BMJ Open

Risk of bias and methodological issues in randomised controlled trials of acupuncture for knee osteoarthritis: a cross-sectional study

Pengli Jia,[1] Li Tang,[1] Jiajie Yu,[1] Andy H Lee,[2] Xu Zhou,[3] Deying Kang,[1] Yanan Luo,[4] Jiali Liu,[1] Xin Sun[1]

To cite: Jia P, Tang L, Yu J, *et al*. Risk of bias and methodological issues in randomised controlled trials of acupuncture for knee osteoarthritis: a cross-sectional study. *BMJ Open* 2018;**8**:e019847. doi:10.1136/bmjopen-2017-019847

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2017-019847).

PJ, LT and JY contributed equally.

Received 29 September 2017
Revised 28 December 2017
Accepted 11 January 2018

[1]Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University, Chengdu, China
[2]School of Public Health, Curtin University, Perth, Western Australia, Australia
[3]Evidence-based Medicine Research Center, Basic Medical College, Jiangxi University of Traditional Chinese Medicine, Chegndu, China
[4]Acupuncture and Tuina School, Chengdu University of Traditional Chinese Medicine, Chengdu, China

Correspondence to
Dr Xin Sun;
sunx79@hotmail.com

ABSTRACT

Objective To assess risk of bias and to investigate methodological issues concerning the design, conduct and analysis of randomised controlled trials (RCTs) testing acupuncture for knee osteoarthritis (KOA).

Methods PubMed, EMBASE, Cochrane Central Register of Controlled Trials and four major Chinese databases were searched for RCTs that investigated the effect of acupuncture for KOA. The Cochrane tool was used to examine the risk of bias of eligible RCTs. Their methodological details were examined using a standardised and pilot-tested questionnaire of 48 items, together with the association between four predefined factors and important methodological quality indicators.

Results A total of 248 RCTs were eligible, of which 39 (15.7%) used computer-generated randomisation sequence. Of the 31 (12.5%) trials that stated the allocation concealment, only one used central randomisation. Twenty-five (10.1%) trials mentioned that their acupuncture procedures were standardised, but only 18 (7.3%) specified how the standardisation was achieved. The great majority of trials (n=233, 94%) stated that blinding was in place, but 204 (87.6%) did not clarify who was blinded. Only 27 (10.9%) trials specified the primary outcome, for which 7 used intention-to-treat analysis. Only 17 (6.9%) trials included details on sample size calculation; none preplanned an interim analysis and associated stopping rule. In total, 46 (18.5%) trials explicitly stated that loss to follow-up occurred, but only 6 (2.4%) provided some information to deal with the issue. No trials prespecified, conducted or reported any subgroup or adjusted analysis for the primary outcome.

Conclusion The overall risk of bias was high among published RCTs testing acupuncture for KOA. Methodological limitations were present in many important aspects of design, conduct and analyses. These findings inform the development of evidence-based methodological guidance for future trials assessing the effect of acupuncture for KOA.

Strengths and limitations of this study

- The literature search was comprehensive, and detailed inclusion and exclusion criteria were developed to ensure transparency and reproducibility in the judgements.
- A standardised and pilot-tested questionnaire was used to systematically assess the methodological issues concerning the design, conduct and analysis of randomised controlled trials testing acupuncture for knee osteoarthritis.
- The insufficient reporting of methodological details among those published trials limited our assessment of the methodological quality and identification of specific methodological gaps.
- In addition, this study did not examine the specification of patient population in those trials, and the findings may not be applicable to other acupuncture trials.

BACKGROUND

Knee osteoarthritis (KOA) is a major health concern worldwide and its disease burden continues to increase due to population ageing.[1 2] Standard treatments for KOA, such as non-steroid anti-inflammatory drugs, glucosamine and intra-articular injections, are often of limited effects.[3–6] Acupuncture has received increasing attention,[7 8] and is becoming popular in many Western countries.[2] Approximately one million American patients with musculoskeletal disorders are now treated with acupuncture annually.[9]

A number of randomised controlled trials (RCTs) have been conducted to test the effects of acupuncture on KOA. However, the evidence so far is weak, and the majority of trials had low methodological quality.[10–12] Due to the lack of compelling evidence, acupuncture has not been recommended as a treatment option for KOA in the most recent clinical practice guidelines.[2 13–15]

Although there is strong consensus that efforts should be made to improve the quality of such trials, no established methodological guidance is available to inform researchers about how to rigorously design, conduct and

analyse a trial of acupuncture for KOA. Unlike classical drug trials, RCTs of acupuncture have many specific features. For instance, acupuncture, as a typical complex intervention, is often expertise/preference-based, varies in practice and contains multiple components. Patient characteristics and study settings are also complex. Blinding may be challenging. Thus, there are special issues concerning the design, conduct and analysis of RCTs of acupuncture for KOA. Our ultimate goal is to develop a methodological guidance for rigorous design, conduct and analysis of acupuncture RCTs for KOA. Understanding methodological characteristics of previously published RCTs testing acupuncture for KOA would be the first and critical step.

Therefore, we conducted a cross-sectional study of published RCTs specifically testing acupuncture for KOA, so as to examine their methodological details. The findings provide critical information and valuable insights to identify important methodological gaps and support the development of methodological guidance.

METHODS

Eligibility criteria

An eligible study had to meet all of the following inclusion criteria: (1) a two-arm RCT, regardless of study type (superiority, non-inferiority or equivalence); (2) enrolled patients diagnosed with KOA; (3) tested the effect of acupuncture in an uncounfunded manner. That is, acupuncture may be used alone or in combination with other treatments, as long as the effects of acupuncture can be parcelled out. The types of acupuncture may include electro-acupuncture, filiform needle, fire needle, silver needle, dry needle, laser acupuncture, ear acupuncture and scalp acupuncture; (4) compared with another type of acupuncture, pharmacological intervention, placebo acupuncture, sham acupuncture, waiting list or physical treatments (eg, exercises and weight loss); and (5) reported findings in the full text of the paper.

