How to design high quality acupuncture trials—a consensus informed by evidence

Yu-Qing Zhang,[1,2,3,4] Rui-Min Jiao,[1,5] Claudia M Witt,[6] Lixing Lao,[7] Jian-Ping Liu,[8,9] Lehana Thabane,[10] Karen J Sherman,[11,12] Mike Cummings,[13] Dawn P Richards,[14] Eun-Kyung Anna Kim,[15] Tae-Hun Kim,[16,17,18] Myeong Soo Lee,[19,20] Michael E Wechsler,[21] Benno Brinkhaus,[22] Jun J Mao,[23] Caroline A Smith,[24] Wei-Juan Gang,[1,5] Bao-Yan Liu,[25] Zhi-Shun Liu,[26] Yan Liu,[27] Hui Zheng,[28] Jia-Ni Wu,[26] Alonso Carrasco-Labra,[29] Mohit Bhandari,[10] Philip J Devereaux,[10] Xiang-Hong Jing,[1,5] Gordon Guyatt[3,10]

For numbered affiliations see end of the article

Correspondence to: X-H Jing xhjingt66@163.com (ORCID 0000-0003-0272-1576)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2022;376:e067476 http://dx.doi.org/10.1136/bmj-2021-067476

Accepted: 23 February 2022

An international panel including patients, clinicians, researchers, acupuncture and surgery trialists, statisticians, and experts in clinical epidemiology and methodology have developed new guidance for randomised controlled trials in acupuncture. It addresses the most prevalent and critical concerns of current acupuncture trials and will help funding agencies, trial registers, and journal editors to evaluate the relevance, importance, and quality of submitted trial proposals and completed trials

Acupuncture therapies are used in 183 countries according to a 2013 survey by the World Federation of Acupuncture-Moxibustion Societies.[1] Despite over 14 000 published randomised controlled trials in acupuncture,[2 3] scepticism remains about the number of studies to support the effectiveness and safety of interventions, as is the case for other complementary therapies. Challenges in designing and conducting high quality acupuncture trials may partly explain the scepticism. Randomised trials provide the most rigorous support for therapeutic and health system decisions for patients, clinicians, payers, and policy makers.[4 5] Multiple meta-epidemiological studies have identified shortcomings in acupuncture trials. Common problems include limitations in study design, applicability, and reporting.[6-9]

In contrast to pharmacological interventions, acupuncture therapies rely on procedural expertise and are complex and multifaceted.[8 9] These characteristics present unique methodological challenges for randomised trials, including standardisation of interventions and controls, dealing with non-specific effects of interventions, and patients' expectations.

The original and revised Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA)[10 11] focus on reporting not optimal trial conduct. Other acupuncture guidance covers type of trial (efficacy,[12] effectiveness[13]), design, and lack of representativeness (include an international panel[14]), but is not informed by comprehensive and systematic evidence synthesis.

To address the most prevalent design and methodological concerns in current acupuncture randomised trials, an international panel including patients, clinicians, researchers, and trialists in acupuncture, surgery, statistics, patient engagement, and clinical epidemiology developed guidance for research teams planning acupuncture trials. To inform the guidance, we conducted a systematic survey[15] exploring characteristics associated with acupuncture treatment effects.

The issues covered in this guidance fall into two categories: those unique to acupuncture trials, and, based on our exhaustive review of acupuncture trials, those that are not specific but still applicable to acupuncture trials. We frame our discussion as a series of questions for readers to consider when designing trials, focusing on choices of the research question, patients, interventions, and outcomes and presenting the implications of these choices using examples from existing trials (box 1).

SUMMARY POINTS

Despite over 14 000 published randomised controlled trials in acupuncture, shortcomings persist, including research corresponding poorly to clinical practice, poor reporting, and high risk of bias

In contrast to pharmacological interventions, acupuncture therapies rely on procedural expertise and are often complex and multifaceted

An international panel including patients, clinicians, researchers, trialists in acupuncture and surgery, and experts in statistics, patient engagement, and clinical epidemiology and methodology developed guidance for trials

This guidance covers methodological issues, including selection of control, blinding, and the challenges particular to acupuncture trials, including four essential design elements for acupuncture based on whether trials are primarily explanatory/mechanistic or pragmatic/practical

Development of the guidance

We established a steering committee including frontline acupuncture clinicians, acupuncture trialists, clinical trial methodologists, and a statistician (XHJ, JPL, LXL, CMW, LT, YQZ, RMJ, and GHG) with extensive experience in acupuncture and trial methodology. The steering committee recruited an international expert panel of 27 experts (from Asia, Europe, America, and Australia), including patients, frontline acupuncture

Box 1: Guidance to consider when designing an acupuncture trial

Choosing the research question

Consideration 1: Is the question important?

Trialists should establish the rationale for the research question informed by a systematic review of the relevant literature.

Enrolling patients

Consideration 2: Who should the participants be?

Trialists with a primary explanatory or mechanistic objective could include people who are most responsive to the intervention. With a primary pragmatic or practical objective, trialists could include a broad population varying in age, severity, comorbidity, and exposure to other interventions.

Consideration 3: How can trialists address possible differences in effect across patient groups?

For pragmatically oriented trials include heterogeneous populations, with prior specification of subgroup analyses to consider hypothesised effect modification.

Selecting the intervention

Consideration 4: Who should perform the intervention?

Trialists should report the expertise of the acupuncturists. A trial that aims to show whether an acupuncture treatment can work under ideal conditions will choose the most expert practitioners available. Trialists aiming to establish the effect of treatment in ordinary practice will select clinicians with typical levels of expertise.

Consideration 5: What specific acupuncture technical features should be considered if aiming to design trials for maximum treatment effect?

If aiming for maximum treatment effect, trialists should select a high frequency of acupuncture treatment sessions and penetrating type of acupuncture (manual and electroacupuncture) over lower frequency or non-penetrating (transcutaneous electrical acupoint stimulation (TEAS), laser, acupressure) acupuncture.

Choosing the comparator

Consideration 6: Are trialists interested in the specific effect or the overall (specific and non-specific) effect of acupuncture?

Trialists should blind data collectors, outcome assessors, and data analysts and carefully consider the desirability of a sham that leads to underestimation of acupuncture's treatment effects in clinical practice.

Fine tuning the flexibility of intervention and comparator

Consideration 7: To what extent should the trialist choose a flexible intervention and comparator?

Trialists with a primary explanatory or mechanistic focus might specify that practitioners administering both the intervention and the control use a highly standardised approach. Trialists with a primary pragmatic or practical focus might include clinicians with varied techniques reflecting practice in the community and instruct them to use their usual treatment approaches.

Consideration 8: How should trialists deal with adherence?

When designing a trial that primarily takes the individual patient's perspective, trialists might consider implementing strategies to achieve optimal adherence. When trialists take a public health perspective, there is no need to implement strategies to increase adherence.

Selecting the outcome measure and follow-up

Consideration 9: What type of outcomes should trialists choose?

Trialists with mainly mechanistic objectives can focus on surrogate outcomes (eg, blood pressure, heart rate). If trialists have a primary pragmatic objective, they should focus on patient important outcomes such as pain. Trialists should continue using symptoms and functional outcomes and increase the use of quality of life and major events (eg, cardiovascular events, hospital admission, infections) as outcomes.

Consideration 10: How should trialists decide the duration of follow-up?

When investigators have primarily explanatory objectives, short follow-up times are likely to be optimal. Acupuncture trials with primarily pragmatic objectives might choose follow-up times of six months or longer.

Engaging patients

Consideration 11: How should triaists engage patients during design and conduct of trials?

To empower patients and conduct more patient centric research, acupuncture trials should consider informing, educating, discussing, engaging, and partnering with patient organisations and with individual patients.

clinicians, acupuncture and surgical trialists, clinical trial methodologists, and statisticians. Acupuncture trialists and methodologists were appointed to the panel on the basis of their h index and assessment of their research expertise, and frontline clinicians on the basis of their clinical experience and reputation.

