PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Is acupuncture effective for knee osteoarthritis? A protocol for a systematic review and meta-analysis
AUTHORS	Liu, Chuanyang; Tu, Jian Feng; Lee, Myeong Soo; Qi, Lingyu; Yu, Fang-Ting; YAN, Shiyan; Li, Jin-Ling; Lin, Lu; Hao, Xiao-Wan; Su, Xin-Tong; Yang, Jing-Wen; Wang, Li-Qiong

VERSION 1 – REVIEW

REVIEWER	Martin Bjørn Stausholm University of Bergen
REVIEW RETURNED	02-Jul-2021

GENERAL COMMENTS	I recommend a minor revision of the methods. I am not qualified to judge the novelty of the study. Introduction/discussion section Please consider citing the following paper: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7023333/ In the strengths and limitations section The authors state that “This study will be the first of its kind to explore the difference in the efficacy between manual acupuncture and electroacupuncture for KOA by synthesizing the evidence from direct comparison and indirect comparison.”. However, it is unclear when and if the results will be published. Types of participants The inclusion criteria for participants are too vague. Criteria for excluding studies The authors stated that they will not include studies with lack of allocation concealment. Why is allocation concealment more important than other risk of bias domains, such as a high drop-out rate? I recommend not excluding studies based on lack of allocation concealment. Please remember to add a table of excluded studies with reasons for exclusion and a reference. Acupuncture adequacy assessment I do not understand the following statement by the authors: “To test the success of the blinding, we will ask the two assessors to guess the provenance of each study.”. Heterogeneity assessment A high level of statistical heterogeneity is only an indication that the studies are methodologically heterogeneous. Therefore, the authors should not dismiss a meta-analysis with high statistical heterogeneity in advance. https://www.meta-analysis-workshops.com/download/common-

	mistakes2.pdf Data synthesis They authors stated that “The fixed-effects model will be used if little heterogeneity.”. This is a common mistake. Low statistical heterogeneity is only an indication that the studies are methodological homogeneous – the studies must be completely methodological homogeneous to be included in a fixed effects meta-analysis. Of note, a random effects meta-analysis will produce the same results as a fixed effects meta-analysis in the absence of statistical heterogeneity. Therefore, I recommend sticking to the random effects meta-analysis. https://www.meta-analysis-workshops.com/download/common-mistakes2.pdf

REVIEWER	Neilson Mathews Lehigh Valley Health Network, Family Medicine
REVIEW RETURNED	17-Jul-2021

GENERAL COMMENTS	I first wanted to give my background to understand my comments on the proposed manuscript. I am a Family Physician who specializes in Sports Medicine. I am in academics now and am also a licensed acupuncturist. I am very interested whenever acupuncture is studied but especially in musculoskeletal disorders. I am not well versed in statistical analysis though. My first impression is you are undertaking a massive endeavor. I will list some major comments and then list some minor recommendations from each page. In general have found that acupuncture articles from Asia seem to always find a positive response to acupuncture in published articles compared to the US and Europe. Therefore, I am concerned about the inclusion of a great deal of potentially biased articles in the meta-analysis that may skew the results in a positive direction incorrectly. I think you are looking for some conclusions in your outcomes that will be very difficult to find. I cannot see how you can "explore the applied law of different doses of acupuncture" from medical trials. There is just too much variability to compare different types of acupuncture and point selection not to mention the use of adjunct techniques like electroacupuncture or stimulation. I also have an issue with using sham as a control as it has been shown not to be a placebo control. There are published articles about the effects from sham acupuncture. I worry that using this as a control will possibly mask a positive effect compared to true placebo or wait list controls. Smaller comments: Page 7, line 17-I find this hard to believe that 85% of OA is in the knee just comparing to my clinical practice and all other stats I have seen Page 7, line 38-43-would amplify this section and use examples-which analgesics have no effect? Page 8, line 38-56-although these mediators might be involved, OA is in general not an inflammatory disorder but a pain disorder; I would like to see more included with pain pathways. Page 9, line 48-change "have" to "has" Page 10, line14-Objectives-for question 2, I would find this almost impossible to study as I have rarely seen the amount of medication tracked or recorded in studies Page 11-Types of participants-I would expand on this section and be more specific on what is considered OA and how diagnosed Page 11-Types of interventions-I often use moxabustion as an