We excluded study protocol from our study. Trials reported in research letters or conference abstracts were also excluded.

Literature search

We searched PubMed, EMBASE, Cochrane Central Register of Controlled Trials as well as four Chinese databases, namely Chinese Biomedical Database, China National Knowledge Infrastructure, Wanfang and China Science and Technology Journal Database (VIP). The search was conducted from inception to February 2017 and for articles published in either English or Chinese. A combination of keywords and Medical Subject Headings related to KOA, acupuncture and RCT was used after consulting an experienced librarian (online supplementary appendix 1). We also manually screened reference lists of eligible papers for articles which were not identified by the computerised search.

Study process

Using standardised and pilot-tested forms, two authors (PLJ and JLL) screened the titles and abstracts independently and in duplicate for potential eligibility. They subsequently read the full texts to determine final eligibility. Then, from the eligible RCTs, two other authors (PLJ and XZ) assessed their risk of bias and collected methodological details on design, conduct and analysis using a study-specific questionnaire comprising 48 items (see the 'Study questionnaire' section below). A third reviewer (XS) was involved as the adjudicator in case of discrepancy throughout the process.

Risk of bias assessment

We assessed the risk of bias of included RCTs using the Cochrane Risk of Bias tool. This tool contains seven items: (1) randomisation sequence generation; (2) allocation concealment; (3) blinding of participants and personnel; (4) blinding of outcome assessors; (5) incomplete outcome data; (6) selective reporting; and (7) other sources of bias. Each item was assigned one of three responses: 'low risk of bias, high risk of bias or unclear risk of bias' according to the Cochrane handbook.

Study questionnaire

In order to fully examine the methodological details of eligible RCTs, the research team, consisting of clinical trial experts, statisticians and acupuncturists, developed a study-specific questionnaire.

In developing the questionnaire, we first reviewed existing tools for assessing methodological quality and reporting (eg, Cochrane Risk of Bias Tool and CONSORT), and drafted a list of items. Based on those items, the research team, through brainstorming, added other potentially useful items. Afterwards, the research team undertook six group discussions to ascertain relevancy of each questionnaire item and clarify its definition. Items that only addressed issues of reporting were deleted due to out of the study scope. The resulting questionnaire was pilot-tested using 10 studies randomly selected from the eligible articles. Thereafter, the items were revised based on the pilot test results.

The final version of the study questionnaire (online supplementary appendix 2) consisted of 48 items covering the following aspects: general information (eg, publication year, scope of journal and funding sources), study design and conduct (eg, randomisation, allocation concealment, standardisation of acupuncture procedures, blinding and sample size calculation) and data analysis (methods for dealing with loss to follow-up, subgroup analysis and adjustment for confounders).

Statistical analysis

We qualitatively summarised the general characteristics about the design, conduct and data analysis of included studies. Number and percentage were used for multinomial or dichotomous outcomes, whereas median and IQR were calculated for continuous variables.