Using the results of our systematic survey,[15] the steering committee drafted the first version of the guidance. The purpose of the survey was to identify the most prevalent trial design and execution issues and factors associated with the magnitude of acupuncture treatment effects.[15] In an iterative process, the expert panel provided feedback on the structure and content of the acupuncture trials guidance draft and on piloting of the guidance and user feedback (web appendix 1).

Box 2: Explanatory and mechanistic frameworks for trials

Explanatory-pragmatic

Schwartz and Lellouch's explanatory-pragmatic framework classified trials as either aiming to achieve the maximum effects of interventions (eg, measure tumour size or cancer related mortality as outcomes) under ideal conditions or to inform decisions in clinical or health system contexts by resembling current practice settings.[4] Witt and colleagues have provided guidance to apply this framework specific to acupuncture trials.[13 16]

Mechanistic-practical

Karanicolas and colleagues provide an alternative framework emphasising the purpose and suggest that trial purpose depends on the decision makers' perspectives (eg, patients and clinicians versus policy makers).[5] They proposed a revised mechanistic-practical framework to design and interpret randomised trials. Mechanistic trials have proof-of-concept objectives and often include surrogate outcomes such as biological measures (eg, forced exhaled volume in 1 second (FEV_1) in patients with asthma[17]); practical trials provide comprehensive information that bears directly on healthcare decisions.

Use of the guidance

In the early stages of designing a randomised controlled trial of acupuncture, researchers should specify their objective. Studies may fall along a continuum from explanatory or mechanistic to pragmatic or practical (box 2).[4 5] Trials with primary explanatory or mechanistic objectives will evaluate the specific effect of acupuncture whereas those with mainly pragmatic or practical objectives will aim to inform clinical or policy decision making. Trialists can refer to the pragmatic-explanatory continuum indicator summary (PRECIS)[18] or PRECIS-2[19] to assess where their trials are positioned.

Once trialists situate their study in the explanatory-pragmatic continuum, they can use the PICOT framework[20] to determine critical design elements. For instance, planning a randomised trial to study the effect of acupuncture on chronic low back pain requires selection of participants (P), intervention (I) (eg, acupoints, stimulation methods, depth of the insertion), comparator (C) (eg, sham acupuncture or medication), outcomes (O), and duration of follow-up (T). If their primary objective is to evaluate the specific effect of acupuncture, trialists will likely choose sham as the comparator. If the primary objective is to explore the effectiveness and safety of acupuncture therapies and inform clinical or policy decision making, commonly used alternatives (such as non-steroidal anti-inflammatory drugs) might be suitable comparators.

In defining their objectives and consequent methods in relation to the explanatory-pragmatic continuum (box 3), trialists should consider a systematic review to establish the rationale for their research question (box 1, consideration 1).

Although the guidance we offer to determine the objective of pragmatic versus explanatory trials generally applies, there will be exceptions. For instance, although pragmatic objectives can generally be achieved by enrolling heterogeneous populations, if the desired focus is a narrower population, it does not preclude a primarily pragmatic objective. Once the objectives are defined trialists should consider the 11 questions below in designing their trial.

Box 3: Acupuncture randomised trial with both pragmatic and explanatory features[21]

- *Population*—Patients with chronic low back pain
- *Intervention*—Clinicians gave true acupuncture treatment twice a week for five weeks (10 sessions). If patients experienced a 10-50% reduction in pain (Von Korff chronic pain grade scale), they received five additional sessions
- *Comparators*—Sham acupuncture (same frequency and session as true acupuncture) or German guideline based multimodal treatment programme (10 sessions of physician or physiotherapist administered physiotherapy or exercise)
- *Outcomes and follow-up*—The primary outcome was pain or functional status of low back pain six months after randomisation. Secondary outcomes were quality of life and patient global assessment of therapy effectiveness, medication use, and adverse events at each session after six months. Trialists defined pain status as 33% improvement on three pain related items on the Von Korff chronic pain grade scale and functional status as 12% improvement on the Hanover functional ability questionnaire

Interpretation

- Use of guideline based conventional therapy and patient important outcomes with long term follow-up indicates a pragmatic design. Having sham acupuncture as the comparator shows a more explanatory feature

Choosing the research question

Consideration 1: Is the question important?

Published acupuncture randomised trials rarely cite previous systematic reviews in describing the rationale for conducting a trial, suggesting a high risk of wasteful use of research resources[22] (web appendix 2, 2.1.1). Deciding on a trial's objective requires an understanding of the available body of evidence on related questions, which requires the availability or conduct of systematic reviews.[22] Are there already sufficient trials exploring the specific effect of acupuncture therapies? Is there already moderate or high certainty evidence establishing treatment effects in particular settings to inform decision making? If existing systematic reviews show little or no evidence on acupuncture therapies for a target condition, investigators can conclude that an optimal randomised trial will provide proof-of-concept information. If sufficient evidence establishes a therapeutic effect on a surrogate outcome or patient important outcome under narrowly defined circumstances, moving to a pragmatic trial is likely to be appropriate (box 1, consideration 1).

Enrolling patients

Consideration 2: Who should the participants include?

Trialists need to determine the population that might be the most responsive to and are interested in the intervention and what are the necessary enrolment restrictions (box 1, consideration 2). Restricting participants to those who are most responsive reflects a proof-of-concept approach. Such a trial will aim to establish the effect of acupuncture therapies in people

Box 4: Three acupuncture trials with inclusion criteria that reflect explanatory and pragmatic goals

Example 1: Patients with knee osteoarthritis and chronic knee pain[23]

- *Inclusion criteria (explanatory or mechanistic)*—Patients were eligible if they fulfilled all five criteria: age ≥50 years; history of knee pain for >3 months; reported knee pain on most days of the past month; average knee pain severity ≥4 on an 11 point scale over the past month; and) reported morning knee stiffness of <30 minutes
- *Exclusion criteria*—The study excluded patients if they fulfilled any of the following 13 criteria: history of any systemic arthritic condition; knee arthroplasty on most painful knee; on waiting list for any knee surgery; any knee surgery in past six months; presence of any other condition affecting lower limb function (eg, trauma, malignancy, neurological condition); any knee injection in past six months (eg, cortisone, hyaluronic acid);) current use of oral or injectable anticoagulant medication; use of acupuncture in past 12 months; any bleeding disorder; allergy to light; referral to a pain clinic or use of morphine or pethidine within the past six months; any other medical condition precluding participation in the trial (eg, kidney or liver disease, deep vein thrombosis); knee pain subject to compensation claim; unable to give written informed consent

Example 2: Patients with chronic knee pain and knee osteoarthritis[24]

- *Inclusion criteria (pragmatic or practical)*—Patients were eligible if they fulfilled all four criteria: diagnosis of knee osteoarthritis by American College of Rheumatology criteria; age ≥45 years; pain in one or both knees for three months or longer; confirmed knee osteoarthritis with radiological evidence (at least grade 1 according to the Ahlbäck classification)

Example 3: Patients with persistent non-specific low back pain using traditional acupuncture or usual care[25]

- *Inclusion criteria (pragmatic or practical)*—Any patients aged 18-65 with non-specific low back pain for 4-52 weeks. Patients with spinal disease were excluded

most likely to benefit, leaving open the question of its value in routine clinical settings serving a diverse population.

Trialists with a primarily explanatory or mechanistic objective might include people who are most responsive, those at highest risk of adverse outcomes, or the most highly compliant and may exclude those with comorbidities that might affect the outcomes or likely to drop out because of severe comorbidity.

Trials with a primarily pragmatic aim would produce widely generalisable results that optimally inform clinical practice. The trial should include a broad population varying in age, severity, comorbidity, and exposure to other interventions and, depending on their target audience, either highly compliant or more general populations (box 4).

Consideration 3: How can trialists address possible differences in effect across patient groups?

A limitation of pragmatic trials is that they include very heterogeneous and rarified populations and run the risk of producing an overall estimate of effect that is not applicable to all enrolled patients. In the extreme, they report an intermediate effect when some patients experience large benefit and some none at all. Trialists can overcome this limitation by prespecifying subgroup analyses with statistical tests of interaction to investigate hypothesised effect modification (box 1, consideration 3).

