Published in final edited form as:
*Stat Med*. 2012 March 30; 31(7): 606–618. doi:10.1002/sim.4034.

# Issues of design and statistical analysis in controlled clinical acupuncture trials: An analysis of English-language reports from Western journals

**Ping Shuai**[a], **Xiao-Hua Zhou**[b,*], **Lixing Lao**[c], and **Xiaosong Li**[a,*]

[a]Dept. Health Statistics, West China School of Public Health, Sichuan University, Chengdu, China

[b]Dept. Biostatistics, School of Public Health and Community Medicine, University of Washington, WA, USA

[c]Dept. Family and Community Medicine, School of Medicine, University of Maryland, MD, USA

## Abstract

**Objective—**To investigate major methods of design and statistical analysis in controlled clinical acupuncture trials published in the West during the past six years (2003–2009) and, based on this analysis, to provide recommendations that address methodological issues and challenges in clinical acupuncture research.

**Method—**PubMed was searched for acupuncture RCTs published in Western journals in English between 2003 and 2009. The keyword used was acupuncture.

**Results—**One hundred and eight qualified reports of acupuncture trials that included more than 30 symptoms/conditions were identified, analyzed, and grouped into efficacy (explanatory), effectiveness (pragmatically beneficial) and other (unspecified) studies. All were randomized controlled clinical trials (RCTs). In spite of significant improvement in the quality of acupuncture RCTs in the last 30 years, these reports show that some methodological issues and shortcomings in design and analysis remain. Moreover, the quality of the efficacy studies was not superior to that of the other types of studies. Research design and reporting problems include unclear patient criteria and inadequate practitioner eligibility, inadequate randomization and blinding, deficiencies in the selection of controls, and improper outcome measurements. Problems in statistical analysis included insufficient sample sizes and power calculations, inadequate handling of missing data and multiple comparisons, and inefficient methods for dealing with repeated-measure and cluster data, baseline value adjustment, and confounding issues.

**Conclusion—**Despite recent advancements in acupuncture research, acupuncture RCTs can be improved, and more rigorous research methods should be carefully considered.

### Keywords

acupuncture; randomized controlled clinical trials; methodology design; statistical analysis

Because conventional medicine has limitations and side effects, increasing numbers of people turn to complementary and alternative modalities, including acupuncture, for treating their diseases and alleviating their symptoms. Since the 1970s numerous clinical trials have

*Corresponding authors: Xiao-Hua Zhou, Dept. Biostatistics, School of Public Health and Community Medicine, University of Washington, WA, USA 98198, azhou@u.washington.edu, Xiaosong Li, Dept. Health Statistics, West China School of Public Health, Sichuan University, No.17, 3rd Sect., Renmin South Road, Chengdu, China, 610041, lixiaosong1101@126.com.

been conducted in the West to evaluate the effects of acupuncture, but the results remain controversial [1].

Recently, the randomized controlled clinical trial (RCT), considered the "gold standard" in conventional clinical research, has been employed extensively to study acupuncture. However, some researchers argue that the RCT might not be the most suitable method for evaluating that modality, which originated in traditional Chinese medicine (TCM), a medical system very different from that of conventional medicine in terms of diagnostic and therapeutic principles [2, 3]. While it may be true that by its nature acupuncture poses difficulties for researchers, the contradictory and inconclusive results of acupuncture RCTs may speak more to their methodological shortcomings [4, 5] than to philosophical distinctions between conventional medicine and TCM. Though the quality of acupuncture RCTs has significantly improved over the last 30 years, many researchers point out that meticulous attention to detail, more rigorous design, and better explanation of results are needed [6]. The purpose of this paper is to identify the main methodological issues of design and statistical analysis in recently published acupuncture RCTs in order to provide useful recommendations for improving the quality of future clinical trials.

## Materials and Methods

Our focus was on acupuncture RCTs published in English in Western journals from August 2003 to December 2009, primarily those referenced in PubMed. Inclusive criteria were as follows: manual acupuncture (an acupuncture technique that involves manually rotating, lifting, and thrusting a needle); treatment primarily based in TCM; treatment performed for curative, palliative, or preventive purposes. We excluded observational studies and reports written in other languages. Considering that certain alternative and complementary therapies related to acupuncture may have different mechanisms that might confound the results, we excluded trials of acupressure, moxibustion, transcutaneous nerve stimulation, electro-acupuncture, and auricular acupuncture. Our analysis is not intended to be a systematic review.