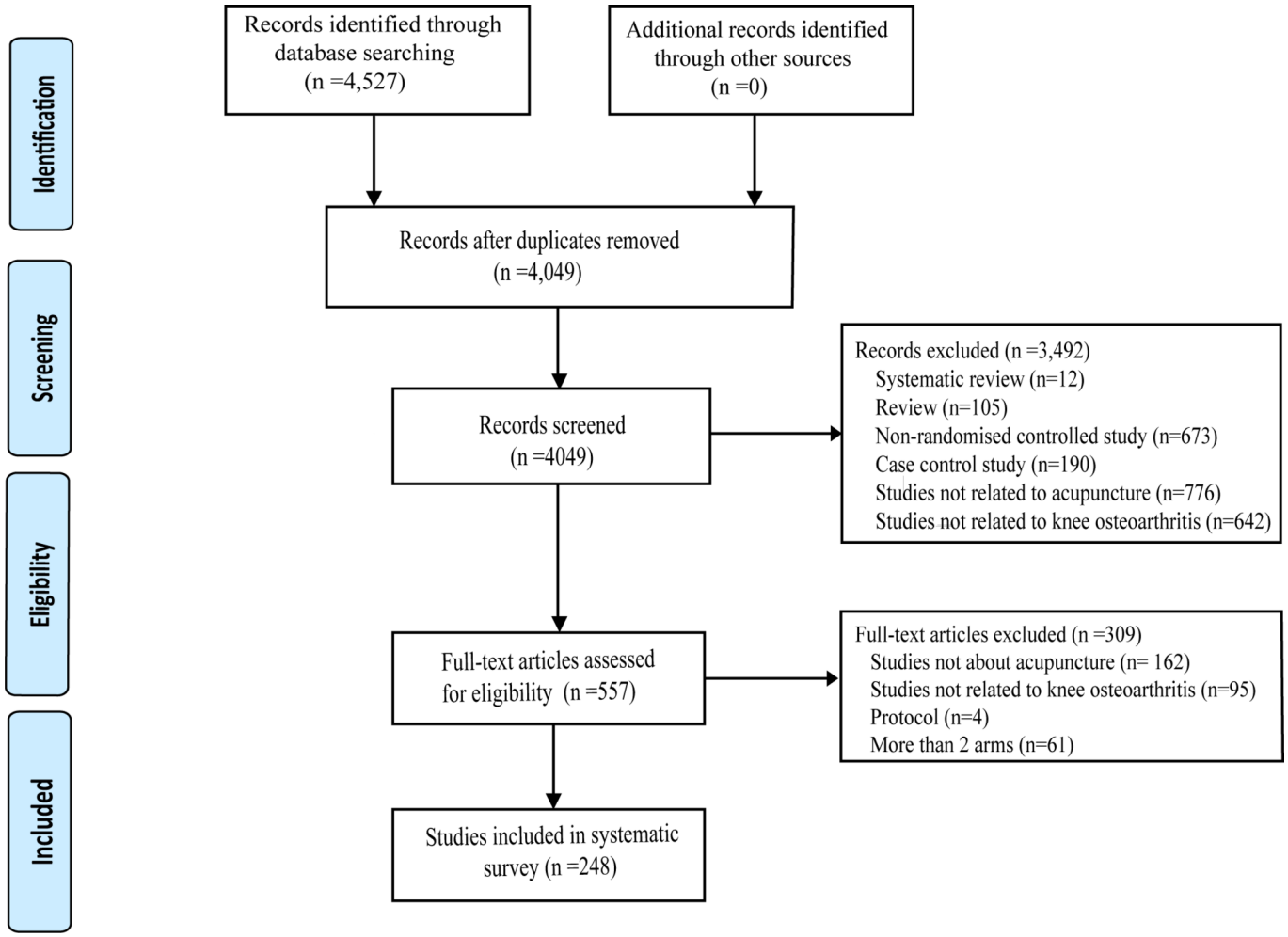

Figure 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart for search and selection processes.

In order to examine the factors associated with major methodological quality indicators, we conducted four separate multivariable regression analyses for the following: adequate generation of randomisation sequence (yes vs no or unclear), allocation concealment (yes vs no or unclear), blinding of participants and personnel (yes vs no or unclear) and blinding of outcome assessors (yes vs no or unclear). For each of the regression analyses, we included four prespecified factors, namely, language (English vs Chinese), multiple centre study (yes vs no or unclear), funding support (yes vs no or unclear) and sample size (categorised according to the median).

RESULTS

Our search resulted in 4527 reports, of which 557 were potentially eligible after title and abstract screening. On reviewing their full texts, 248 RCTs were finally included (figure 1). No additional RCTs were identified through screening of their reference lists.

General characteristics of included trials

The 248 RCTs (219 in Chinese and 29 in English) were published between 1992 and 2016. Table 1 summarises general characteristics of the included trials. The majority were single-centre trials (92.7%) published in a non-acupuncture related journal (76.6%). The median sample size was only 76 (range 20–632), and the median duration of follow-up was 4 weeks (range 1–48 weeks). Only one-quarter of the included trials reported sources of funding and all of them were funded by non-profit funding agencies. The median number of authors was 3 (range 1–13).

Risk of bias assessment of included trials

We found that almost half of the 248 trials were at low risk of bias for randomisation sequence generation (119, 48.0%) and incomplete outcome data (105, 42.3%). But only a small proportion of trials were at low risk of bias for allocation concealment (31, 12.5%), blinding of participants and personnel (15, 12.5%), blinding of outcome assessors (15, 12.5%), selective reporting (9, 3.6%) and other sources of bias (7, 2.8%) (figure 2, online supplementary appendix 3). In general, the trials had a high risk of bias.

Table 1 General characteristics of included trials

Features of included RCTs	Total (n=248,%)	Chinese (n=219,%)	English (n=29,%)
No. of authors (median (IQR))	3 (2–5)	5 (4–7.5)	3 (1–4)
Sample size (median (IQR))	76 (60–100)	80 (60–105)	60 (40–90.5)
≤76*	126 (50.8)	105 (47.9)	21 (72.4)
>76	122 (49.2)	114 (52.1)	8 (27.6)
Single versus multicentre study			
Single centre	230 (92.7)	202 (92.2)	28 (96.6)
Multicentre	18 (7.3)	17 (7.8)	1 (3.4)
Length of follow-up			
≤4 weeks	162 (65.3)	153 (69.9)	9 (31.0)
>4 weeks	82 (33.1)	64 (29.2)	18 (62.1)
Unclear	4 (1.60)	2 (0.9)	2 (6.9)
Scope of journals			
Acupuncture-related	58 (23.4)	48 (21.9)	10 (34.5)
Others (non-acupuncture)	190 (76.6)	171 (78.1)	19 (65.5)
Sources of funding			
Non-profit funding agencies	62 (25.0)	46 (21.0)	16 (55.2)
Unclear	186 (75.0)	173 (79.0)	13 (44.8)

*Median.
RCT, randomised controlled trial.

Methodological details of included trials

Randomisation and allocation concealment

Of those 248 RCTs, 99 (39.9%) did not provide details on randomisation, and a large proportion (217, 87.5%) failed to specify the method for allocation concealment (table 2). Among those which provided details on randomisation (149, 60.1%), only 39 (26.2%) used a computer-generated randomisation sequence (the preferred approach), 70 (47.0%) used a random number table, whereas others (40, 26.8%) adopted inadequate methods (eg, coin tossing, visit order) to generate the randomisation sequence. Very few trials used blocked randomisation (15, 6.0%) and stratification at randomisation (9, 3.6%).

Of the 31 trials (12.5%) that provided details about allocation concealment, only 1 (3.2%) reported central randomisation, whereas 30 trials (96.8%) used sequentially numbered, opaque or sealed envelope methods. No details were available from other trials (217, 87.5%). Given their applied methods of generation, we judged that most of trials were unlikely to conceal treatment allocation (see table 2).