Participants' expectation serves as an example of subgroup effect in the context of acupuncture trials. Trialists hypothesised that acupuncture would show larger effects in patients who expect a benefit from acupuncture than in those who do not.[26] Across therapeutic areas and health outcomes (eg, pain relief, symptom improvement, return to work), results have consistently confirmed this hypothesis: participants who have positive expectations of interventions consistently show a larger effect of acupuncture than the more pessimistic.[27-29]

The repeated finding that patient expectations influence magnitude of effect suggests trialists should assess participants' expectations at the time of randomisation. Trialists can use validated instruments such as the acupuncture expectancy scale,[30] expectation for treatment scale,[31] and the expectations for complementary and alternative medicine treatments questionnaire.[32]

Published acupuncture trials have rarely explored effect modification (web appendix 2, 2.1.2). Trialists can consider exploring effect modifiers when designing acupuncture trials with sufficiently large sample sizes and, if they identify possible subgroup effects, assess their credibility criteria using the rigorously constructed and user tested instrument to assess the credibility of effect modification analyses (ICEMAN).[33] Choosing a small number of hypotheses, ideally identified as promising (ie, with moderate credibility on ICEMAN) in previous trials, represents an optimal approach (box 5).[33]

Selecting the intervention

Consideration 4: Who should perform the intervention?

The effect of acupuncture may differ depending on various characteristics of the acupuncturists (box 1, consideration 4). These include education and training (whether they had systematic training such as undergraduate, graduate, diploma training, or short term training), length of practice, and how well the acupuncturists perform a particular complex technique (eg, fire needles and intradermal needle).

Differential expertise can create a systematic difference[35-37] across interventions and between trials.[38] A bibliometric study investigated 7085 Chinese randomised trials of acupuncture and found that over 20% of the trials compared one acupuncture treatment with another or with an alternative treatment that required expertise (box 6).[3]

In a patient level analysis with 9990 patients treated by 2781 physicians in Germany, neither training

Box 5: Patients' expectation as an effect modifier and credibility of effect modification assessment

Study of acupuncture versus sham acupuncture for patients with chronic low back pain.[34] The study used the expectations for relief scale to measure patients' expectations at baseline and conducted a subgroup analysis in patients with high versus low expectations. Expectations were a significant predictor of only true acupuncture response (P=0.002), such that those with greater expectations had more substantial pain relief.

We used ICEMAN[33] to assess the credibility of potentially relevant effect modification:

Was the direction of the effect modification correctly hypothesised a priori? (definitely no, probably no or unclear, probably yes, or definitely yes)

Definitely yes: subgroups (patients' expectation) were prespecified in published protocol, and publications included an explicit statement that they had correctly predicted the direction of effect modification

Was the effect modification supported by prior evidence? (inconsistent with prior evidence, little or no support or unclear, some support, strong support)

Strong support: the main result was consistent with prior solid evidence directly applicable to the clinical scenario

Does a test for interaction suggest that chance is an unlikely explanation of the apparent effect modification (consider irrespective of the number of effect modifiers)? (chance a very likely explanation, chance a likely explanation or unclear, chance may not explain, chance an unlikely explanation)

Chance may not explain: P=0.002 (interaction P values ≤0.01 and >0.005)

Did the authors test only a small number of effect modifiers or consider the number in their statistical analysis? (definitely no, probably no or unclear, probably yes, definitely yes)

Probably yes: three effect modifiers were prespecificed but applied to other outcomes than specified in the protocol

If the effect modifier is a continuous variable, were arbitrary cut-points avoided? (definitely no, probably no or unclear, probably yes, definitely yes)

Definitely yes: The analysis was based on a full continuum—eg, assuming linear or logarithmic relationships. A repeated-measures analysis of variance (ANOVA) was used to characterise the relationship between psychiatric group and analgesia

Optional: Are there any additional consideration that may increase or decrease credibility? (yes, probably decrease credibility; yes, probably increase credibility)

Yes, probably increase credibility: the effect modification persisted after adjustment for other potential effect modifiers (interventions and psychopathology)

How would you rate the overall credibility of the proposed effect modification (very low—very likely no effect modification, use overall effect for each subgroup; low—likely no effect modification, use overall effect for each subgroup but note remaining uncertainly; moderate—likely effect modification, use separate effects for each subgroup but note remaining uncertainly; high—very likely effect modification, use separate effects for each subgroup)?

High credibility very likely: none of the response options definitely or probably reduced credibility

Interpretation

We judged the credibility of the potential effect modification as high considering a small number of clearly a priori hypotheses with a specified direction; support from previous evidence; a low P value in the test of interaction; and adjustment for other potential subgroup effects

duration nor experience affected the extent of the acupuncture effect.[40] Unfortunately, exploration of the effect of expertise is limited because acupuncture trials often fail to report this information (web appendix 2, 2.1.3 and 2.2.1).

Box 6: Who should perform the intervention?

Consider a randomised controlled trial in which 80 participants with post-stroke pseudobulbar palsy were randomised to an acupuncture intervention—a quick needle insertion at the Aqiang point—or a comparator of routine acupuncture in combination with western medicine.[39] Clinicians implementing the intervention needed to perform a quick needle insertion with strong stimulation at the Aqiang point (located in the anterior midline of the neck, between the thyroid cartilage and cricoid cartilage) that may induce the patient to cough violently. In the comparator group, clinicians needed only to stimulate CV 23 (Lian quan) point with the standard acupuncture manipulation. If practitioners in the trial had a greater skill or more experience in performing the special acupuncture techniques, the results might not apply to those with less expertise in the Aqiang point approach.

This trial failed to find a difference in swallow function between the two groups, but it did not establish the expertise of those administering the more challenging intervention. The effects of expert administration therefore remain uncertain, although the results may be representative of suboptimal expertise in the community. The trial's small sample size might also explain the failure to detect a difference.

For studies investigating pain as an outcome, greater experience and training of practitioners (systematic acupuncture or complementary medicine education versus short term training) is associated with greater treatment effect.[15]

Whether a trial has explanatory or pragmatic objectives affects trialists' decisions about the extent of expertise. A trial that primarily aims to show whether an acupuncture treatment can work under ideal conditions will choose the most expert practitioners. On the other hand, trialists primarily aiming to establish the effect of treatment in everyday practice will select clinicians with typical levels of expertise (box 7).

One way to deal with the issue of expertise is by using a trial that randomises by expertise. For example, investigators interested in the relative effect of Japanese meridian therapy versus Korean constitutional acupuncture might randomise patients to clinicians who specialised in meridian therapy or those who specialised in Korean constitutional acupuncture. Such expertise based designs may enhance applicability, encourage practitioners' participation in trials,[43] result in less crossover,[35] and provide patients the best quality care. To date,[44] no acupuncture trial has used

Box 7: Examples of acupuncture trials using highly skilled and variously skilled practitioners

Auricular acupuncture versus sham auricular acupuncture for musculoskeletal disorders related chronic back pain[41]

In this explanatory mechanistic trial all participating acupuncturists had at least 10 years of experience. Having highly skilled acupuncturists perform the intervention may enhance treatment effects but raises questions for a patient consulting a practitioner who has just completed acupuncture training

Acupuncture plus advice and exercise versus advice and exercise for patients with knee osteoarthritis[42]

In this pragmatic trial, acupuncture was delivered by 67 physiotherapists trained by the Acupuncture Association of Chartered Physiotherapists (35 hours of training), with three to over 10 years of practice. A positive result of such a trial would suggest a widely applicable intervention; a negative result would leave the question of the effect with more expert practitioners uncertain

an expertise based design, but it is an attractive option for future trials of alternative acupuncture approaches that take a practical perspective.[45 46]

In any case, trialists should report the expertise of the acupuncturists who administered intervention and comparator treatments. More transparent reporting will enable further investigation into how expertise affects the effectiveness of acupuncture treatment.

Consideration 5: What specific acupuncture technical features should be considered if aiming to design trials for the maximum treatment effect?

The effect of acupuncture may differ depending on the choice of techniques, including underlying acupuncture theory (eg, western medical acupuncture, traditional Chinese medicine acupuncture, dry needling) and type of acupuncture (manual versus electroacupuncture).