## Results

Accordingly, 108 identified reports [7–114] qualified for review and analysis. They encompass a wide range of conditions/symptoms, including: osteoarthritis/knee pain, migraine, low back pain, pelvic pain, hot flashes, labor pain and labor initiation, analgesic and sedative effects in surgery/ cancer pain, tension-type headache, shoulder/arm pain, embryo transfer, irritable bowel syndrome, allergic rhinitis, depression, fibromyalgia, myofascial pain, insomnia, inflammatory breast in lactating women, stroke rehabilitation, anxiety, nausea and vomiting, dry eyes, hypertension, PTSD (Post-Traumatic Stress Disorder), recurrent cystitis, psychogenic erectile dysfunction, chronic post-stroke leg spasticity, overactive bladder, refractory heartburn, rheumatoid arthritis, disabling breathlessness, impingement syndrome, cervical spondylosis, COPD (Chronic Obstructive Pulmonary Disease), diabetic bladder dysfunction, balance function after stoke. The sample size of those articles varied from sixteen to over three thousand. The reports had problems both in design and statistical analysis.

### Part I. Issues in Design

For the sake of analysis and comparison, we roughly categorized the 108 studies into three groups, based on design and author's statement of purpose. There were 45 efficacy studies, 44 effectiveness studies, and 19 that could not be defined. Clear statement of the study objective contributes to the use of appropriate research models and methods.

Efficacy trials are used to determine whether an intervention works and is safe when delivered under controlled conditions, while effectiveness trials evaluate the usefulness of an intervention's pragmatic "real-world" applicability [115]. Efficacy studies should be well controlled and more rigorously designed than effectiveness studies, while the latter are focused on how much benefit patients gain from an intervention in actual practice. In our analysis, we found only a few efficacy studies that conformed to strict and precise designs [7, 21, 37, 46, 96].

**1. Eligibility of Patients and Acupuncturists—**Patient source is one of the most important issues in patient eligibility and is especially crucial to efficacy studies. Some reports did not state clearly the source of patients or used simple sampling. In total, only 26 -- 14 (29%) efficacy studies, 9 (19%) effectiveness studies, and 3 (15%) unspecified studies -- described in detail whether their patients were recruited through advertisements or other media.

Adequate and detailed inclusion and exclusion criteria can ensure a homogeneous sample. Although most of these studies clarified their selection criteria, nineteen, including four efficacy studies, did not report inclusion and exclusion criteria in detail.

Considering the uniqueness of acupuncture, in studies using non-needle-insertion sham acupuncture as control, subjects ideally should have no previous experience of acupuncture. A patient who has had true acupuncture treatment will easily recognize the sham as a fake treatment, thus confounding the blinding, so some researchers require that included patients be acupuncture naïve. Table 1 is a breakdown of how the issue of previous experience was handled in these studies.

The following categories of practitioners were reported as having performed the treatments: acupuncturists, physiotherapists, physicians, researchers, nurses, and midwives. STRICTA (Standards for Reporting Interventions in Controlled Trials of Acupuncture) recommendations by MacPherson et al. state that practitioner qualifications should be carefully established and clearly reported [116]. While point selection rationale, needling details, treatment regimen (frequency and duration of treatment) and co-interventions are important items in STRICTA, adequate training is particularly important. However, 53 reports, including 21 efficacy reports, provided no detailed information on practitioners' acupuncture education, training, background, or experience. Some simply stated that the acupuncturists were "experienced" or "qualified" without providing specific information.

A well-trained and experienced practitioner is likely to deliver a treatment with adequate point selection, needling technique, and treatment regime. As acupuncture is a needling-skill-driven therapy, the use of qualified acupuncturists improves the quality and credibility of a study. Furthermore, adequate training of the acupuncturists in the specific treatment protocol will enhance standardization and minimize variation in their interactions with patients [117]. In addition, consultation with acupuncture experts or consensus among a group of experienced acupuncturists regarding the acupuncture treatment procedure would be a good approach for designing a clinically valid study treatment protocol.