Intervention and blinding

All 248 trials described interventions for each group (table 3). Types of acupuncture used in the interventions group were filiform/ordinary needle (107, 43.1%), electro-acupuncture (98, 39.5%), fire needle (14, 5.7%),

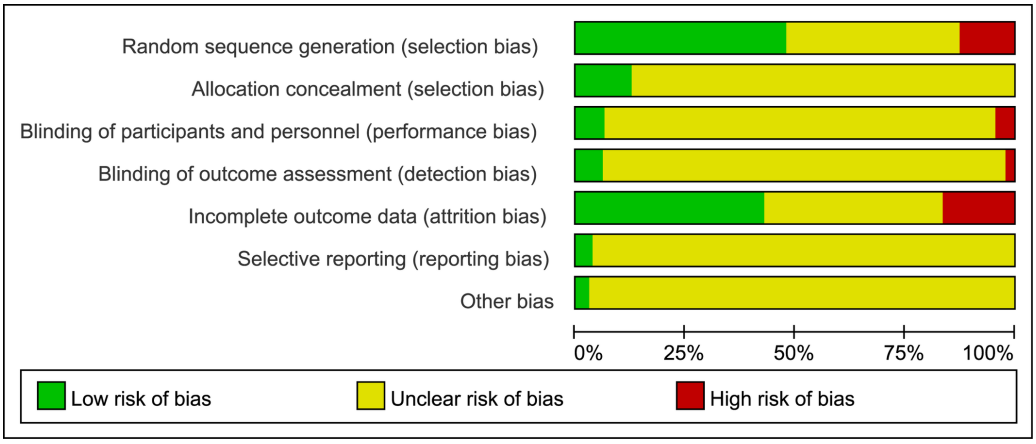

Figure 2 Risk of bias assessment of randomised controlled trials of acupuncture for knee osteoarthritis (n=248).

Table 2 Randomisation and allocation concealment of included trials

Items	Total (n=248,%)	English (n=29,%)	Chinese (n=219,%)
Methods of randomisation sequence generated			
Reported	149 (60.1)	17 (58.6)	132 (60.3)
A random number table	70 (47.0)	4 (23.5)	66 (50.0)
Computer-generated randomisation number	39 (26.2)	11 (64.7)	28 (21.2)
Coin tossing	2 (1.3)	1 (5.9)	1 (0.8)
Shuffling cards or envelopes	2 (1.3)	0 (0.0)	2 (1.5)
Drawing of lots	6 (4.1)	0 (0.0)	6 (4.5)
Visit order	30 (20.1)	1 (5.9)	29 (22.0)
Unclear	99 (39.9)	12 (41.4)	87 (39.7)
Methods of concealment of allocation			
Reported	31 (12.5)	10 (35.5)	21 (9.6)
Central allocation (telephone, web-based)	1 (3.2)	1 (10.0)	0 (0.0)
Sequentially numbered, opaque, sealed envelopes	30 (96.8)	9 (90.0)	21 (100.0)
Unclear	217 (87.5)	19 (65.5)	198 (90.4)
Use of block randomisation			
Yes	15 (6.0)	11 (37.9)	4 (1.8)
Unclear	233 (94.0)	18 (62.1)	215 (98.2)
Use of stratification at randomisation			
Yes	9 (3.6)	4 (13.8)	5 (2.3)
Unclear	239 (96.4)	25 (86.2)	214 (97.7)

laser acupuncture (4, 1.6%) and non-conventional acupuncture (20, 8.1%), such as silver, Fu's subcutaneous, blade and longitudinal needle. For the control group, the majority (184, 74.2%) administered a single intervention: acupuncture and moxibustion (51, 20.6%), pharmacological intervention (45, 18.1%), needle-knife (29, 11.7%), physical treatments (22, 8.9%), acupuncture (21, 8.5%), sham acupuncture (8, 3.2%), placebo acupuncture (5, 2.0%), moxibustion (2, 0.8%) and waiting list (1, 0.4%). Multiple interventions, such as pharmacological intervention in conjunction with acupuncture/physical treatments, were used in the remaining trials (64, 25.8%).

Standardisation of acupuncture procedures was poor across trials. Only 25 (10%) trials stated that acupuncture procedure was standardised, and only 18 (7.3%) explicitly specified that the standardisation was conducted according to manual, specific guidelines or a standardised protocol. A small proportion of the trials (18, 7.3%) clearly stated that acupuncture was performed by acupuncturists, physicians and/or physiotherapists.

The great majority of trials (233, 94%) stated that blinding was in place, but most of them (204, 87.6%) did not clarify who was blinded. The blinding of patients, acupuncturists, outcome assessors or data analysts was mentioned in <10% of the RCTs. Only 12 trials (4.8%) stated that blinding was achieved by sham procedures, such as placebo acupuncture or sham acupuncture (table 3).

Sample size calculation

Only a small number of trials (17, 6.9%) included details regarding the sample size calculation (table 4). Among those 17 trials, 10 performed the calculation based on a defined primary outcome, 9 defined the α level, 12 gave the β value, 9 took into account the potential loss to follow-up and 11 used data from a pilot/previous study. None of the trials preplanned an interim analysis and associated stopping rule.

Primary outcome and data analysis

Only 27 (10.9%) RCTs specified a primary outcome of interest (eg, pain, function and quality of life, table 5). Among them, six mentioned the statistical analytic method, and seven reported intention-to-treat analysis for the primary outcome.

In total, 145 trials (58.5%) provided information regarding loss to follow-up, and 46 (31.7%) trials explicitly stated that loss to follow-up occurred. Among these 46 RCTs, 37 (80.4%) trials reported loss to follow-up separately for each arm, but only one compared baseline characteristics of loss to follow-up patients between treatment and control groups. Only six trials (13.1%) provided some information to deal with the problem. However, they did not specify the exact analytical methods for dealing with loss to follow-up in the primary outcome. Furthermore, none of trials prespecified, conducted or reported any subgroup analysis or adjustment for confounding variables (table 5).