The multivariate analysis in our systematic survey[15] found that penetrating types of acupuncture (manual and electroacupuncture) showed a larger treatment effect than non-penetrating (transcutaneous electrical acupoint stimulation, laser, acupressure) acupuncture. In addition, a higher frequency of acupuncture treatment sessions was associated with larger effects than a lower frequency. Trialists can establish an expert committee to define the optimal strategy for acupuncture related details (eg, acupoints selection, stimulation technique) (box 1, consideration 5).

Choosing the comparator

Consideration 6: Are trialists interested in the specific effect or overall effect of acupuncture?

Blinding refers to those involved in randomised trials being unaware of treatment assignment after randomisation. Blinding is relevant to participants, clinicians, data collectors, outcome assessors, and data analysts. Blinding of clinicians and patients avoids non-specific effects and bias arising from co-intervention; blinding of others minimises bias in the assessment of outcome. Blinding clinicians and patients raises both theoretical and practical challenges, but to minimise bias trialists should blind data collectors, outcome assessors, and data analysts (box 1, consideration 6, web appendix 2, 2.2.2(1)).

Although legitimate scepticism exists about tests of blinding, in seven of 10 studies we identified in which authors tested for blinding (one to four times during the trial), study participants were unable to distinguish active and sham treatment, providing potentially useful reassurance of the success of blinding.[15] Thus, the choice of whether to conduct tests for blinding remains, in our view, a matter of investigator judgment in which both options have merit.

In drug trials, which are closer to a explanatory or mechanistic design, trialists often use a placebo to achieve blinding. Acupuncture trials, however, face multiple challenges when using placebo or sham acupuncture for blinding.[15] It is nearly impossible to blind clinicians who deliver acupuncture, although one device exists that allows practitioner blinding but limits the acupuncture to a rather superficial version (maximum 5 mm insertion).[45] Participants who have had previous acupuncture experience may also be hard to blind.[47 48]

Whether to attempt blinding in acupuncture trials depends on the hypotheses and objectives of a trial. Here, we will refer to a treatment's biological effects as specific effects and placebo effects as acupuncture's non-specific effects. Trialists examining explanatory questions should include sham acupuncture control with adequately blinded participants to differentiate acupuncture's specific and non-specific effects.[45]

The results of such explanatory studies require careful interpretation, however. The clinical and basic science literature support the possibility of specific effects generated by sham acupuncture[49 50]— that is, a failure to show a difference between real and sham acupuncture may be because the sham has effects closely related to that of the intervention. Therefore, when using sham control, trialists must consider the possibility of a specific effect generated by the sham and thus underestimating the effects of acupuncture in clinical practice compared with no intervention or other interventions such as drugs.

Previous studies of shams have focused on pain and chronic pain and conducted univariable analyses (web appendix 2, 2.2.2 (2)). We used multivariable analyses in our systematic survey, adjusting for other potential factors such as treatment frequency, flexibility of the acupuncture regimen, and sample size. Like a previous systematic review,[51] we found the type of sham did not influence acupuncture's effect.[15] This is perhaps unsurprising considering the finding of meta-epidemiological studies in areas beyond acupuncture: results of such studies have proved inconsistent, with one recent thorough review showing no systematic effect of blinding.[52]

Considering all the evidence, the effect of the type of sham remains uncertain and might vary with the type of medical condition. Over 90% of acupuncture trials focus on pain, quality of life, function, and other symptoms. These are all subjective outcomes in which blinding may be particularly important.[15] Trialists

Box 8: Choosing the comparator: examples of acupuncture trials investigating specific and overall (specific and non-specific) effects

Specific and overall effects

Trial 1—Patients with knee osteoarthritis were assigned to true acupuncture plus education and attentional therapy, sham acupuncture plus education and attentional therapy, or education and attentional therapy alone. Sham acupuncture (combination of non-insertion and needle-insertion sham acupuncture) was received with two needles at sham points and tapping at nine real points.[53]

Specific effects

Trial 2—Patients with knee osteoarthritis were assigned to drug treatment alone, drug treatment plus true acupuncture, or drug treatment plus sham acupuncture.[24] Sham acupuncture was administered with the same duration and frequency and by the same specialist who performed the true acupuncture. Retractable needles were placed into small adhesive cylinders to support the needles but did not perforate the skin. The acupuncturist placed the needles at the same points as the non-sham group and used the same pairs of electrodes to simulate the electrical connection.

Trial 3—Patients with knee osteoarthritis were randomised to intensive acupuncture or sham acupuncture.[54] Sham acupuncture (combined non-insertion and needle insertion sham acupuncture) was achieved by choosing acupoints that were away from the conventional acupoints or meridians, superficially penetrated (2-3 mm in depth), and without manipulation of the needles after insertion for "de qi" responses.

Overall (specific and non-specific) effect

Trial 4—Patients with persistent non-specific low back pain received either a short course of traditional acupuncture or usual care only. Patients in the usual care group received NHS treatment according to their general practitioner's assessment of need.[25]

should therefore carefully consider the desirability of a sham in randomised trials investigating efficacy and, if desirable, the nature of the sham in relation to their trial's specific objectives. Trialists should also be aware that the use of non-penetrating needle shams mandates the use of the same device in the real acupuncture group. These devices may impede real acupuncture treatment effects (box 1, consideration 6; box 8).[55]

Fine tuning the flexibility of intervention and comparator

Consideration 7: To what extent should the trialist choose a flexible intervention and comparator?

Acupuncture therapies are multifaceted interventions that incorporate patient-practitioner interactions, treatment theories, tools to deliver the stimulations, manipulation techniques, and points selection. The trial objectives should determine the flexibility of the intervention or control.

Suppose trialists aim to assess the maximum treatment effect of a particular acupuncture technique or regimen (closer to explanatory approach). In that case, they should specify that practitioners use a highly standardised approach to administer both the intervention and the control. In-person or video training for all participating practitioners regarding types of needles, the number of needles, needling details (eg, direction, depth of insertions, needle manipulation), needle retention time, and responses (eg, de qi or muscle twitches) may facilitate standardisation and may be particularly important in multicentre studies.

Box 9: Examples of fine tuning the flexibility of intervention and comparator in acupuncture trials

Restrictive and highly standardised intervention and comparison

Example 1: Postmenopausal women with early stage breast cancer taking an aromatase inhibitor randomised to acupuncture, sham acupuncture, or waiting list control[56]

- Acupuncture intervention—A group of acupuncturists developed a standard operating protocol based on previous acupuncture studies for aromatase inhibitor related arthralgias adhering to the STRICTA recommendation. Trialists described the frequency of treatment, acupoint prescription, whether "de-qi" occurred, depth of needle insertion, and acupoint manipulation in detail
- Comparison (sham acupuncture)—The sham acupuncture consisted of a standardised prescription of minimally invasive, shallow needle insertion using thin and short needles at the regimen of non-acupuncture points. The sham acupuncture protocol also included joint specific treatments and an auricular sham based on application of adhesives to non-acupuncture points on the ear

Varied intervention and comparator

Example 2: A pragmatic acupuncture randomised controlled for the management of chronic low back in older adults[57]

- Acupuncture intervention—An acupuncture advisory panel produced a consensus acupuncture intervention protocol. Acupuncture point sites must include both local and distal points with 6–20 insertions per session, but selection of specific acupoints was at the practitioner's discretion. The needle retention time and length f the session ranged between 0 and 40 minutes and 45-60 minutes, respectively

Example 3: Patients visiting emergency departments with an ankle sprain, migraine, or lower back pain were randomised to acupuncture alone or combined with pharmacotherapy[58]

Comparison (pharmacotherapy)—Pharmacotherapy can be any first or second line analgesia drug.

Comments

The example of the highly standardised intervention and comparator reflects an explanatory goal for the study. The studies with flexible intervention and comparator reflect a pragmatic goal, and inferences from these studies might be highly applicable to clinical practice.

Trialists can monitor standardisation in participating centres by visiting sites regularly or reviewing the recorded video reports. In addition, standardisation should include the number, frequency, and duration of treatment sessions (box 1, consideration 7; box 9, example 1).