**2. Randomization and Masking—**Randomization helps control for variations among patients and ensures that both known and unknown confounding factors are evenly distributed between groups. Every reviewed report mentioned using randomization to allocate patients.

Hedden et al. [110], recommends that in a clinical trial with a small sample size ($n<200$), where the probability of imbalance between comparison groups is high, use of modern

methods of randomization, such as stratification, is best. Stratification is a procedure for achieving balance between groups in baseline characteristics, while blocking is used to achieve balance in size. Blocking can be used to ensure close numerical balance in groups at any time during a trial. To maintain blinding by keeping participants from deducing their treatment from the block size, larger or randomly varied block sizes may be used. In about half of the reviewed trials the researchers considered this issue and reported their blocking procedure. Stratified randomization provides increased confidence that groups are comparable with respect to known prognostic factors. A total of 8 (18%) of the efficacy, 5 (11%) of the effectiveness, and 2 (11%) of the unspecified studies used stratified randomization for such factors as clinic center, sex, and symptom baseline scores.

Masking patients and practitioners is done to minimize possible bias in general RCTs. Clearly masking is important. However, given the nature of acupuncture, some researchers claim that masking acupuncture practitioners is difficult or impossible and posit study designs with a blinded evaluator and blinded patients as alternatives to the standard double-blinded study used in drug trials [119,120].

In the 108 studies, two main forms of masking were used: patient-only blinding, and patient/evaluator blinding. The different methods used in masking are listed in Table 2. Three efficacy studies and two effectiveness studies attempted to blind both patients and practitioners. In the latter, the investigator used two teams of acupuncturists that were blinded to patient allocation. The first team, responsible for diagnosis and prescription, sealed the point prescriptions in envelopes given to a research assistant. Depending on patient assignment, the sealed envelopes that the acupuncturists in the second team received either contained an appropriate prescription made by the first team or a set of irrelevant acupuncture points for control. Practitioner blinding has clear limitations, since acupuncture treatment is highly skill dependent. An experienced acupuncturist can easily tell whether an acupuncture protocol suits the patient's condition. Thus the double blinding used in standard drug trials may not be appropriate in trials of such non-pharmacologic interventions such as acupuncture and physical therapy, and it remains controversial.

Evaluation of the credibility of blinding, however, can be useful since inadequate blinding is associated with bias. Of the trials reviewed, 18 (40%) of the efficacy, 14 (32%) of the effectiveness, and 1 (5%) of the unspecified studies included such a test. The most common approach was to ask subjects which treatment they believed they had received. This approach has its limitations; there is no methodological consensus on whether such tests are appropriate [130], and some studies show that they are problematic since the guess may depend on clinical outcome if the credibility test is conducted at the end of a trial [133]. On the other hand, if this method is used at an early stage before the patient has noticed any treatment effect, it may be adequate for validating patient blinding to treatment procedure since influence from treatment result is excluded [132,133].

Better methods for evaluating the success of blinding are urgently needed [130]. To achieve better patient blinding, it is suggested that investigators make every effort to ensure that subjects from different groups do not meet in the waiting room, where they might compare treatment experiences. Additionally, point location diagrams and reading materials that discuss acupuncture, and thus potentially may reveal treatment assignment, should be removed from the clinical setting.

In addition to the blinding efforts discussed above, treatment concealment can be better achieved if the study practitioners use neutral language and avoid discussing the treatment protocol. Using a previously prepared script for acupuncturist/patient conversation might

helpful to minimize the potential bias by an acupuncturist as he or she is usually not blinded to the treatment assignment. [121].

**3. Control Group Choice—**Control group choice is dependent on study purpose and the acupuncture modality to be tested. In these reports, control groups varied widely. Sham acupuncture was the most common control; 17 (38%) efficacy, 14 (32%) effectiveness, and 7 (37%) unspecified studies compared true acupuncture with sham acupuncture. Less frequent were reports comparing acupuncture to standard therapy or to no treatment, acupuncture plus standard treatment to standard treatment, and three-arm design comparing acupuncture, sham acupuncture, and standard or no treatment. In the effectiveness studies, the acupuncture arm was often compared to either standard care [16, 18, 23, 35, 41, 62, 67, 89, 113], or other established treatments such as physical therapy [9, 14, 38]. Although the pitfall of this type of trial is that patients cannot be blinded to treatment assignment, the advantage is that such trials reflect real-world clinical practice, and their data can help clinicians determine the choice of therapy in daily practice.