Table 3 Interventions and blinding of included trials

Items	Total (n=248,%)	English (n=29,%)	Chinese (n=219,%)
Type of the acupuncture of the experimental group			
Filiform/ordinary needle	107 (43.1)	20 (69.0)	87 (39.7)
Electro-acupuncture	98 (39.5)	4 (13.8)	94 (42.9)
Fire needle	14 (5.7)	0 (0)	14 (6.4)
Laser acupuncture	4 (1.6)	4 (13.8)	0 (0.0)
Non-conventional acupuncture*	20 (8.1)	1 (3.4)	19 (86.8)
Combination of different types of acupuncture	5 (2.0)	0 (0)	5 (2.3)
The intervention of the control group			
Multiple interventions†	64 (25.8)	2 (6.9)	62 (28.3)
Acupuncture and moxibustion	51 (20.6)	2 (6.9)	49 (22.4)
Pharmacological interventions	45 (18.1)	2 (6.9)	43 (19.6)
Needle-knife	29 (11.7)	1 (3.4)	28 (12.8)
Physical treatments	22 (8.9)	3 (10.3)	19 (8.7)
Acupuncture	21 (8.5)	5 (17.2)	16 (7.3)
Filiform needle	9 (42.9)	1 (3.4)	8 (3.7)
Electro-acupuncture	8 (38.1)	1 (3.4)	7 (3.2)
Fire needle	4 (19.0)	0 (0)	4 (1.8)
Sham acupuncture	8 (3.2)	8 (27.6)	0 (0)
Placebo acupuncture	5 (2.0)	5 (17.2)	0 (0)
Moxibustion	2 (0.8)	0 (0)	2 (0.9)
Waiting list	1 (0.4)	1 (3.4)	0 (0)
Standardisation of the procedure			
Acupuncture procedure was standardised	25 (10.1)	9 (31.0)	16 (7.3)
Unclear	223 (89.9)	20 (69.0)	203 (92.7)
How was the acupuncture procedure standardised?			
Reported	18 (7.3)	8 (27.6)	10 (4.6)
According to written manuals or specific guidelines	11 (61.1)	2 (25.0)	9 (90.0)
Standardised acupuncture protocol	7 (38.9)	6 (75.0)	1 (10.0)
Unclear	230 (92.7)	21 (72.4)	209 (95.4)
Reasoning for the acupuncture procedure			
Reported	84 (33.9)	13 (44.8)	71 (32.4)
Theory	60 (71.4)	7 (53.8)	53 (74.7)
Literature	3 (3.6)	2 (15.4)	1 (1.4)
Consensus methods	21 (25.0)	4 (30.8)	17 (23.9)
Unclear	164 (66.1)	16 (55.2)	148 (67.6)
Who undertook acupuncture procedures?			
Acupuncturist/doctor/physiotherapists	18 (7.3)	15 (51.7)	3 (1.4)
Unclear	230 (92.7)	14 (48.3)	216 (98.6)
Blinding			
Any blinding	233 (94.0)	23 (79.3)	210 (95.9)
Patient			
Yes	16 (6.9)	14 (60.9)	2 (1.0)
No	11 (4.7)	6 (26.12)	5 (2.4)
Acupuncturist			

Continued

Table 3 Continued

Items	Total (n=248,%)	English (n=29,%)	Chinese (n=219,%)
Yes	0 (0.0)	0 (0.0)	0 (0.0)
No	21 (9.0)	16 (69.6)	5 (2.4)
Outcome assessor			
Yes	15 (6.4)	12 (52.2)	3 (1.4)
No	5 (2.1)	5 (21.7)	0 (0.0)
Data analyst			
Yes	10 (4.3)	7 (30.4)	3 (1.4)
No	5 (2.1)	5 (21.7)	0 (0.0)
Details not reported‡	204 (87.6)	7 (30.4)	197 (93.8)
Unclear	15 (6.0)	6 (20.7)	9 (4.1)

*Non-conventional acupuncture including silver needle, Fu's subcutaneous, blade needle, longitudinal needle, ear needle and head needle.
†Multiple interventions including acupuncture or/and pharmacological or/and physical treatments in the control group.
‡Trials just stated as 'single-blinded/double-blinded/blinded' without giving further information.

Factors associated with methodological quality indicators

The multiple regression analysis showed that no prespecified factors were statistically associated with better use of method for generating randomisation sequence; results of which are summarised in table 6. However, trials published in English (OR 4.99, 95% CI 1.72 to 14.47) and conducted in multiple centres (OR 4.49, 95% CI 1.39 to 14.53) were associated with a higher likelihood of achieving allocation concealment. Similarly, trials published in English were more likely to achieve blinding of participants and personnel or outcome assessors (OR 230.74, 95% CI 19.93 to 2671.65; OR 199.82, 95% CI 22.97 to 1738.10, respectively) after adjusting for the influence of funding, sample size and study centres. Multicentre trials were associated with a higher likelihood of achieving blinding of outcome assessors (OR 65.47, 95% CI 6.03 to 710.78).

DISCUSSION

Main findings

In this study, we found that the published RCTs testing acupuncture for KOA were at high risk of bias. The methodological quality issues were, however, complicated with inadequate reporting of trials. We found that 66.1% of trials were unclear in the specification of acupuncture procedure, 92.7% in procedure standardisation, 39.9% in randomisation and 87.5% in concealment. The apparent lack of rigorous methods was in part due to the poor reporting of trials.

However, one may reasonably infer that adequate use of rigorous methodologies for design, conduct and analyses is less likely among trials with serious limitations in reporting. We found that, among those trials which provided methodological details, many used inadequate methods to generate a randomisation sequence, such as coin tossing and alternate visit number; 12.5% of trials performed allocation concealment, but only one used central randomisation, the optimal approach to achieve adequate allocation concealment. In addition, very few trials considered block randomisation and stratification at randomisation, the two important approaches to achieving prognostic balance. These two techniques are particularly useful for acupuncture trials, since the study population is often present with a variety of characteristics

Acupuncture represents a practitioner-dependent, complex intervention.[16] Acupuncture interventions are subjected to a high degree of variability in terms of the acupuncturist's expertise and styles.[16 17] Careful consideration of standardisation of acupuncture procedures is an important aspect in conducting high-quality trials, and explicit description of the acupuncture procedure is essential to determine whether the results are generalisable.[18] Our study found that many types of acupuncture were used. However, only 10% of trials explicitly stated whether and how acupuncture procedures were standardised, and 7.3% of trials specified who would be qualified for operating the procedures.