Trialists with primarily pragmatic objectives will take a different approach. They will probably recruit a wide variety of clinicians representing the variability in technique in the community and instruct practitioners to use their usual treatment approaches (box 9, examples 2 and 3).

The multivariable analysis in our systematic survey found that single centred randomised trials of acupuncture have larger treatment effects than multicentre trials.[15] The greater standardisation in single centres might be responsible for the larger effects.

Reporting all technical details of the intervention and controls transparently according to the STRICTA recommendations would enhance the interpretation and applicability of acupuncture trials (web appendix 2, 2.1.4).[10 59]

Consideration 8: How should trialists deal with issues of adherence?

Although it is often inevitable that some trial patients do not attend all treatments, how this affects inferences from the evidence may vary depending on your perspective. When patients face the decision to use acupuncture therapies, they will be interested in the effectiveness and safety of the treatment. Thus, substantial non-adherence limits the applicability of the trial from the patients' point of view.

From a public health perspective, however, decision makers might be more interested in the treatment effect in a population with the suboptimal adherence that will actually occur in typical clinical practice. The non-adherence, if indeed typical of what happens in population implementation, will actually strengthen applicability from the public health point of view.

Thus, when designing a trial that primarily takes the individual patient's perspective, trialists should consider implementing strategies to achieve optimal adherence. One strategy is to engage patients to help design the trial. Patients can identify and improve onerous trial design, therefore increasing participation.[60] Another strategy is to include a run-in period before randomisation and only randomise patients who are highly compliant in the run-in.

Including a run-in period is highly feasible in drug trials when patients may be offered a placebo during the run-in or when the trial is focused on morbid or fatal events and short term exposure will have no influence on the outcome. The situation is more complicated in acupuncture trials focusing on symptoms in which initial exposure to treatment during a run-in may complicate inferences about what occurs after randomisation. This is probably why acupuncture trials rarely use run-in periods to enhance adherence.

As an alternative to run-ins, acupuncture trials can assign tasks to patients before randomisation to assess their adherence (web appendix 2, 2.2.3). Trialists can select patients based on completeness and timeliness of fulfilling these tasks. Trials could also provide additional support and adequate incentives to improve adherence (box 1, consideration 8; box 10).[45 62]

If trialists take a public health perspective when considering patients' adherence, there is no need to implement strategies to increase adherence. The minimal enforcement approach closely resembles standard practice and provides a reasonable estimation of how treatment effects might be in the setting in which trials are conducted.

Selecting outcome measure and follow-up

Consideration 9: What type of outcomes should trialists choose?

Trialists with mainly mechanistic objectives can usually focus on surrogate outcomes (eg, blood pressure, heart rate) to estimate the effect on patient important outcomes. If trialists' have a pragmatic objective primarily of guiding healthcare decision making, they should usually select patient important outcomes.[63] Trialists can consider taking advantage of minimal sets of agreed and standardised collection of outcomes in their clinical areas, known as core outcome set, on which investigators have agreed (https://www.comet-initiative.org/, web appendix 2, 2.2.4 (1)).[64-66] Including all outcomes in a core outcome set will ensure maximum applicability of trial results to patients, the public, healthcare professionals, and other decision makers. In addition, trialists can consider outcomes recommended by research organisations and medical societies' guidelines on conducting trials for specific diseases (web appendix 2, 2.2.4(2)). [64 65 67]

As well as using symptoms and functional outcomes trialists should increase the use of quality-of-life outcomes (web appendix 2, 2.1.5). Including major morbid events (eg, serious cardiovascular events, hospital admission, infections) might greatly enhance the importance and applicability of acupuncture trials when relevant but would likely require far larger sample sizes than is typical of current trials (web appendix 2, 2.2.4 (3)). Trials should also report adverse events outcomes (box 1, consideration 9; box 11).

Box 10: Example of acupuncture trial that implemented strategies to improve patient adherence

Acupuncture versus waiting list control for menopausal patients with vasomotor symptoms[61]

Patients in the acupuncture group could have up to 20 acupuncture treatments by one of four acupuncturists over six months. The acupuncturist and patients decided treatment dates together, respecting the treatment frequency to increase compliance. Patients were paid $35 (£26; €31) for their baseline visit and $25 to complete each follow-up visit. Those who completed all six visits were paid an additional $50.

Box 11: Selecting outcome measure and follow-up in acupuncture randomised trials

Randomised controlled trials on acupuncture versus sham acupuncture or no acupuncture for exercise induced asthma in children[17]

This trial measured forced expiratory flow in first second (FEV_1), forced vital capacity (FVC), and peak expiratory flow rate (PEFR) at 0, 5, 10, 15, 20 minutes after treatment. As the trial had an explanatory function trialists used biological or surrogate outcomes

Randomised trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain[68]

This trial used symptoms and dysfunction as the primary outcomes and disability, utilisation (provider visits, radiological procedures, operations, and hospital admissions), and cost as secondary outcomes. On a 0 to 10 scale, patients rated how "bothersome" their symptoms (back pain, leg pain, and numbness or tingling) were. Patients reported their dysfunction using the modified Roland disability scale. Blinded interviewers measured all outcomes at 4, 10, and 52 weeks after randomisation.

The aim of the study was to inform clinical or policy decision making. The follow-up selection informs the long term impact of acupuncture intervention. It also sheds light on the possible need for "booster treatments" or "maintenance treatment," therefore helping to achieve practical objectives.

Consideration 10: How to decide the duration of follow-up

Acupuncture effects often emerge immediately or shortly after treatment sessions, although specific timing depends on the disease or conditions acupuncture treats and the outcome measures. For example, acupuncture shows a 20% reduction in chronic low back pain after 4-5 sessions,[69] and it might take up to three months to show effects on post-stroke motor functions.[70] When investigators with explanatory objectives primarily, short follow-up times are likely to be optimal. Such design choices do not answer the question of interest to patients—can I expect a longer term benefit from acupuncture? Thus, acupuncture trials with mainly pragmatic objectives are likely to choose the longer follow-up times most important for clinical or policy decision making.[71]

Acupuncture trials published in the past five years mainly used short to medium term follow-up times (web appendix 2, 2.1.6). Trialists should take into account the course of disease or condition for optimal follow-up times. In many instances, when designing pragmatic or practical trials, trialists need to consider measuring outcomes after six months or longer to optimally inform clinical decision making (box 1, consideration 10; box 11).

Engaging patients

Consideration 11: How should trials engage patients when designing and conducting RCTs?

Patients are among the most critical stakeholders in clinical care and health research. Involving patients when designing trials can enhance the research question's relevance and importance, which helps to improve adherence (box 1, consideration 11).[72-74] Furthermore, having patients involved as members of the trial team can facilitate the understanding of research findings and therefore facilitate dissemination and application.[73 74]

As complex interventions, acupuncture therapies focus on patient engagement, though this is rarely included in randomised trials. Previous guidance on practical acupuncture trials also recommended engagement of relevant stakeholders (eg, practitioners, clinicians, patients, payers, researchers).[13] Our systematic survey[15] showed that over half of the acupuncture trials did not mention even the minimal extent of patient involvement—signing the informed consent form.

To empower patients and conduct more patient centric research, acupuncture trials should consider informing, educating, discussing, engaging, and partnering with patient engagement organisations such as Clinical Trials Ontario, which helps trial teams engage with patients.[75 76]

Conclusion

We have produced guidance to consider before and during the design of acupuncture trials to deal with the unique challenges and the most prevalent concerns in current acupuncture trials (box 1). With the rapid growth of published acupuncture trials and upcoming funding opportunities worldwide in integrative medicine, failure to enhance the rigour of acupuncture trials will result in additional research waste.[77] This guidance highlights the importance of granting agencies and researchers carefully considering the need for research in the trial planning and funding processes. Our guidance will help both researchers and granting agencies consider the critical decisions linking trial design and its inferences and application to produce trials that facilitate healthcare decision making.