	adjunct, was this considered as another type since it is used often? In a recent review we did on this same topic, we found moxa used often Page 12, line 14-again, how would you find this data to make any decent conclusions? Line 19-how would you find the cost effectiveness? Line 25-I would change "symptoms" to "criteria" Page 13, line 12-delete patellofemoral syndrome-it is not OA Page 14, line 43-why study choice of points? it will be next to impossible to make any strong conclusions and would be more appropriate in an acupuncture journal or TCM journal, not a prestigious medical journal. Also, with so many different types of acupuncture (TCM, Japanese, Scalp, Ear, French, etc.) and different ways to diagnose, how can you make conclusions? Page 15, line 30-which methods? Page 18, line 7-grammatically incorrect sentence Page 18, line 38-this is the first time you mentioned specific grading systems of OA, should include earlier if going to use it Page 19-once again, I would be careful using sham as a control Line 35-you likely need to explain the "DeQi" response; I find it hard to believe that a trial was done with enough patients to discern DeQi was associated with a treatment effect. References-question use of #2-its 8 years old and an expert's concensus paper #12, 14, 15, 22, 23, 25 are older papers-think these topics have much better and more recent research/papers #48-ACR has more updated guidelines than 1995

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Comments to the Author:

I recommend a minor revision of the methods. I am not qualified to judge the novelty of the study.

Introduction/discussion section

1. Please consider citing the following paper: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7023333/

Response: We thank the reviewer for the comments. The paper is valuable for us and mentioned the question about “the challenge of adequacy of dose”. We have cited it. We also plan to assess the adequacy of acupuncture and the dosage of acupuncture in our review. Our group built a scoring instrument to calculate the dose of acupuncture from four parameters.[1] Based on the sum of the scores, we defined three doses of acupuncture treatment: high dosage, medium dosage and low dosage. And we designed three subgroups according the three kinds of dosage to explore the relationship between doses of acupuncture and efficacy. (line 145-146, 365-366 and references 33 highlight in red)

[1]. Sun N, Tu J, Lin L, et al. Correlation between acupuncture dose and effectiveness in the treatment of knee osteoarthritis: a systematic review. *Acupuncture in medicine* 2019;37:261-67.

2. In the strengths and limitations section

The authors state that “This study will be the first of its kind to explore the difference in the efficacy between manual acupuncture and electroacupuncture for KOA by synthesizing the evidence from direct comparison and indirect comparison.”.

However, it is unclear when and if the results will be published.

Response: We apologize for our inaccurate statement so we have deleted the “first”. As you said, it is unclear when and if this section will be published. If manual acupuncture and eletroacupuncture are effective for KOA compared with sham acupuncture at the same time, we will conduct the exploratory research to compare the difference in the effectiveness between MA and EA by synthesizing the evidence from direct comparison and indirect comparison. (line 80-82 highlight in red)

3. Types of participants

The inclusion criteria for participants are too vague.

Response: We are sorry about the unclear description so we have made the section more precise. In the revised manuscript, we have stated “studies enrolling participants diagnosed as KOA will be included. The diagnostic criteria should be based on the American College of Rheumatology clinical criteria, National Institute for Health and Clinical Excellence guidelines or any other accepted guidelines. There will be no restrictions on their age, sex, race, education, economic status, Kellgren-Lawrence score or Outbridge score.” (line 182-186 highlight in red)

4. Criteria for excluding studies

The authors stated that they will not include studies with lack of allocation concealment. Why is allocation concealment more important than other risk of bias domains, such as a high drop-out rate? I recommend not excluding studies based on lack of allocation concealment. Please remember to add a table of excluded studies with reasons for exclusion and a reference.

