocol and results of the trial. This paper reports a protocol of the trial.

3. Trial Registration: It is not up to date. Please update (last update October 3, 2017. Recruitment status: not yet recruitment) the manuscript said: "Patient enrolment started late October 2017"

[Response] Thank you for your suggestion. We have uploaded our progress of the trial in the ClinicalTrials.gov. Once the staff of the ClinicalTrials.gov checks the revision, the updated trial status will appear on the website.

4. Strengths and limitations of this study: In Strengths and limitations of this study the authors should consider and discuss whether it is a limitation neither a placebo nor sham acupuncture will be employed as an active control. Currently, sham or placebo acupuncture is used to assess the efficacy of the specific component of the acupuncture while reducing any possible influence from clinical contexts and other treatment-related processes.

[Response] Thank you for your suggestion. We agree that this is a limitation and have added to the discussion, as below:

"Our study also has a few limitations should be considered. First, we did not include the shame procedure as an active control. Thus, the placebo effect may not be well parceled out. However, our primary objective is to identify whether there is difference between in the acupuncture between highly sensitive and no/low sensitive points. Thus, the no/low sensitive points would thus serve as the primary active control. The inclusion of waiting list will also be used a non-treatment control."

5. Another limitation is the absence of stratification. The objective of stratified randomization is to ensure balance of the treatment groups with respect to the various combinations of the prognostic variables. Simple randomization will not ensure that these groups are balanced within these strata so permuted blocks are used within each stratum are used to achieve balance.

Specifying subgroups before the trial is conducted does not mitigate this bias; mitigation would require stratification according to the subgroup variable before randomization, so that patient characteristics would be balanced in the two groups within each subgroup stratum.

A trial's credibility is weakened if the groups are not matched for important baseline characteristics

[Response] Thank you for your suggestion. In our study, we applied stratification by the four participating sites. We also had thought very carefully about additional variables for stratification during the planning of the trial. However, after three rounds of panel discussions, we decided not to include additional baseline variables. We have added this to the limitation of the discussion section.

6. Why do you use a waiting list instead of placebo or sham acupuncture?

[Response] In our study, the primary hypothesis is to differentiate the effect between the acupuncture on high-sensitive versus low/non-sensitive points. Thus, the low/non-sensitive points will be used as an active control. We additionally included waiting list as a non-treatment control. This approach would help differentiate the effects between experimental group versus active control and no-treatment control. Nevertheless, we included a statement in the discussion that the non-use of sham procedure may be a limitation of our study.

7. The authors seemed to have included participants with any level of pain. However, it is recommended to include people with pain intensity of at least 4 out 10 on the VAS. Please consider reviewing as this is an important limitation. Please refer to OARSI Clinical Trials

Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis DOI: http://dx.doi.org/10.1016/j.joca.2015.03.005

[Response] Thank you for your suggestion. However, we are unable to amend this, because we have recruited over 100 patients for the study. We recognized that the inclusion of patients with VAS less than 4 may be a limitation, and have added this point to the limitation of discussion section, as below.

"Second, we included participants with any level of pain. The inclusion of patients with VAS less than 4 may not be responsive enough to allow for detection of change. However, the VAS is a secondary outcome, and would not affect our primary analysis"

8. Introduction

Pilot randomized trial of 36 patients. (NCT03008668): in the study design the target number of participants that you need for the study is 664 participants, there are only two arms but it doesn´t say nothing about pilot study. It is not up to date (the date of first posted and last update posted is the same January 2 2017). Could you explain that?

[Response] We apologize for the confusion. In our plan, we will conduct two trials – one is the pilot trial, and the other is the definitive trial. We have actually completed the registration for both trials at the ClinicalTrials.gov (the pilot trial: NCT03008668 and the definitive trial NCT03299439). During the registration process, however, our research staff made a mistake, in which she registered the pilot trial with a wrong sample . We apologize for the error, and have corrected this on the clinicaltrial.gov. Meanwhile, we have updated the information in the ClinicalTrials.gov.

We have updated the title of the pilot trial as: "A pilot trial of acupuncture for knee osteoarthritis with differential functional status of acupoints" to reminder the reader that it was a pilot trial. The staff of the ClinicalTrials.gov will inform us the update when they check the revision.

9. Method and analysis

Please consider my comment about including a reference to the study is conforming to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), OARSI Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis and STRICTA (Standards for Reporting Interventions in Clinical Trials of Acupuncture) guidelines for acupuncture studies.

[Response] Thank you for your suggestion. We have added these references and included these guidelines in our study. It reads as below.