Another relevant issue is blinding. Our study found that 94% of trials reported blinding, but 88% of them did not specify who were blinded; only 5% of trials attempted to use sham procedures. Blinding appeared to be challenging in the setting of acupuncture trials. Sometimes, caregivers, who administer acupuncture procedures, and patients may not be blinded if interventions between treatment and control groups have varying features. However, trialists should always consider blinding of both caregivers and patients under all circumstances. Unfortunately, nearly no trials clearly justified why blinding was not implemented. In addition, others parties such as outcome assessors can always be masked. This is particularly relevant for trials of KOA, since the main outcomes are often patient centred (such as pain, disability, quality of life and global perceived effect) and subjective.[19]

Table 4 Sample size calculation of included trial

Items	Total (n=248, %)	English (n=29, %)	Chinese (n=219, %)
Was sample size calculation conducted?			
Reported	17 (6.9)	14 (48.3)	3 (1.4)
Yes, details of the calculation reported	17 (100.0)	14 (100.0)	3 (100.0)
Yes, but no statistical details	0 (0.0)	0 (0.0)	0 (0.0)
No	0 (0.0)	0 (0.0)	0 (0.0)
Unclear	231 (93.1)	15 (51.7)	216 (98.6)
Was the sample size calculated based on the defined primary outcome?			
Reported	17 (100.0)	14 (100.0)	3 (100.0)
Yes	10 (58.8)	9 (64.3)	1 (33.3)
Unclear	7 (41.2)	5 (35.7·)	2 (66.7)
Did the trial define α?			
Reported	17 (100.0)	14 (100.0)	3 (100.0)
Yes	9 (52.9)	9 (64.3)	0 (0.0)
Unclear	8 (47.1)	5 (35.7)	3 (100.0)
Did the trial define β?			
Reported	17 (100.0)	14 (100.0)	3 (100.0)
Yes	12 (70.6)	12 (85.7)	0 (0.0)
Unclear	5 (29.4)	2 (14.3)	3 (100.0)
What basis was the sample size calculated?			
Reported	11 (64.7)	9 (64.3)	2 (66.7)
Pilot study	2 (18.2)	2 (22.2)	0 (0.0)
Previous study	9 (81.8)	7 (77.8)	2 (100.0)
Unclear	6 (35.3)	5 (35.7)	1 (33.3)
Was the loss to follow-up considered?			
Reported	11 (65.7)	8 (57.1)	3 (100)
Yes	9 (81.8)	6 (75.0)	3 (100.0)
No	2 (18.2)	2 (25.0)	0 (0.0)
Unclear	6 (35.3)	6 (42.9)	0 (0.0)

The lack of specifying primary outcomes and calculating sample sizes are two other concerning issues. Our study showed that only 10.9% of trials specified a primary outcome, much lower than previous studies (45% and 73%),[20 21] and only 6.9% clearly articulated sample size calculation. Such lack of explicit specification of primary outcomes suggested that trialists did not carefully consider this issue during design, despite the specification of primary outcome is highly relevant to the main study objective and subsequent sample size calculation. Moreover, there is strong evidence that significant outcomes are more likely to be reported.[22] When the primary outcomes were not prespecified, the authors tended to selectively report outcomes based on statistical significance instead of the predefined clinical importance,[20] thereby portraying an overestimation of the treatment effects. Similarly, very few (10 trials) determined the sample size based on the defined primary outcome, and included previous data to support sample size calculation. None of the trials considered interim analyses and associated stopping rule, even though a few trials had adopted large sample sizes and should have the opportunity to conduct close monitoring of trials through interim analysis.

Analysis of trial data has posed another issue. Our study found that most trials failed to adequately deal with loss to follow-up. Very few had prespecified how the primary outcome was analysed. Even more surprisingly, no trial explicitly considered subgroup analyses and adjustment for potential confounding effects. Although these conditions are not always present in RCTs, particularly those with smaller sample sizes, their absence of consideration further implicates the weak planning and analysis of these trials.

Overall, our study revealed that the published trials testing acupuncture for KOA are at high risk of bias and identified that many methodological details are not carefully considered during the various stages of planning, design, conduct and analysis.

Table 5 Primary outcome and data analyses of included trials

Items	Total (n=248,%)	English (n=29,%)	Chinese (n=219,%)
Specified primary outcome			
Yes	27 (10.9)	11 (37.9)	16 (7.3)
Unclear	221 (89.1)	18 (62.1)	203 (92.7)
ITT principle used for primary outcome			
Reported	27 (100.0)	11 (100.0)	16 (100.0)
ITT analysis	7 (25.9)	5 (45.5)	2 (12.5)
Not reported	20 (74.1)	6 (54.5)	14 (87.5)
Type of the primary outcome(s)			
Reported	27 (100.0)	11 (100.0)	16 (100.0)
Patient-reported outcome	25 (92.6)	9 (81.8)	16 (100.0)
Laboratory results	1 (3.7)	1 (9.1)	0 (0.0)
Patient recruitment rate	1 (3.7)	1 (9.1)	0 (0.0)
Did LTFU explicitly stated			
Reported	145 (58.5)	27 (93.1)	118 (53.9)
Yes, explicit statement: LTFU occurred	46 (31.7)	20 (74.1)	26 (22.0)
Yes, explicit statement: LTFU did not occur	99 (68.3)	7 (25.9)	92 (78.0)
No, no explicit statement about LTFU, unclear	103 (41.5)	2 (6.9)	101 (46.1)
LTFU reported separately for the all arms			
Reported	46 (100.0)	20 (100.0)	26 (100.0)
Yes	37 (80.4)	18 (90.0)	19 (73.1)
No	9 (19.6)	2 (10.0)	7 (26.9)
Comparison of baseline characteristic of LTFU			
Reported	46 (100.0)	20 (100.0)	26 (100.0)
Yes	1 (2.2)	1 (5.0)	0 (0.0)
No	45 (97.8)	19 (95.0)	26 (100.0)
Analytical methods for LTFU data			
Reported	6 (13.1)	6 (30.0)	0 (0.0)
Last observation carried forward	4 (66.6)	4 (66.6)	0 (0.0)
The worst of the scores obtained for the intervention group and the best obtained for the control group	1 (16.7)	1 (16.7)	0 (0.0)
Baseline observation carried forward was used for missing data	1 (16.7)	1 (16.7)	0 (0.0)
Unclear	40 (86.9)	14 (70.0)	26 (100.0)