Although sham acupuncture was the most common control, there is controversy over how to choose an appropriate sham [122,123]. In the studies reviewed, four main types of sham acupuncture were used based on needling and point selection: non-insertion at non-acupuncture points, non-insertion at true points, insertion or superficial insertion at non-acupuncture points, and superficial insertion at true points. The most popular type was insertion or superficial insertion at non-acupuncture points (see Table 3). But minimal acupuncture, defined as needles inserted into non-acupuncture points and/or superficially used as placebo control, is not necessarily inert and may have both specific and non-specific effects [131]. Indeed, there is general agreement that invasive sham acupuncture is not inert. As Langevin et al. [123] stated, enhanced benefit from sham acupuncture, especially likely in the case of invasive sham needling, may result from locally induced changes in microcirculation, immune system responses, or nonspecific pain modulation mechanisms. And more invasive sham procedures, which have been shown by neuroimaging to affect pain processing regions, should be also avoided [123].

In most of the reviewed studies, non-acupuncture control points were not well defined. Some researchers use true acupuncture points unrelated to the condition being treated as non-acupuncture controls, while others used points a few centimeters away from the true points. As the size of a real acupuncture point is still not clearly defined and some schools of thought do not regard traditional acupuncture points as fixed loci, these methods of determining a non-acupuncture control point are problematical.

Different acupuncture controls in clinical trials may answer different research questions. In clinical drug trial design, the specific effect of a targeted receptor is usually clearly defined, and the placebo control pill is known to be inert. In contrast, in an acupuncture trial an investigator who chooses a non-specific sham control to determine whether acupuncture produces a non-specific effect must first clearly understand how acupuncture works and what the specific effect of acupuncture is. Because the validity of acupuncture theories, mechanisms, point indications, locations, and needling methods has not been established, we suggest that non-insertion at non-acupuncture points is generally the best sham control, as it minimizes non-specific physiological confounding effects. However, care needs to be taken even with this procedure, since some researchers maintain that touch or pressure on any point may produce unintended physiological effects [123]. Some investigators state that noninvasive sham procedures, such as tapping an empty needle guide tube on the skin surface, may trigger physiological regulatory responses, especially if done at acupuncture points, and according to some classical theories, that of certain Japanese schools for example, a non-penetrating needle is considered to have a therapeutic effect [122].

**4. Outcome—**Adequate choice of appropriate outcomes measurements is important. However, not every report we reviewed clearly stated outcomes. A total of 13 (29%) efficacy, 13 (30%) effectiveness and 11 (58%) unspecified studies made no mention of their outcomes in their method sections, although some reported these in their results sections.

In clinical studies, objective outcome measurement is preferred to subjective [124], and acupuncture trials using objective outcome measurements are encouraged. However, in studies on certain conditions, such as pain, only subjective measurement can be used. Because most of the trials we reviewed were on chronic pain, subjective measurements were the most common (see Table 4). The researchers often asked patients to use the visual analog scale (VAS) or numerical rating scales (NRS) at baseline and in post-intervention assessments to rate the intensity of their pain. Although these are standard and validated instruments, self-reported pain intensity can vary greatly from patient to patient, and a large sample size is needed to address this variation.

Some of the reviewed reports used questionnaires to obtain outcome measurements, some standard and widely used, others self-designed. However, only a few authors discussed the reliability and validity of those questionnaires: 13 (29%) in efficacy studies, 14 (32%) in effectiveness studies and 2 (11%) in unspecified studies. We suggest that the reliability and validity of outcome measurement instruments should be discussed in research reports, although in the case of widely used instruments, reference citations are sufficient.

## Part II. Issues of Statistical Analysis

**1. Sample Size and Power Calculation—**Sample size and power calculations are necessary when designing trials and planning for resource utilization. These were not well reported in the reviewed studies. A total of 20 (44%) efficacy, 12 (27%) effectiveness, and 9 (47%) unspecified studies did not mention them at all. Calculating sample size requires mean pain score, standard deviation, and effect size. In all, 59 (55%) reports gave the effect size, and 23 (21%) described the standard deviation.