"We followed the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials)[27], OARSI (Clinical Trials Recommendations: Design, conduct, and reporting of clinical trials for knee osteoarthritis)[28] and STRICTA (Standards for Reporting Interventions in Clinical Trials of Acupuncture) [29]guidelines."

10. Inclusion criteria

Consider my comment above about including people with pain intensity of at least 4 out of 10 (VAS for pain) "refractory knee pain" in the manuscript

[Response] Thank you very much for your helpful suggestion. Please see our response to question 7.

11. Selection of acupoints

Why do you choose only five points in the treatment? Explain that.

[Response] The reason for choosing only five points in the treatment is based on literature and expert consensus. Firstly, we identified 13 candidate acupoints for treating KOA. Then all of the 13 candidate acupoints or identified tender points will be ranked based on their pressure-pain threshold. The five points with the lowest pain threshold are identified as the highly sensitive points, whereas the five points with the highest pain threshold are selected as the lowly/non-sensitive points, which is designed to primarily examine if acupuncture at highly sensitive points, compared with low/non-sensitive points or no acupuncture (waiting-list), can result in improving pain, joint function and quality of life, among patients with KOA. Please see the identification of candidate acupoints and tender points for sensitization measurement section on page 10-11.

12. The specific point locations used in treatments where standardised should be described in terms of an accepted nomenclature (e.g. GB21) or in terms of anatomical location where there is no accepted name. (Not in capital letters). (eg. Yanglingquan (GB34) or GB34 (Yanglingquan)).

[Response] Thank you for your suggestion. We have revised the nomenclature as below.

"We identified 13 candidate acupoints for treating KOA, namely Heding (EX-LE2), Neixiyan (EX-LE4), Dubi (ST35), Xuehai (SP10), Liangqiu (ST34), Yinlingquan (SP9), Yanglingquan GB34), Zusanli (ST36), Weizhong (BL40), Yingu (KI10), Xiguan (LR7), Ququan (LR8), Weiyang (BL39)."

13. Interventions

Do you use a manual stimulation during 30 minutes? The manuscript: "The acupuncture stimulation lasts for 30 min". Explain that.

[Response] Thank you for your suggestion. By stating "The acupuncture stimulation lasts for 30 min", we meant that "the needles placed in the acupoints were manually stimulated every 15 minutes and were removed after 30 minutes. We have revised the statement in the manuscript.

14. Table 1.

The table should include: Schedule of Enrolment, Interventions, And Assessments (SPIRIT). Though various presentation formats exist, key information to convey includes the timing of each visit, starting from initial eligibility screening through to study close-out; time periods

during which trial interventions will be administered; and the procedures and assessments performed at each visit.

[Response] Thank you for your suggestion. We have revised the flow diagram as below in this revision:

Table 1 Measurements to be taken at each point in trial

TIMEPOINT	STUDY PERIOD					
	Enrolment	Allocation	Post-allocation			Post-allocation
	-1 week (-7~0 day)	Day 0	4 weeks (±3 days)	8 weeks (±3 days)	12 weeks (±3 days)	16 weeks (±3 days)
ENROLMENT						
Eligibility screen	X					
Informed consent	X					
Allocation		X				
INTERVENTIONS						
High sensitization group			X			
Low/non-sensitization group			X			
Waiting-list group			X			
ASSESSMENTS						
X-ray examination of the knee joint	X					
Measurement of sensitization intensity	X					
			X	X	X	X
Measurement of pressure-pain threshold of the five selected points	X		X	X	X	X
	X		X	X	X	X
WOMAC[a] score	X		X	X	X	X
SF-12[b] score			X	X	X	X
Knee ranges of motion	X		X	X	X	X
Adverse events						
Other treatments received for						

knee osteoarthritis						

X=Measurements to be taken at this point

[a] Western Ontario and McMaster Universities Osteoarthritis index

[b] Short Form-12 health survey

15. Primary outcome

The current WOMAC survey is comprised of 24 items divided into three subscales: Pain (5 items) with a score range of 0–20, stiffness (2 items) score range of 0–8, and physical function (17 items) score range of 0–68. The physical functioning questions cover everyday activities. These scales will be used separately and will not be summed. Patients will respond orally to the five levels with the following criteria: "none" = 0, "a bit" = 1, "quite a bit" = 2, "a lot" = 3, and "very much" = 4. If two or more questions are left unanswered, the scale will be declared invalid. If the patient does not respond to one question, a mean will be taken from the results of the other questions. The range will be 0–98. Could you explain the Chinese version of WOMAC ?, in the manuscript you said that « Each of the 24 items will be graded on a visual analog scale ranging from 0 to 10, with higher scores reflecting more pain, stiffness and poorer physical function » What is the subscale and the total range ? It is important because you use this in the sample size.