Response: Thank you very much for your advice. According to your recommendation, we decide not to exclude studies based on lack of allocation concealment. Maybe we can exclude studies with lack of allocation concealment during the sensitivity analysis. Besides, we will add a table of exclude studies with reasons for exclusion to the appendix of our meta-analysis. (line 248-249, 322-323highlight in red)

5. Acupuncture adequacy assessment

I do not understand the following statement by the authors: “To test the success of the blinding, we will ask the two assessors to guess the provenance of each study.”.

Response: We are sorry about the unclear description and we have changed another understandable description. The acupuncture adequacy assessment instrument we used was invented by Berman’s group.[1] When we finish the data extraction, we will ask experienced acupuncturists to assess adequacy from 4 aspects: choice of acupuncture, number of sessions, needling technique, and experience of the acupuncturist. The whole assessment is only based on the study population and the acupuncture procedure. They make a judgment according to their own experience and the assessment is subjective. So the two acupuncturists will be blinded to the results of the study and publication. And we ask them to guess the provenance of each study in order to test the success of blinding.[1] (line 279-287)

[1]. Manheimer E, Linde K, Lao L, et al. Meta-analysis: acupuncture for osteoarthritis of the knee. *Annals of internal medicine* 2007;146:868-77.

6. Heterogeneity assessment

A high level of statistical heterogeneity is only an indication that the studies are methodologically heterogeneous. Therefore, the authors should not dismiss a meta-analysis with high statistical heterogeneity in advance.

https://www.meta-analysis-workshops.com/download/common-mistakes2.pdf

Response: We thank the reviewer for the suggestion. According to your recommendation, all the analyses will be based on the random-effects model because the RCTs included by us came from different populations. We have also planned to explore the possible sources of heterogeneity via meta-regression and subgroup analyses if high heterogeneity. (line 289-293 highlight in red)

7. Data synthesis

They authors stated that “The fixed-effects model will be used if little heterogeneity.”. This is a common mistake. Low statistical heterogeneity is only an indication that the studies are methodological homogeneous – the studies must be completely methodological homogeneous to be included in a fixed effects meta-analysis. Of note, a random effects meta-analysis will produce the same results as a fixed effects meta-analysis in the absence of statistical heterogeneity. Therefore, I recommend sticking to the random effects meta-analysis.

https://www.meta-analysis-workshops.com/download/common-mistakes2.pdf

Response: Thank you for your advice. We have read the articles recommended by you carefully. According to your recommendation, all the analyses will be based on the random-effects model because the RCTs included by us came from different populations. (line 303-304 highlight in red)

Reviewer: 2

Comments to the Author:

I first wanted to give my background to understand my comments on the proposed manuscript. I am a Family Physician who specializes in Sports Medicine. I am in academics now and am also a licensed acupuncturist. I am very interested whenever acupuncture is studied but especially in musculoskeletal disorders. I am not well versed in statistical analysis though.

My first impression is you are undertaking a massive endeavor. I will list some major comments and then list some minor recommendations from each page.

In general have found that acupuncture articles from Asia seem to always find a positive response to acupuncture in published articles compared to the US and Europe. Therefore, I am concerned about the inclusion of a great deal of potentially biased articles in the meta-analysis that may skew the results in a positive direction incorrectly.

Response: We thank the reviewer for the comments. As you said, the acupuncture trials from Asia are likely to show a higher positive rate than the acupuncture trials from western countries. One of the reasons is many negative result articles were not published in Asia. We have planned to search several clinical trials registry to collect gray literature as much as possible. Moreover, we will conduct a subgroup analysis based on the different studies' area.

I think you are looking for some conclusions in your outcomes that will be very difficult to find. I cannot see how you can "explore the applied law of different doses of acupuncture" from medical trials. There is just too much variability to compare different types of acupuncture and point selection not to mention the use of adjunct techniques like electroacupuncture or stimulation.