Sample size calculations are approximate because the calculation equations themselves are often based on approximations to the exact statistical distributions, and some of the parameters in the equation depend on guess work [125]. Some of the studies could not achieve the original power expected because of patients who dropped out or were lost to follow-up. Only 22 (20%) trials accounted for such problems and allotted 10% to 30% more participants to compensate for possible dropouts or missing data.

**2. Dropouts and Missing Data—**Dropout reasons varied and included moving out of town, intolerance to adverse effects, and feeling worse after treatment. Some of these would lead to biased results if ignored by the analyst. A total of 22 (49%) efficacy, 26 (59%) effectiveness, and 12 (63%) unspecified studies gave detailed descriptions of their dropouts, including number of dropouts and time and reason for dropping out. Characteristics or baseline values between dropouts and remaining patients were compared in 5 (11%) efficacy, 2 (5%) effectiveness, and 2 (11%) unspecified studies. If a comparison test shows a significant result, then the missing data in the dropouts are not missing completely at random (MCAR). MCAR is one of the three main mechanisms for characterizing missing data, missing at random (MAR) and missing not at random (MNAR) being the other two. If comparison shows that the result is not significant, the missing data still may not be MCAR, since the samples may differ in some unmeasured characteristics, but the comparison does serve as an informal check when determining the best method for dealing with dropouts.

In reporting a trial, it is essential to disclose any missing data and state how they were handled. In the reviewed studies, the most popular methods of dealing with missing data

problems were replacing missing values with baseline data and 'last observation carried forward' (LOCF). Six trials combined two or more methods, such as LOCF plus multiple imputation or LOCF plus replacement by mean scores obtained. Table 5 shows the methods used. Although using LOCF, mean score, or baseline score to replace a missing value is convenient and easy to handle, those methods are neither efficient nor valid. Since more powerful statistical methods for dealing with missing data, for example multiple imputations and weighted estimating equations, have been developed, we suggest using more sophisticated ways to handle this problem.

**3. Issues in Multiple Comparison—**Acupuncture studies may have multiple arms. For example, a design may compare a true acupuncture group to several different sham acupuncture groups, or true and sham acupuncture to conventional therapy. In such cases multiple comparison tests should be used to avoid the potential increase of type I error that occurs when statistical tests are used repeatedly. Table 6 showed that multiple comparison tests were not always conducted when a trial consisted of more than two groups that should have been compared to each other. Instead, some used *t* test to do pairwise comparisons, which may lead to faulty conclusions. Multiple comparison methods used in these studies include the Bonferroni correction test, a closed testing procedure, Scheffe's test, and *a priori*-ordered two-sided null hypotheses using Student's *t* test.

**4. Controlling Baseline Measurements—**The reporting of baseline demographic and clinical characteristics of each group is one of the requirements of the CONSORT (Consolidated Standards of Reporting Trials) statement [126]. Knowing baseline characteristics allows the reader to assess how closely matched the patients are and to judge the validity of the trials. In studies where important baseline factors appear well balanced, it is likely that any differences in outcome between the intervention and control groups are a real effect of treatment. Although most acupuncture clinical trials include randomization and some baseline comparisons between groups, certain unknown and important characteristics may be unbalanced in the baseline measurements. Thus Vickers [127] agrees that using co-variance (ANCOVA) to adjust for baseline scores is the preferred method of analysis in most acupuncture pain trials. He described four trials in which less powerful, unadjusted methods were used and reanalyzed them with ANCOVA. Reanalysis produced different conclusions. Therefore, in addition to comparing baseline characteristics, it is also necessary to use efficient statistical methods such as ANCOVA and regression to control potentially unbalanced baseline values between groups and to achieve more powerful conclusions. Besides ANCOVA, some studies reviewed also used multiple linear regressions, multivariate longitudinal analysis, or the linear mixed-effect model to ensure that baseline scores be well controlled and adjusted.