[Response] Thank you for your suggestion. We used the Chinese version of WOMAC, which is validated. In our Chinese version of WOMAC, the score ranges for the pain, stiffness and physical function subscale are, respectively, 0-50, 0-20 and 0-170, resulting in a total range of 0-240.

In the manuscript, we have revised as below:

The Chinese version of WOMAC consists of 24 items assessing the KOA patients' pain (5 items), stiffness (2 items), and physical function (17 items). Each of the 24 items will be graded on a visual analog scale ranging from 0 to 10, with higher scores reflecting more pain, stiffness and poorer physical function. The score ranges for the pain, stiffness and physical function subscale are, respectively, 0-50, 0-20 and 0-170, resulting in a total range of 0-240.

16. Sample size

A change of more than 12 units by 14 weeks is an almost 40% improvement from baseline (WOMAC function not in the WOMAC total score) page 906.

Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial. Ann Intern Med. 2004 Dec 21;141(12):901-10.

[Response] We apologize for the confusion. In our practice of sample size calculation, we did the initial calculation with very limited prior information. By 12 units, we meant the standard deviation of the score at baseline, not the change itself. We assumed that one-third of 12 units (that is, baseline

deviation) would represent a clinically important difference (i.e. treatment effects). The sample size calculation at the time was based on a scale of total 98 points. In our study, we used the scale of 240 points.

.

Last year, we conducted a pilot trial, and the results are available now. We thus updated the sample size calculation, and included in this version. It reads as below:

Our primary study hypothesis is that acupuncture on high-sensitive points (i.e. experimental group) would achieve more reduction in the total WOMAC score than acupuncture on non/low-sensitive points (i.e. active control group) or waiting list group (i.e. no treatment group).

The sample size estimation was based on the mean difference in the change of total WOMAC score from baseline given the estimates obtained from our pilot trial. The following assumptions were made to calculate the sample size: a mean difference of 12 between the high and low/non-sensitization groups, standard deviation of total score of 33, a two-sided significance level of 0.025 (adjusted for multiple testing), and a power of 0.9. With these assumptions, a sample size 189 patients per arm is required to provide a power of 90% at the alpha level of 0.025 to detect a difference of 12 points between the high and low/non-sensitization groups. This sample size would provide adequate power to detect the difference between high-sensitization group versus waiting list group, on the ground that the treatment effect between high-sensitization group and non/low-sensitization group would be smaller that between high-sensitization group and waiting list group.

To allow for a loss to follow up of 10%, a minimum sample size of 666 patients (222 patients per arm) at baseline was required.

17. Discussion

Page 16 line 45. Could you explain better this idea I don´t understand: Were this hypothesis proved, the findings would have profound impact on the theory and practice of acupuncture on KOA.

[Response] Thank you for your comments. As stated in our introduction section, according to the theory of traditional Chinese medicine (TCM), there are connections between the disease conditions and their respective points (i.e. traditional acupoints and tender points (ashi points) on the surface of the human body). These points become sensitized when the body suffers from a disease state. Stimulation of the sensitive points could lead to an improvement of disease conditions. Therefore, by identifying sensitive points associated with KOA, one would hypothesize that acupuncture at more sensitized points would achieve better treatment effects on KOA. Were this hypothesis proved, the findings may change the regular practice of acupuncture on KOA.

18. Page 16 line 46. Replace Frist by First You don´t say nothing about data sharing, (to facilitate reproducibility and data reuse).

[Response] Thank you for your suggestion. We have revised this issue as below:

"First, the trial would contribute to a better understanding of the effectiveness of acupuncture for KOA…"