Response: We are sorry about the inaccurate description. Our group built a scoring instrument to calculate the dose of acupuncture from four parameters[1]. Based on the sum of the scores, we defined three doses of acupuncture treatment: high dosage, medium dosage and low dosage. In the meta-analysis, we will design three subgroups according the three kinds of dosage to explore the relationship between doses of acupuncture and efficacy. (line 365-373 highlight in red)

[1]. Sun N, Tu J, Lin L, et al. Correlation between acupuncture dose and effectiveness in the treatment of knee osteoarthritis: a systematic review. *Acupuncture in medicine* 2019;37:261-67.

I also have an issue with using sham as a control as it has been shown not to be a placebo control. There are published articles about the effects from sham acupuncture. I worry that using this as a control will possibly mask a positive effect compared to true placebo or wait list controls.

Response: We are sorry about the unclear description. The reason we choose sham acupuncture as a common comparator is to compare the difference in the effectiveness between MA and EA by synthesizing the evidence from direct comparison and indirect comparison. This section is only an exploratory research. For the whole meta-analysis, we have planned to compare needle acupuncture with sham acupuncture, analgesic, usual care or waiting list control groups in the “Types of control groups” section. (line 192-194)

Smaller comments:

1. Page 7, line 17-I find this hard to believe that 85% of OA is in the knee just comparing to my clinical practice and all other stats I have seen

Response: We are sorry about the inaccurate description. What we really want to describe is the disease burden instead of the prevalence. In the revised manuscript, we changed into “Knee osteoarthritis accounts for approximately 85% of global osteoarthritis burden”.[1](line 93-94 highlight in red)

[1].Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet (London, England)* 2016;388:1545-602.

2. Page 7, line 38-43-would amplify this section and use examples-which analgesics have no effect?

Response: We are sorry about the inaccurate description and we have amplified this section. In the revised manuscript, we changed into “acetaminophen (paracetamol) is not associated with long-term pain improvement”. Acetaminophen is an inexpensive analgesic, but it was not associated with long-term pain improvement in the present study.[1]But acupuncture was proved to have persistence of the effects for chronic pain.[2](line 103-104 highlight in red)

[1]. Gregori D, Giacovelli G, Minto C, et al. Association of Pharmacological Treatments With Long-term Pain Control in Patients With Knee Osteoarthritis: A Systematic Review and Meta-analysis. *JAMA* 2018;320:2564-79.

[2]. MacPherson H, Vertosick EA, Foster NE, et al. The persistence of the effects of acupuncture after a course of treatment: a meta-analysis of patients with chronic pain. *Pain* 2017;158(5):784-93.

3. Page 8, line 38-56-although these mediators might be involved, OA is in general not an inflammatory disorder but a pain disorder; I would like to see more included with pain pathways.

Response: In the revised manuscript, we have added some mechanisms related to pain pathways of acupuncture for KOA. See “In addition, CBR1-GABA-5-HT may be a novel pathway contributed to the

effect of EA on KOA pain. EA down-regulated IL-1β expression via activating the peripheral CBR2 to inhibit the KOA pain". (line 136-139 highlight in red)

4. Page 9, line 48-change "have" to "has"

Response: In the revised manuscript, we stated that "one systematic review has looked at the comparative effectiveness of manual acupuncture and electroacupuncture, but considered only direct evidence". (line 150 highlight in red)

5. Page 10, line14-Objectives-for question 2, I would find this almost impossible to study as I have rarely seen the amount of medication tracked or recorded in studies

Response: As you said, we found there are few studies recording the amount of medication during screening literatures. One paper reported the number of people using emergency analgesics (paracetamol).[1] Maybe we can make a qualitative description about the section instead of considering it as a primary objective. So we changed the "Is acupuncture associated with a reduction in medication use in patients with KOA?" into "Is there a difference in the efficacy between manual acupuncture and electroacupuncture?" in "Objectives" section. (line 165-166 highlight in red)

1. Tu JF, Yang JW, Shi GX, et al. Efficacy of Intensive Acupuncture Versus Sham Acupuncture in Knee Osteoarthritis: A Randomized Controlled Trial. *Arthritis Rheumatol* 2021;73:448-58.