AUTHOR AFFILIATIONS

[1]Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences, Beijing, China

[2]CEBIM (Center for Evidence-Based Integrative Medicine)-Clarity Collaboration, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China

[3]Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada

[4]Nottingham Ningbo GRADE Centre, University of Nottingham Ningbo, China

[5]China Center for Evidence-Based Traditional Chinese Medicine, China Academy of Chinese Medical Sciences, Beijing, China

[6]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland

[7]Virginia University of Integrative Medicine, Fairfax, VA, USA

[8]Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

[9]Institute for Excellence in Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

[10]Department of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[11]Kaiser Permanente Washington Health Research Institute, Seattle WA, USA

[12]Department of Epidemiology, University of Washington, Seattle WA, USA

[13]British Medical Acupuncture Society, London, UK

[14]Patient and Public Engagement, Clinical Trials Ontario, Toronto, ON, Canada

[15]Department of Western Medicine, Virginia University of Integrative Medicine, Fairfax, VA, USA

[16]Korean Medicine Clinical Trial Center, Seoul, Republic of Korea
[17]Korean Medicine Hospital, Seoul, Republic of Korea
[18]Kyung Hee University, Seoul, Republic of Korea
[19]Clinical Medicine Division, Korea Institute of Oriental Medicine, Daejeon, Republic of Korea
[20]Korean Convergence Medicine, University of Science and Technology, Daejeon, Republic of Korea
[21]National Jewish Health, Denver, CO, USA
[22]Institute of Social Medicine, Epidemiology and Health Economics, Charité Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
[23]Integrative Medicine Service, Memorial Sloan Kettering Cancer Center, Bendheim Integrative Medicine Center, New York, NY USA
[24]Translational Health Research Institute, Western Sydney University, Penrith, New South Wales, Australia
[25]China Academy of Chinese Medical Sciences, Beijing, China
[26]Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China
[27]Key Laboratory of Chinese Internal Medicine of Ministry of Education, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China
[28]School of Acupuncture-Moxibustion and Tuina/The Third Affiliated Hospital, Chengdu University of Traditional Chinese Medicine, Sichuan, China
[29]Center for Integrative Global Oral Health, School of Dental Medicine, University of Pennsylvania, Philadelphia, PA, USA

We thank Qi Zhou, McMaster University, Canada, for statistical and data interpretation support from the survey studies; Klaus Linde, Technical University of Munich, Andrew J Vickers, Memorial Sloan-Kettering Cancer Center, and Kim Bennell, University of Melbourne, for feedback on the manuscript; Yu-Tong Fei, Hui-Juan Cao, Ru-Yu Xia, Qian-Yun Chai, Chang-Hao Liang, YuTing Feng, and Yi-Ran Du, from Centre for Evidence-Based Chinese Medicine and Dongzhimen Hospital Beijing for several key references manusript; and Jie Wang, Guang'anmen Hospital, for methodological feedback.

Contributors: Y-QZ and R-MJ contributed equally to this work. Y-QZ and X-HJ are the initiators of this series of articles, defined the scope, content, structure, drafted the manuscript, and coordinated the research team. R-MJ identified the acupuncture trial examples, discussed the manuscript's structure, revised tables and main text, organised references, and coordinated the research team. GG was involved in planning this series of articles and designing the content and structure of this article. X-HJ, Y-QZ, GG, LT, CW, J-PL, R-MJ, and LL are steering committee members. They discussed the structure and content of the paper in four teleconferences and revised versions of the manuscript. JJM, LL, X-HJ, and DPR discussed and debated the paper's content at the Society of Acupuncture Research 2021 international research conference and contributed further to the content. W-JG provided statistical and data interpretation support for the survey studies. All authors reviewed, revised, and contributed to the creation of the manuscript. All authors approved the final version of the manuscript. X-HJ is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: The China Academy of Chinese Medical Sciences (CACMS) Innovation Fund (CI 2021A03503 and CI2021B011) and " The Belt and Road" Traditional Chinese Medicine Cooperation Project of China Academy of Chinese Medical Sciences (GH201901) supported this work. The funders had no role in considering the study design or in the collection, analysis, interpretation of data, writing of the report, or decision to submit the article for publication.

Competing interests: All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: CW received a grant at the University of Zurich from the Kelm Foundation for a randomized controlled trial on acupuncture for dysgeusia during chemotherapy. CW is an unpaid board Member of the Society of Acupuncture Research. KJS received a grant from NIH (National Center for Complementary and Integrative Health). KJS seats on a Data Safety Monitoring Board or Advisory Board from NIH-DoD (Department of Defence)-VA(Veterans Affairs) Collaboratory DSMB, and Advisory Board of the National Center for Complementary and Integrative Health of the US NIH with honorarium. MC receives royalties from non commercially used Medical Acupuncture textbooks.MC works as the Medical Director of the British Medical Acupuncture Society (BMAS) full time. DPR is a volunteer vice president of the Canadian Arthritis Patient Alliance, a patient-led and run arthritis organization that accepts arms-length funding from pharma companies, and an employee of Five02 Labs Inc, under Clinical Trials Ontario to support its patient and public engagement efforts. DPR received an honorarium attend meetings and travel support from Eli Lilly Canada for her speech about living with arthritis and arthritis patient organizations in Canada. MW received grants from the NIH, Sanofi, Astrazeneca Teva, and Jin Hua Foundation Cohero. MW received consulting fees from AstraZeneca, Amgen, Glaxosmithkline, Sanofi, Genzyme, Regeneron, Boehringer Ingelheim, Novartis, genentech, Pulmatrix, Teva, Equillium, cytoreason , Restorbio, Cohero Health, cerecor, incyte, sound biologics, and kinaset. MW received grants for payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from AstraZeneca, Glaxosmithkline, Sanofi, and Regeneron. PJD received grants from Abbott Diagnostics, Roche Diagnostics, and Siemens, the monitoring device of CloudDX, and the medical device of Philips Healthcare. PJD reports as a Member of the Scientific Advisory Board of Bayer and Quidel. The study was supported by the China Academy of Chinese Medical Sciences Innovation Fund (CI 2021A03503 & CI2021B011) and "The Belt And Road" Traditional Chinese Medicine Cooperation Project of China Academy of Chinese Medical Sciences; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; and no other relationships or activities that could appear to have influenced the submitted work of all authors.

Patient and public involvement: Dawn Richards, a patient with rheumatoid arthritis and osteoarthritis and director of patient and public engagement of Clinical Trials Ontario, and Eun-Kyung Anna Kim, a patient representative from Virginia University of Integrative Medicine, reviewed, revised, and provided feedback on the manuscript.

Provenance and peer review: Commissioned; externally peer reviewed.

This article is part of a collection funded by the special purpose funds for the belt and road, China Academy of Chinese Medical Sciences, National Natural Science Foundation of China, the National Center for Complementary and Integrative Health, the Innovation Team and Talents Cultivation Program of the National Administration of Traditional Chinese Medicine, the Special Project of "Lingnan Modernization of Traditional Chinese Medicine" of the 2019 Guangdong Key Research and Development Program, and the Project of First Class Universities and High-level Dual Discipline for Guangzhou University of Chinese Medicine. *The BMJ* commissioned, peer reviewed, edited, and made the decision to publish. Kamran Abbasi was the lead editor for *The BMJ*. Yu-Qing Zhang advised on commissioning for the collection, designed the topic of the series, and coordinated the author teams. Gordon Guyatt provided valuable advice and guidance.