ITT, intention to treat; LTFU, loss to follow-up.

Implications for future trials

Our findings have important implications for future trials of acupuncture for KOA. First, one should centrally randomise patients based on a computer-generated randomisation sequence. Blocked randomisation with variable sizes may also improve balance of sample sizes between groups. In case of important prognostic factors, stratification should be introduced to further improve the balance of these factors.

All acupuncture trials of KOA should predefine primary outcomes in the protocol and follow the CONSORT guidelines (2010) to register the trial at inception. Given the predefined primary outcome, careful calculation of sample sizes should be performed with an appropriate alpha level and statistical power, and including important clinical parameters such as expected treatment effect, baseline risk and expected loss to follow rate. In all possible cases, trial investigators should justify the choices of the primary outcome and parameters for the sample size calculation.

Designing and practicing acupuncture interventions are important. Trial investigators should be explicit in their design of acupuncture procedures. In any case, one should consider standardising issues such as the type of acupuncture, acupoints and other procedural issues (eg, number of needle insertions/treatment sessions

Table 6 Factors associated with methodological quality indicators

Items	OR and 95% CI			
	Crude	P value	Adjusted*	P value*
Randomisation sequence generated				
English vs Chinese (ref)	1.39 (0.64 to 3.02)	0.41	1.09 (0.45 to 2.62)	0.85
Multiple vs single centre (ref)	1.78 (0.66 to 4.75)	0.25	1.63 (0.57 to 4.67)	0.37
Funding vs unclear (ref)	1.88 (1.05 to 3.37)	0.04	1.75 (0.94 to 3.26)	0.08
Sample size>76 vs ≤76 (ref)	0.75 (0.45 to 1.23)	0.25	0.73 (0.43 to 1.22)	0.23
Allocation concealment				
English vs Chinese (ref)	4.96 (2.04 to 12.08)	<0.01	4.99 (1.72 to 14.47)	<0.01
Multiple vs single centre (ref)	3.02 (0.99 to 9.17)	0.05	4.49 (1.39 to 14.53)	0.01
Funding vs unclear (ref)	1.79 (0.80 to 3.99)	0.15	0.98 (0.36 to 2.62)	0.97
Sample size>76 vs ≤76 (ref)	0.53 (0.24 to 1.15)	0.11	0.55 (0.24 to 1.26)	0.16
Blinding of participants and personnel				
English vs Chinese (ref)	203.47 (24.93 to 1660.49)	<0.01	230.74 (19.93 to 2671.65)	<0.01
Multiple vs single centre (ref)	0.91 (0.11 to 7.35)	0.93	3.81 (0.59 to 24.46)	0.16
Funding vs unclear (ref)	3.79 (1.31 to 10.95)	0.01	0.90 (0.22 to 3.74)	0.89
Sample size>76 vs ≤76 (ref)	0.35 (0.10 to 1.15)	0.08	0.80 (0.16 to 3.89)	0.78
Blinding of outcome assessors				
English vs Chinese (ref)	50.82 (13.03 to 198.18)	<0.01	199.82 (22.97 to 1738.10)	<0.01
Multiple vs single centre (ref)	3.63 (0.92 to 14.32)	0.07	65.47 (6.03 to 710.78)	<0.01
Funding vs unclear (ref)	2.11 (0.72 to 6.19)	0.18	0.29 (0.06 to 1.49)	0.14
Sample size>76 vs ≤76 (ref)	0.24 (0.07 to 0.87)	0.03	0.24 (0.04 to 1.40)	0.11

*From separate logistic regression models adjusting for other factors.
ref, reference category.

and frequency and duration of treatment sessions). The experience of practitioners would be another factor for consideration, and when possible, for standardisation. Ideally, the qualification, training and experience of the acupuncturists providing the treatment should be more or less similar, especially for an exploratory trial.

In parallel, trial investigators should judge whether a blinding technique can be used according to the actual setting because in many situations of acupuncture trials, it is impractical or infeasible to blind the participants or study personnel. Nevertheless, blinding should be encouraged in an explanatory trial. They may adopt appropriate blinding method, such as sham acupuncture, and monitor the performance of blinding regularly to ensure that it is well implemented.

In specifying interventions, trial investigators should also account for the impact of co-interventions, and justify their use during the course of a trial. They must make an effort to minimise cross-overs, so as to maintain interventions as planned, and develop strategies to maximise patient follow-up. In all cases, these issues should be documented in the study protocol.

A statistical analysis plan should be in place prior to undertaking data analysis. Trial investigators should prespecify the analytic strategies to deal with loss to follow-up and the missing data problem. These should be explicit for both primary and secondary outcomes. It would be highly desirable to prespecify any subgroup hypotheses, if relevant, and plan the corresponding adjusted analyses.

There are other issues that trial investigators should take into consideration prior to their studies, such as safety monitoring and specification of interim analysis and stopping rules.