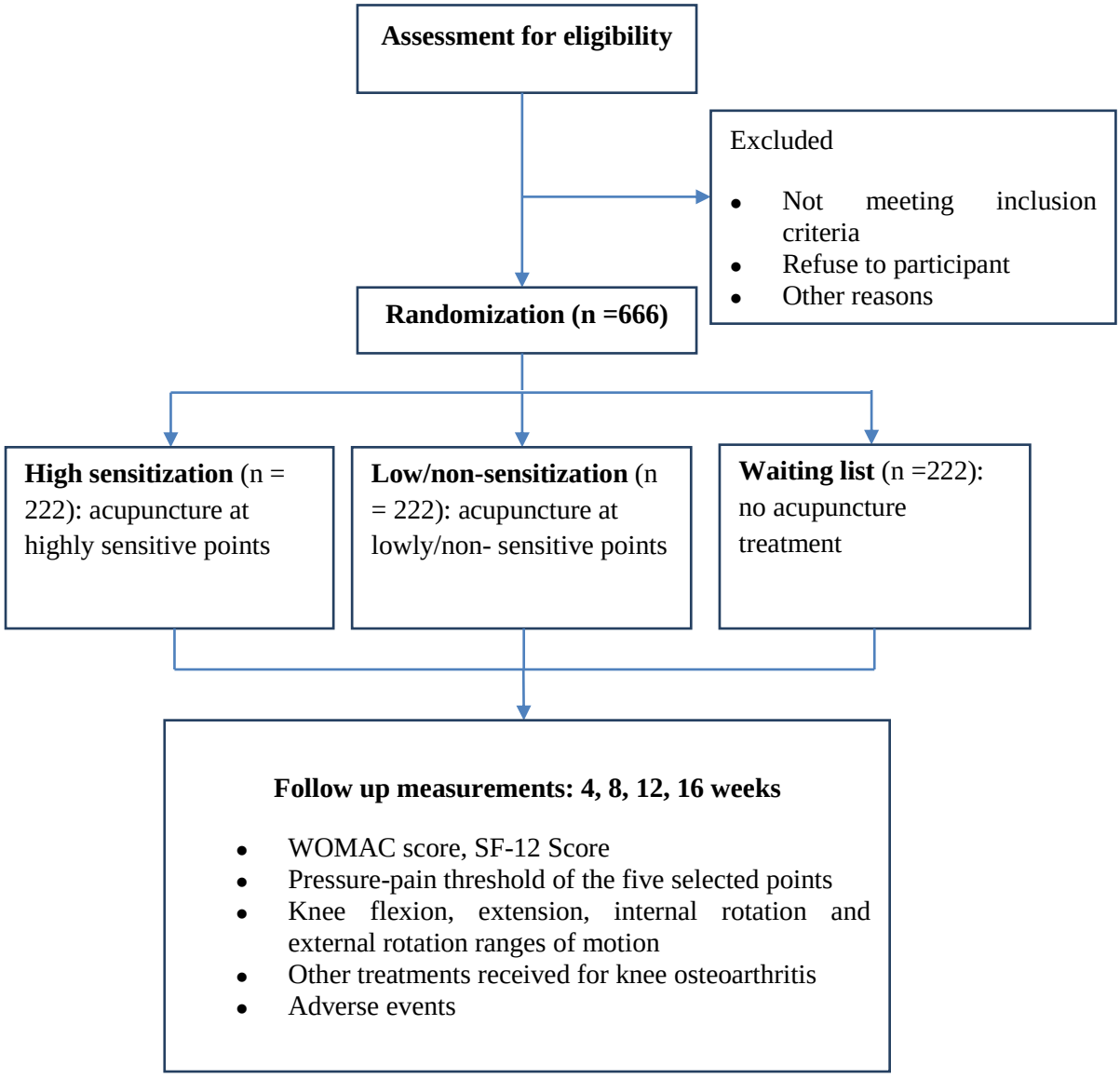
We have added a data sharing statement in the manuscript. It now read as below：

*"**Data sharing statement** The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request."*

19. Figure 1. Flow diagram

The flow diagram should present a complete timeline of the study visits, enrolment process, interventions, and assessments performed on participants.

[Response] Thank you for your suggestion. We have revised the flow diagram as below in this revision:



20. References

List the names and initials of all authors if there are 3 or fewer; otherwise list the first 3 and add 'et al.' Journals from BMJ use a slightly modified version of Vancouver referencing style. (http://www.citethisforme.com).

References must be complete, including initial(s) of author(s) cited, title of paper, journal, year of publication, and volume and page numbers. Revise all the references.

[Response] Thank you for your suggestion. We have revised the references in this revision.

Reviewer: 2

1. In the introduction, there are a few instances where the sentence tense does not agree, it should be in the present and sometimes it is in the past tense. Also the report of the incidence of KOA is somewhat confusing between actual incidence and those comparing to "person-years". Can you convert the person years to incidence mathematically, then it would be easier for the reader to compare?

[Response] Thank you for your suggestion. We have checked instances in the introduction and modified the false sentence tense, as below.