1 State Council Information Office of the People's Republic of China. Traditional Chinese medicine in China. 2016. http://www.china.org.cn/chinese/2017-01/17/content_40120942.htm
2 Su CX. Evaluation of bias in acupuncture randomized controlled trials. [In Chinsese]. Beijing University of Traditional Chinese Medicine, 2015.
3 Wang YY. Bibliometric study on randomized controlled trials of acupuncture in chinese journals and the influencing factors of positive results. [In Chinese.] Beijing University of Traditional Chinese Medicine, 2014.
4 Oxman AD, Lombard C, Treweek S, Gagnier JJ, Maclure M, Zwarenstein M. Why we will remain pragmatists: four problems with the impractical mechanistic framework and a better solution. *J Clin Epidemiol* 2009;62:485-8. doi:10.1016/j.jclinepi.2008.08.015
5 Karanicolas PJ, Montori VM, Devereaux PJ, Schünemann H, Guyatt GH. A new "mechanistic-practical" framework for designing and interpreting randomized trials. *J Clin Epidemiol* 2009;62:479-84. doi:10.1016/j.jclinepi.2008.02.009
6 Shuai P, Zhou XH, Lao L, Li X. Issues of design and statistical analysis in controlled clinical acupuncture trials: an analysis of English-language reports from Western journals. *Stat Med* 2012;31:606-18. doi:10.1002/sim.4034
7 Zhuang Y, Xing JJ, Li J, Zeng BY, Liang FR. History of acupuncture research. *Int Rev Neurobiol* 2013;111:1-23. doi:10.1016/B978-0-12-411545-3.00001-8

8 Langevin HM, Wayne PM, Macpherson H, et al. Paradoxes in acupuncture research: strategies for moving forward. *Evid Based Complement Alternat Med* 2011;2011:180805. doi:10.1155/2011/180805
9 Paterson C, Britten N. Acupuncture as a complex intervention: a holistic model. *J Altern Complement Med* 2004;10:791-801. doi:10.1089/acm.2004.10.791
10 MacPherson H, Altman DG, Hammerschlag R, et al, STRICTA Revision Group. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. *PLoS Med* 2010;7:e1000261. doi:10.1371/journal.pmed.1000261
11 MacPherson H, White A, Cummings M, Jobst K, Rose K, Niemtzow R, STandards for Reporting Interventions in Controlled Trails of Acupuncture. Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations. Standards for Reporting Interventions in Controlled Trails of Acupuncture. *Acupunct Med* 2002;20:22-5. doi:10.1136/aim.20.1.22
12 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. . doi:10.1054/ctim.2001.0489
13 Witt CM, Aickin M, Baca T, et al, Effectiveness Guidance Documents Collaborators. Effectiveness guidance document (EGD) for acupuncture research - a consensus document for conducting trials. *BMC Complement Altern Med* 2012;12:148. doi:10.1186/1472-6882-12-148
14 Chen H, Yang M, Ning Z, et al. A guideline for randomized controlled trials of acupuncture. *Am J Chin Med* 2019;47:1-18. doi:10.1142/S0192415X19500010
15 Gang WJ, Xiu WC, Shi LJ, et al.Factors associated with the magnitude of acupuncture treatment effects (Famous): a meta-epidemiological study of acupuncture randomized controlled trials.*BMJ Open* (submitted).
16 Witt CM, Aickin M, Cherkin D, et al, Collaborators. Effectiveness guidance document (EGD) for Chinese medicine trials: a consensus document. *Trials* 2014;15:169. doi:10.1186/1745-6215-15-169
17 Fung KP, Chow OK, So SY. Attenuation of exercise-induced asthma by acupuncture. *Lancet* 1986;2:1419-22. doi:10.1016/S0140-6736(86)92733-9
18 Riddle DL, Johnson RE, Jensen MP, et al. The pragmatic-explanatory continuum indicator summary (PRECIS) instrument was useful for refining a randomized trial design: experiences from an investigative team. *J Clin Epidemiol* 2010;63:1271-5. doi:10.1016/j.jclinepi.2010.03.006
19 Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147. doi:10.1136/bmj.h2147
20 Haynes RB, Sackett DL, Guyatt GH, et al. How to do clinical practice research: a new book and a new series in the Journal of Clinical Epidemiology. *J Clin Epidemiol* 2006;59:873-5. doi:10.1016/j.jclinepi.2006.06.005
21 Haake M, Müller HH, Schade-Brittinger C, et al. German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups. *Arch Intern Med* 2007;167:1892-8. doi:10.1001/Archinte.167.17.1892
22 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86-9. doi:10.1016/S0140-6736(09)60329-9
23 Hinman RS, McCrory P, Pirotta M, et al. Acupuncture for chronic knee pain: a randomized clinical trial. *JAMA* 2014;312:1313-22. doi:10.1001/jama.2014.12660
24 Vas J, Méndez C, Perea-Milla E, et al. Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: randomised controlled trial. *BMJ* 2004;329:1216. doi:10.1136/bmj.38238.601447.3A
25 Thomas KJ, MacPherson H, Thorpe L, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain. *BMJ* 2006;333:623. doi:10.1136/bmj.38878.907361.7C
26 Constantino MJ, Ametrano RM, Greenberg RP. Clinician interventions and participant characteristics that foster adaptive patient expectations for psychotherapy and psychotherapeutic change. *Psychotherapy (Chic)* 2012;49:557-69. doi:10.1037/a0029440
27 Iles RA, Davidson M, Taylor NF, O'Halloran P. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *J Occup Rehabil* 2009;19:25-40. doi:10.1007/s10926-008-9161-0
28 Bialosky JE, Bishop MD, Cleland JA. Individual expectation: an overlooked, but pertinent, factor in the treatment of individuals experiencing musculoskeletal pain. *Phys Ther* 2010;90:1345-55. doi:10.2522/ptj.20090306
29 Mondloch MV, Cole DC, Frank JW. Does how you do depend on how you think you'll do? A systematic review of the evidence for a relation between patients' recovery expectations and health outcomes. *CMAJ* 2001;165:174-9.
30 Mao JJ, Xie SX, Bowman MA. Uncovering the expectancy effect: the validation of the acupuncture expectancy scale. *Altern Ther Health Med* 2010;16:22-7.
31 Barth J, Kern A, Lüthi S, Witt CM. Assessment of patients' expectations: development and validation of the Expectation for Treatment Scale (ETS). *BMJ Open* 2019;9:e026712. doi:10.1136/bmjopen-2018-026712
32 Jones SM, Lange J, Turner J, et al. Development and validation of the Expect questionnaire: assessing patient expectations of outcomes of complementary and alternative medicine treatments for chronic pain. *J Altern Complement Med* 2016;22:936-46. doi:10.1089/acm.2016.0242
33 Schandelmaier S, Briel M, Varadhan R, et al. Development of the Instrument to assess the Credibility of Effect Modification Analyses (ICEMAN) in randomized controlled trials and meta-analyses. *CMAJ* 2020;192:E901-6. doi:10.1503/cmaj.200077
34 Wasan AD, Kong J, Pham LD, Kaptchuk TJ, Edwards R, Gollub RL. The impact of placebo, psychopathology, and expectations on the response to acupuncture needling in patients with chronic low back pain. *J Pain* 2010;11:555-63. doi:10.1016/j.jpain.2009.09.013
35 Devereaux PJ, Bhandari M, Clarke M, et al. Need for expertise based randomised controlled trials. *BMJ* 2005;330:88-91. doi:10.1136/bmj.330.7482.88
36 Scholtes VA, Nijman TH, van Beers L, Devereaux PJ, Poolman RW. Emerging designs in orthopaedics: expertise-based randomized controlled trials. *J Bone Joint Surg Am* 2012;94(Suppl 1):24-8. doi:10.2106/JBJS.K.01626
37 Evaniew N, Carrasco-Labra A, Devereaux PJ, et al. How to use a randomized clinical trial addressing a surgical procedure: users' guide to the medical literature. *JAMA Surg* 2016;151:657-62. doi:10.1001/jamasurg.2016.0072
38 Johnston BC, da Costa BR, Devereaux PJ, Akl EA, Busse JW, Expertise-Based RCT Working Group. The use of expertise-based randomized controlled trials to assess spinal manipulation and acupuncture for low back pain: a systematic review. *Spine (Phila Pa 1976)* 2008;33:914-8. doi:10.1097/BRS.0b013e31816b4be4
39 Ma P, Xu S, Tian W, et al. [Efficacy observation of post-stroke pseudobulbar palsy treated with quick needle insertion therapy at *Aqiang* point][in Chinese]. *Zhongguo Zhen Jiu* 2016;36:1027-30.
40 Witt CM, Lüdtke R, Wegscheider K, Willich SN. Physician characteristics and variation in treatment outcomes: are better qualified and experienced physicians more successful in treating patients with chronic pain with acupuncture?*J Pain* 2010;11:431-5. doi:10.1016/j.jpain.2009.08.010
41 Moura CC, Chaves ECL, Chianca TCM, Ruginsk SG, Nogueira DA, Iunes DH. Effects of auricular acupuncture on chronic pain in people with back musculoskeletal disorders: a randomized clinical trial. *Rev Esc Enferm USP* 2019;53:e03418. doi:10.1590/s1980-220x2018009003418
42 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. doi:10.1136/bmj.39280.509803.BE
43 Bednarska E, Bryant D, Devereaux PJ, Expertise-Based Working Group. Orthopaedic surgeons prefer to participate in expertise-based randomized trials. *Clin Orthop Relat Res* 2008;466:1734-44. doi:10.1007/s11999-008-0273-9
44 Xu ZG, Ge SQ, Lu LM. [Application of expertise-based pragmatic randomized controlled trial in acupuncture-moxibustion clinical research][in Chinese]. *Zhongguo Zhen Jiu* 2020;40:538-42.
45 Takakura N, Yajima H. A placebo acupuncture needle with potential for double blinding - a validation study. *Acupunct Med* 2008;26:224-30. doi:10.1136/aim.26.4.224
46 Vas J, Cintado MC, Aranda-Regules JM, Aguilar I, Rivas Ruiz F. Effect of ear acupuncture on pregnancy-related pain in the lower back and posterior pelvic girdle: A multicenter randomized clinical trial. *Acta Obstet Gynecol Scand* 2019;98:1307-17. doi:10.1111/aogs.13635
47 Schlaeger JM, Takakura N, Yajima H, et al. Double-blind acupuncture needles: a multi-needle, multi-session randomized feasibility study. *Pilot Feasibility Stud* 2018;4:72. doi:10.1186/s40814-018-0265-9
48 Takakura N, Yajima H. A double-blind placebo needle for acupuncture research. *BMC Complement Altern Med* 2007;7:31. doi:10.1186/1472-6882-7-31
49 Zucker NA, Tsodikov A, Mist SD, Cina S, Napadow V, Harris RE. Evoked pressure pain sensitivity is associated with differential analgesic response to verum and sham acupuncture in fibromyalgia. *Pain Med* 2017;18:1582-92. doi:10.1093/pm/pnx001
50 Vickers AJ, Vertosick EA, Lewith G, et al. Do the effects of acupuncture vary between acupuncturists? Analysis of the Acupuncture Trialists' Collaboration individual patient data meta-analysis. *Acupunct Med* 2021;39:309. doi:10.1177/0964528420959089
51 Madsen MV, Gøtzsche PC, Hróbjartsson A. Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups. *BMJ* 2009;338:a3115. doi:10.1136/bmj.a3115