6. Page 11-Types of participants-I would expand on this section and be more specific on what is considered OA and how diagnosed

Response: We are sorry about the unclear description so we have made the section more precise. We don't need to diagnose these patients because we include studies not patients. The RCTs included by us have detailed diagnostic criteria. In the revised manuscript, we have stated "studies enrolling participants diagnosed as KOA will be included. The diagnostic criteria should be based on the American College of Rheumatology clinical criteria, National Institute for Health and Clinical Excellence guidelines or any other accepted guidelines. There will be no restrictions on their age, sex, race, education, economic status, Kellgren-Lawrence score or Outbridge score." (line 182-186 highlight in red)

7. Page 11-Types of interventions-I often use moxabustion as an adjunct, was this considered as another type since it is used often? In a recent review we did on this same topic, we found moxa used often

Response: Thank you for your suggestion. Although both acupuncture and moxibustion belong to TCM, they are different in some ways. Their stimulation to acupoints is different. The acupuncture is mainly mechanical stimulation and moxibustion mainly involves light and heat stimulation. Our group has been committed to the studies of acupuncture. So only RCTs related to acupuncture will be included. And we think moxibustion for KOA can be a separate study in the future.

8. Page 12, line 14-again, how would you find this data to make any decent conclusions?

Response: As you said, we found there are few studies recording the amount of medication during screening literatures. One paper reported the number of people using emergency analgesics (paracetamol).[1] Maybe we can make a qualitative description about the section instead of considering it as a primary objective. So we stick to retaining the “drug use” section.

1. Tu JF, Yang JW, Shi GX, et al. Efficacy of Intensive Acupuncture Versus Sham Acupuncture in Knee Osteoarthritis: A Randomized Controlled Trial. *Arthritis Rheumatol* 2021;73:448-58.

9. Line 19-how would you find the cost effectiveness?

Response: The cost effectiveness of acupuncture is very important. What we want to do is a comprehensive meta-analysis from effectiveness, safety and cost effectiveness three different aspects. Some studies focused on the cost-effectiveness of acupuncture treatment.[1-3] But we only found one study related to acupuncture for KOA.[4] Maybe we can make a qualitative description to appeal to more health economics research in the future if few literatures are found.

[1]. Molassiotis A, Dawkins B, Longo R, et al. Economic evaluation alongside a randomised controlled trial to assess the effectiveness and cost-effectiveness of acupuncture in the management of chemotherapy-induced peripheral neuropathy. *Acupuncture in medicine* 2021;39(1):41-52.

[2]. Kim D, Lee Y, Park K, et al. The effectiveness and cost-effectiveness of motion style acupuncture treatment (MSAT) for acute neck pain: A multi-center randomized controlled trial. *Medicine* 2020;99(44):e22871.

[3]. Nicolian S, Butel T, Gambotti L, et al. Cost-effectiveness of acupuncture versus standard care for pelvic and low back pain in pregnancy: A randomized controlled trial. *PloS one* 2019;14(4):e0214195.

[4]. Reinhold T, Witt C, Jena S, et al. Quality of life and cost-effectiveness of acupuncture treatment in patients with osteoarthritis pain. *The European journal of health economics : HEPAC : health economics in prevention and care* 2008;9(3):209-19.

10. Line 25-I would change "symptoms" to "criteria"

Response: We have revised our manuscript. See “Participants with knee pain but no other criteria of KOA”. (line 211 highlight in red)

11. Page 13, line 12-delete patellofemoral syndrome-it is not OA

Response: We think chondromalacia patellae belongs to KOA and have changed “patellofemoral pain syndrome” into “chondromalacia patellae”. (line 225-226 highlight in red)

12. Page 14, line 43-why study choice of points? it will be next to impossible to make any strong conclusions and would be more appropriate in an acupuncture journal or TCM journal, not a prestigious medical journal. Also, with so many different types of acupuncture (TCM, Japanese, Scalp, Ear, French, etc.) and different ways to diagnose, how can you make conclusions?