Strengths and limitations

The main strength of this study was its comprehensive and systematic investigation of the methodological details about the design, conduct and analysis of acupuncture RCTs for KOA. Nonetheless, several limitations should be considered. The data extraction form was developed by our research team. Despite the use of trained assessors and pilot testing of the questionnaire, it is likely that some extracted items from the original article source might still incur a certain degree of misinterpretation which could have affected the results. The insufficient reporting of methodological details among the included trials limited our assessment of the methodological quality and identification of specific methodological gaps. In addition, this study did not examine the specification of patient population in the included trials, and the findings may not be applicable to other

acupuncture trials. Lastly, this study did not investigate methodological characteristics of trials published in languages other than Chinese and English, which were expected to be scarce in the literature.

CONCLUSION

The overall risk of bias was high among published RCTs testing acupuncture for KOA. When examining methodological details, important issues were present in every aspect of design, conduct and analyses. These included, but were not limited to, randomisation, standardisation and operation of procedures, blinding, outcome specification and assessment, sample size calculation, setting up early stopping rules and data analysis (such as dealing with loss to follow-up, subgroup and adjusted analyses). Identifying and understanding these important methodological gaps have profound implications for future trials. Our findings support the development of an evidence-based methodological guidance for the design, conduct and analysis of future trials testing acupuncture for KOA.

Contributors XS conceived the study and acquired the funding. PLJ, LT, JJY, DYK and XS developed and tested the questionnaire. PLJ, XZ, JJY and JLL acquired the data and performed risk of bias assessment. PLJ, LT and YNL performed data analysis. PLJ drafted the manuscript. XS, LT, AHL and JJY critically revised the manuscript. All authors read and approved the final manuscript. XS is the guarantor.

Funding This project was supported by the National Natural Science Foundation of China grant No. 81590955.

Disclaimer The funding has no role on any processes of the study (ie, study design, statistical analysis and results reporting).

Competing interests None declared.

Patient consent Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement All relevant data are within the paper and its Supporting Information files.

REFERENCES

1. Bastick AN, Belo JN, Runhaar J, *et al*. What are the prognostic factors for radiographic progression of knee osteoarthritis? A meta-analysis. *Clin Orthop Relat Res* 2015;473:2969–89.
2. Foster NE, Thomas E, Barlas P, *et al*. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. *BMJ* 2007;335:436.
3. Soni A, Joshi A, Mudge N, *et al*. Supervised exercise plus acupuncture for moderate to severe knee osteoarthritis: a small randomised controlled trial. *Acupunct Med* 2012;30:176–81.
4. McGettigan P, Henry D. Cardiovascular risk and inhibition of cyclooxygenase: a systematic review of the observational studies of selective and nonselective inhibitors of cyclooxygenase 2. *JAMA* 2006;296:1633–44.
5. Blower AL, Brooks A, Fenn GC, *et al*. Emergency admissions for upper gastrointestinal disease and their relation to NSAID use. *Aliment Pharmacol Ther* 1997;11:283–91.
6. Bannuru RR, Natov NS, Obadan IE, *et al*. Therapeutic trajectory of hyaluronic acid versus corticosteroids in the treatment of knee osteoarthritis: a systematic review and meta-analysis. *Arthritis Rheum* 2009;61:1704–11.
7. Chen R, Chen M, Kang M, *et al*. The design and protocol of heat-sensitive moxibustion for knee osteoarthritis: a multicenter randomized controlled trial on the rules of selecting moxibustion location. *BMC Complement Altern Med* 2010;10:32.
8. 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;23:10–17.
9. Paramore LC. Use of alternative therapies: estimates from the 1994 robert wood johnson foundation national access to care survey. *J Pain Symptom Manage* 1997;13:83–9.
10. Ezzo J, Hadhazy V, Birch S, *et al*. Acupuncture for osteoarthritis of the knee: a systematic review. *Arthritis Rheum* 2001;44:819–25.
11. White AR, Filshie J, Cummings TM; International Acupuncture Research Forum. Clinical trials of acupuncture: consensus recommendations for optimal treatment, sham controls and blinding. *Complement Ther Med* 2001;9:237–45.
12. Derry C, Derry S, McQuay H, *et al*. Systematic review of systematic reviews of acupuncture published 1996-2005. *Clin Med* 2006;6:381–6.
13. Jevsevar DS, Brown GA, Jones DL, *et al*. The American academy of orthopaedic surgeons evidence-based guideline on: treatment of osteoarthritis of the knee, 2nd edition. *J Bone Joint Surg Am* 2013;95:1885–6.
14. McAlindon TE, Bannuru RR, Sullivan MC, *et al*. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartilage* 2014;22:363–88.
15. National Clinical Guideline Centre. Osteoarthritis care and management in adults: clinical guideline CG177, methods, evidence and recommendations. https://en.wikipedia.org/wiki/National_Clinical_Guideline_Centre (accessed 29 Nov 2017).
16. Lu W, Dean-Clower E, Doherty-Gilman A, *et al*. The value of acupuncture in cancer care. *Hematol Oncol Clin North Am* 2008;22:631–48.
17. Ahn AC, Bennani T, Freeman R, *et al*. Two styles of acupuncture for treating painful diabetic neuropathy--a pilot randomised control trial. *Acupunct Med* 2007;25:11–17.
18. Kim KH, Kang JW, Lee MS, *et al*. Assessment of the quality of reporting in randomised controlled trials of acupuncture in the Korean literature using the CONSORT statement and STRICTA guidelines. *BMJ Open* 2014;4:e005068.
19. Purepong N, Jitvimonrat A, Sitthipornvorakul E, *et al*. External validity in randomised controlled trials of acupuncture for osteoarthritis knee pain. *Acupunct Med* 2012;30:187–94.
20. Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in pubmed journals. *Lancet* 2005;365:1159–62.
21. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987;317:426–32.
22. Dwan K, Gamble C, Williamson PR, *et al*. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One* 2013;8:e66844.