52 Moustgaard H, Clayton GL, Jones HE. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ* 2020;368:l6802. doi:10.1136/bmj.l6802
53 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;141:901-10. doi:10.7326/0003-4819-141-12-200412210-00006
54 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. doi:10.1002/art.41584
55 Kim TH, Lee MS, Alraek T, Birch S. Acupuncture in sham device controlled trials may not be as effective as acupuncture in the real world: a preliminary network meta-analysis of studies of acupuncture for hot flashes in menopausal women. *Acupunct Med* 2020;38:37-44. doi:10.1136/acupmed-2018-011671
56 Hershman DL, Unger JM, Greenlee H, et al. Effect of acupuncture vs sham acupuncture or waitlist control on joint pain related to aromatase inhibitors among women with early-stage breast cancer: a randomized clinical trial. *JAMA* 2018;320:167-76. doi:10.1001/jama.2018.8907
57 Nielsen A, Ocker L, Majd I, et al. Acupuncture intervention protocol: consensus process for a pragmatic randomized controlled trial of acupuncture for management of chronic low back pain in older adults: an NIH HEAL initiative funded project. *Glob Adv Health Med* 2021;10:21649561211007091. doi:10.1177/21649561211007091
58 Wenden AJ. Acupuncture for analgesia in the emergency department: a multicentre, randomised, equivalence and non-inferiority trial. *Med J Aust* 2018;208:189. doi:10.5694/mja17.00623
59 Li J, Ma H, Wu X. The key point and implementation point of international acupuncture clinical research [in Chinese]. Cross Strait Medical and Health Exchange Association, 2015.
60 DeVore AD, Granger BB, Fonarow GC, et al. Care optimization through patient and hospital engagement clinical trial for heart failure: rationale and design of Connect-HF. *Am Heart J* 2020;220:41-50. doi:10.1016/j.ahj.2019.09.012
61 Avis NE, Coeytaux RR, Isom S, Prevette K, Morgan T. Acupuncture in menopause (AIM) study: a pragmatic, randomized controlled trial. *Menopause* 2016;23:626-37. doi:10.1097/GME.0000000000000597
62 Du Y. Thoughts on clinical research of acupuncture and moxibustion[in Chinese]. *Zhongguo Zhenjiu* 2000;20:312-4.
63 Goldfarb D. *S., et al, Patients at the center: in our practice, and in our use oflanguage.* ACP J Club, 2004.
64 Clarke M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007;8:39. doi:10.1186/1745-6215-8-39
65 Deyo RA, Dworkin SF, Amtmann D, et al. Report of the NIH Task Force on research standards for chronic low back pain. *Pain Med* 2014;15:1249-67. doi:10.1111/pme.12538
66 Smith TO, Hawker GA, Hunter DJ, et al. The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. *J Rheumatol* 2019;46:981-9. doi:10.3899/jrheum.181194
67 Smith CA, de Lacey S, Chapman M, et al. Erratum to acupuncture to improve live birth rates for women undergoing in vitro fertilization: a protocol for a randomized controlled trial. *Trials* 2017;18:248. doi:10.1186/s13063-017-1986-4
68 Cherkin DC, Eisenberg D, Sherman KJ, et al. Randomized trial comparing traditional Chinese medical acupuncture, therapeutic massage, and self-care education for chronic low back pain. *Arch Intern Med* 2001;161:1081-8. doi:10.1001/archinte.161.8.1081
69 Pach D, Yang-Strobel X, Lüdtke R, et al. Standardized versus individualized acupuncture for chronic low back pain: a randomized controlled trial. *Evid Based Complement Alternat Med* 2013;2013:125937. doi:10.1155/2013/125937
70 Wayne PM, Krebs DE, Macklin EA, et al. Acupuncture for upper-extremity rehabilitation in chronic stroke: a randomized sham-controlled study. *Arch Phys Med Rehabil* 2005;86:2248-55. doi:10.1016/j.apmr.2005.07.287
71 Santosa KB, Qi J, Kim HM, Hamill JB, Wilkins EG, Pusic AL. Long-term patient-reported outcomes in postmastectomy breast reconstruction. *JAMA Surg* 2018;153:891-9. doi:10.1001/jamasurg.2018.1677
72 Domecq JP, Prutsky G, Elraiyah T, et al. Patient engagement in research: a systematic review. *BMC Health Serv Res* 2014;14:89. doi:10.1186/1472-6963-14-89
73 Brett J, Staniszewska S, Mockford C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. *Health Expect* 2014;17:637-50. doi:10.1111/j.1369-7625.2012.00795.x
74 Crocker JC, Ricci-Cabello I, Parker A, et al. Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. *BMJ* 2018;363:k4738. doi:10.1136/bmj.k4738
75 Skovlund PC, Nielsen BK, Thaysen HV, et al. The impact of patient involvement in research: a case study of the planning, conduct and dissemination of a clinical, controlled trial. *Res Involv Engagem* 2020;6:43. doi:10.1186/s40900-020-00214-5
76 Price A, Albarqouni L, Kirkpatrick J, et al. Patient and public involvement in the design of clinical trials: An overview of systematic reviews. *J Eval Clin Pract* 2018;24:240-53. doi:10.1111/jep.12805
77 Glasziou P, Chalmers I. Research waste is still a scandal. *BMJ* 2018;363:k4645. doi:10.1136/bmj.k4645

Web appendix 1: Details of the methods for development of this guidance
Web appendix 2: Data and supporting literature for the guidance