Response: Like every meta-analysis related to acupuncture before us, we will describe the choice of points of each trial in appendix. Those who interested in the choice of acupoints can read the appendix.

13. Page 15, line 30-which methods?

Response: We are sorry about the unclear description. We will use a adequacy assessment instrument which was invented by Berman’s group[1]. When we finish the data extraction we will ask experienced acupuncturists to assess adequacy independently from 4 aspects: choice of acupuncture, number of sessions, needling technique, and experience of the acupuncturist. Then the adequacy will be graded “Adequate”, “Inadequate”, and “Unknown” according to the two acupuncturists’ experience. (line 279-283 highlight in red)

1. Manheimer E, Linde K, Lao L, et al. Meta-analysis: acupuncture for osteoarthritis of the knee. *Annals of internal medicine* 2007;146:868-77.

14. Page 18, line 7-grammatically incorrect sentence

Response: In the revised manuscript, we stated that “There will be no patients or public directly involved in this review”. (line 169 highlight in red)

15. Page 18, line 38-this is the first time you mentioned specific grading systems of OA, should include earlier if going to use it

Response: We have revised our manuscript. See “There will be no restrictions on their age, sex, race, education, economic status, Kellgren-Lawrence score or Outbridge score.” (line 186 highlight in red)

16. Page 19-once again, I would be careful using sham as a control

Response: We are sorry about the unclear description. The reason we choose sham acupuncture as a common comparator is to compare the difference in the efficacy between MA and EA by synthesizing the evidence from direct comparison and indirect comparison. This section is only an exploratory research. For the whole meta-analysis, we have planned to compare needle acupuncture with sham acupuncture, analgesic, usual care or waiting list control groups in the “Types of control groups” section.

17. Line 35-you likely need to explain the "DeQi" response; I find it hard to believe that a trial was done with enough patients to discern DeQi was associated with a treatment effect.

Response: Following your suggestion, we gave an explanation about DeQi in our revised manuscript. One RCT (n=338) suggested Among patients with Bell palsy, acupuncture with strong stimulation that elicited DeQi had a greater therapeutic effect, and stronger intensity of DeQi was associated with the better therapeutic effects.[1] We don’t want to explore the association between DeQi response and acupuncture effect. DeQi plays a role in acupuncture dosage so it is only one dimension of acupuncture dosage in our met-analysis. (line 370-373 highlight in red)

1. Xu S, Huang B, Zhang C, et al. Effectiveness of strengthened stimulation during acupuncture for the treatment of Bell palsy: a randomized controlled trial. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* 2013;185(6):473-9.

18. References-question use of #2-its 8 years old and an expert's concensus paper

#12, 14, 15, 22, 23, 25 are older papers-think these topics have much better and more recent research/papers

Response: Many thanks for the reviewer’s good suggestion. We have updated these citations. (line references 2, 11, 18, 19, 20, 27, 28 highlight in red)

19. #48-ACR has more updated guidelines than 1995

Response: As you said, the latest version of ACR is 2019. But the ACR2019 is a guideline related to treatment and the diagnostic criteria section of KOA has not been renewed since 1995. So we think maybe citing ACR1995 is more appropriate, like some studies did.[1, 2]

1. Wang Q, Lv H, Sun Z, et al. Effect of Electroacupuncture versus Sham Electroacupuncture in

Patients with Knee Osteoarthritis: A Pilot Randomized Controlled Trial. *Evidence-based complementary and alternative medicine : eCAM* 2020;2020:1686952.
2. Tu JF, Yang JW, Shi GX, et al. Efficacy of Intensive Acupuncture Versus Sham Acupuncture in Knee Osteoarthritis: A Randomized Controlled Trial. *Arthritis Rheumatol* 2021;73:448-58.