*"The age-standardized incidence rates for knee-replacement surgery **were** estimated at 150 per 100,000 person-years in western countries."*

*"In this study, we **aim** to design a definitive trial to primarily examine if acupuncture at highly sensitive points, compared with low/non-sensitive points or no acupuncture (waiting-list), can result in improving pain, joint function and quality of life, among patients with KOA."*

Based on the raw data in the original reference, we have converted the 240 per 100,000 person-years to incidence rate using the formula: CI = 1 - EXP(-IR*T), and revised the sentence as below.

" the US, where the incidence rate was estimated nearly 0.02%, KOA often results in early retirement and joint replacement."

We also tried to use this formula: CI = 1 - EXP(-IR*T) to convert the 150 per 100,000 person-years to incidence rate, however, the lack of more detailed raw data, we could not convert it. We are sorry for this problem.

2. The three arms include two treatments and one control. Can you justify why there is no placebo arm included. I assume this is because there is enough literature to substantiate that placebo has an effect but actual needling has the placebo plus a physiological effect and therefore does not need to have a group, but I think it needs to be said.

[Response] Thank you for your suggestion. We agree with your statement, and have added the information to the discussion section. Please see response to comment 4 made by reviewer 1.

3. On page 10 lines 7-16. Are the acupoints listed the same as the tenderpoints or ashi points. Or will the tender points be needled even if they are not the acupoints?

[Response] Thank you. In this study, we aim to examine if acupuncture at highly sensitive points, compared with low/non-sensitive points or no acupuncture (waiting-list), can result in improving pain, joint function and quality of life, among patients with KOA. All of the 13 candidate acupoints or identified tender points will be ranked based on their pressure-pain threshold. The five points with the lowest pain threshold are identified as the highly sensitive points, whereas the five points with the highest pain threshold are selected as the lowly/non-sensitive points. Therefore, the tender points will be needled if they are identified as the highly or lowly/non-sensitive points. Detailed information can be found in the identification of candidate acupoints and tender points for sensitization measurement section of manuscript (page 10-11).

4. Could you clarify how often or even if the needles are manually stimulated during the 30 minute treatment?

[Response] Thank you. We have added information that the needles are manually stimulated during the 30 minute treatment, as below.

"The acupuncture stimulation lasts for 30 minutes, and the needles placed in the acupoints were manually stimulated every 15 minutes and were removed after 30 minutes."

5. It is curious that even once the study is complete, you will offer the control group, the non-study standard acupuncture treatment. If the high sensitive points are superior, why wouldn't you establish the treatment points in the same way for the control group's crossover treatment.

[Response] Thank you for your suggestion. It is currently our hypothesis that acupuncture at the high sensitive points will achieve better treatment effects on KOA, and has not been confirmed. However, the results would not be available until we have completed follow up of all patients, at which time most of patients in the waiting list group would have received standard acupuncture treatment.

6. Could you point out the STRICTA criteria and how the study follows this criteria?

[Response] We have listed the STRICTA criteria, and described the extent to which the study would meet the criteria, as below.

Item	Detail	Page
1. Acupuncture rationale	1a) Style of acupuncture	6
	1b) Reason for the treatment provided	6
	1c) Extent to which treatment was varied	6
2. Details of needling	2a) Number of needle insertions per subject	11
	2b) Names of the points	11

	2c) Depth of insertion	11
	2d) Responses sought	11
	2e) Needle stimulation	11
	2f) Needle retention time	11
	2g) Needle type	11
3. Treatment regimen	3a) Number of treatment sessions	11
	3b) Frequency and duration of treatment sessions	11
4. Other components of treatment	4a) Details of other interventions	12
	4b) Setting and context of treatment	7
5. Practitioner background	5) Description of participating acupuncturists	11
6. Control or comparator interventions	6a) Rationale for the control or comparator	6
	6b) Precise description of the control or comparator	11

7. Why are you allowing the control group to undergo non-acupucnture treatments? What if these are quite successful and there is less difference between the two acupucnture treatment arms and the control group. I think you should either allow the acupucnture groups the same non-acupucnture treatments or discontinue any treatments even for the control group. After all you are allowing NSAIDS to all participants so why not the non-pharmacological treatment of their pain. If you think there is a difference then just justify it on evidence-informed grounds.

[Response] Thank you. The use of non-acupuncture treatments (e.g. NSAIDs) for all patients was mainly due to ethical considerations. However, in order to reduce variations in the use of such treatments, we request that patients use NSAIDs only. We were concerned that the variations of co-interventions would diminish the hypothesized effects.