VERSION 2 – REVIEW

REVIEWER	Neilson Mathews Lehigh Valley Health Network, Family Medicine
REVIEW RETURNED	30-Sep-2021

GENERAL COMMENTS	I think this is a good research design for a complicated question I do think a brief discussion on the different types of acupuncture might be useful-not all types need a DeQi response during sessions Might want to clarify its assessing TCM types of treatments I also believe a comment on the use of sham acupuncture as a control should be included-there are several articles on this topic Some of the items in bibliography are older-#41, 42,46, 48-many of these topics have been updated and should have more recent references

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Comments to the Author:

I think this is a good research design for a complicated question

1. I do think a brief discussion on the different types of acupuncture might be useful-not all types need a DeQi response during sessions

Might want to clarify its assessing TCM types of treatments

Response: Many thanks for the reviewer's good suggestion. We have added a brief discussion about the topic and explore the association between effectiveness and TCM types of KOA in subgroup analyses. See line 315, 370-377 highlight in red

There are many different types of acupuncture including manual acupuncture, electroacupuncture, auricular acupuncture, wrist-ankle acupuncture, etc. On the one hand, not all types of acupuncture need a Deqi response during sessions. For example, manual acupuncture and electroacupuncture are required to cause a Deqi response while wrist-ankle acupuncture is not. Furthermore, electrical stimulation can enhance Deqi response elicited by manipulation of needles. Based on the different duration of active stimulation, it is necessary to compare the effectiveness of manual acupuncture and electroacupuncture.[1] On the other hand, Deqi response is more emphasized in China than Western.[2] In our study, we don't want to explore the association between DeQi response and acupuncture effect. So Deqi is only one dimension of our scoring instrument for dosage of acupuncture.

[1]. Langevin H, Schnyer R, MacPherson H, et al. Manual and electrical needle stimulation in acupuncture research: pitfalls and challenges of heterogeneity. *Journal of alternative and complementary medicine (New York, NY)* 2015;21(3):113-28.
[2]. Li Y. Who Has the Final Say on the Dose of Acupuncture? Comment on the Article by Tu et al. *Arthritis & rheumatology (Hoboken, NJ)* 2021;73(6):1089-90.

2. I also believe a comment on the use of sham acupuncture as a control should be included-there are several articles on this topic

Response: Thank you for your suggestion. We have added a comment on the use of sham acupuncture as a control. See line 378-385 highlight in red

Acupuncture has both specific effects caused by intervention itself and non-specific effects including patient-acupuncturist relationship, patient expectations, etc. Sham acupuncture group has usually been set in order to eliminate non-specific effects. The sham acupuncture can be divided into superficial insertion and non-penetrating insertion at traditional acupuncture points or not.[1] Superficial insertion is not a physiologically inert procedure and thus decreases the difference between groups.[2] Therefore, more and more trials choose non-penetrating sham acupuncture at non-acupoints as control to minimize the physiological effects of sham acupuncture.

[1]. Ho R, Wong C, Wu J, et al. Non-specific effects of acupuncture and sham acupuncture in clinical trials from the patient's perspective: a systematic review of qualitative evidence. *Acupuncture in medicine* 2021;39(1):3-19.
[2]. Sun Y, Liu Y, Liu B, et al. Efficacy of Acupuncture for Chronic Prostatitis/Chronic Pelvic Pain Syndrome : A Randomized Trial. *Annals of internal medicine* 2021

3. Some of the items in bibliography are older-#41, 42, 46, 48-many of these topics have been updated and should have more recent references

Response: Thank you for your suggestion. For references #41, 42, the two previous meta-analyses are old and we have deleted them. For #46, 48, the latest version of ACR is 2019. But the ACR2019 is a guideline related to treatment and the diagnostic criteria section of KOA has not been renewed since 1995. And the Outbridge score has also not been renewed since 1961. So we think maybe citing ACR1995 and Outbridge 1961 are more appropriate.