**5. Issues in Repeated-Measure Analysis—**In acupuncture trials, researchers would like to see long-term effects as well as short-term ones. This means that follow-up is done at intervals after treatment. Data collected in this way is called repeated-measures data or longitudinal data and involves measurements made on the same subject at different time points. Measurements made on the same variable for the same subject are likely to be correlated. Statistical models appropriate for longitudinal or repeated-measures data involve estimation of covariance parameters to capture correlations [128]. However, of those reports we reviewed, less than half made the proper analysis of repeated-measure data, using instead simple *t* or chi square tests to compare differences between groups. This neglected the aspect of time-treatment interaction and lowered the statistical efficiency.

Of the 72 trials with repeated-measure data, 19 (50%) efficacy, 8 (29%) effectiveness, and 3 (50%) unspecified studies considered repeated-measure problems and applied the corresponding analysis. Methods used to deal with repeated-measure data in these studies

included repeated-measure ANOVA, multivariate repeated-measure ANOVA (MANOVA), the linear mixed model, and the GEE model. Repeated-measure ANOVA is the traditional way to manage longitudinal data. It also was the most popular in those reports we reviewed and was used in 63% of the efficacy, 63% of the effectiveness, and 67% of the unspecified studies. However, recent research suggests that the linear mixed model is a more efficient method to manage such data. This model allows users to construct designs with a broad selection of covariance structures, thus offering greater efficiency than the multivariate normal distribution model and more flexibility than models assuming sphericity. Additionally, the linear mixed model also allows subjects to have missing time points, while software for traditional repeated-measures ANOVA drops a subject from analysis if it has missing data at a single time point. Therefore we suggest that researchers consider more sophisticated and efficient methods such as the linear mixed-model to handle repeated-measure data.

**6. Issues in Analysis of Cluster Data—**In the acupuncture trials reviewed, patients were collected from several different clinic centers or areas, and patients in certain acupuncture groups were treated by several different acupuncturists. Patient data actually may be grouped, or sorted into clusters. The major issue in the analysis of cluster data is that observations within a cluster are not independent. The degree of similarity is typically measured by the intracluster correlation coefficient (ICC) [129]. Ignoring intracluster correlation could lead to incorrect P values, biased estimates, and very small confidence intervals, all of which could cause incorrect interpretation of associations among variables. The cluster effect is a combination of both intracluster correlation and cluster size. Small intracluster correlations coupled with large cluster size can affect the validity of conventional statistical analyses. Under such circumstances, special analysis of cluster data is essential. However, few researchers in the studies we reviewed addressed this problem. Of the 108 reports, 57 had cluster data involving either patients from multiple clinic centers, patients of a single group being treated by several physicians, or both. But cluster data analysis was presented in only 9 (16%).

Methods used to deal with cluster data are generally called multilevel analyses and are intended to explore, in greater depth, interrelationships among variables at various levels, such as clinic level, physician level, and so forth. For example, in a data set with patient and physician clusters, a comprehensive multilevel data analysis can directly assess the effects of patient- and physician-level variables on outcome and can determine whether physician-level variables moderate patient-level relationships by examining cross-level interactions between variables from both levels. In the 9 trials reporting cluster analysis, methods used included the random-effect regression model, the linear mixed model, and the GEE model.

**7. Confounding Issues—**In pain-related trials such as those of head pain, low back pain, and osteoarthritis, patients are usually allowed to take analgesics before, during, and after treatment and during the follow-up period. The taking of analgesics and the amount of intake are important variables which can affect treatment effect assessment greatly, and they also can be potential confounding factors for acupuncture effect analyses. For example, if patients in a true acupuncture group who seldom take analgesics have almost the same pain evaluation scores after treatment as those in a sham group taking analgesics often and in high dosages, how can investigators determine whether treatment effects are due to the acupuncture or the analgesics? Even when researchers in our sample compared analgesic usage situation between or among groups, they did not control for analgesic usage in primary outcome measures, which may invalidate their conclusions. Using advanced statistical techniques such as multivariate analysis and making analgesics usage a covariant can control for this possibly confounding effect.

## Discussion

The recently published acupuncture trials illustrate the wide range of applications for which acupuncture is being used. Trials aiming to evaluate the efficacy of acupuncture are little better in quality and credibility than the effectiveness and unspecified studies, although the efficacy studies should have been rigorously designed and conducted. Good RCT design involves many aspects. Here we covered only the most common and essential of these, drawing from examples of current acupuncture RCTs, which point up the fact that much more attention needs to be given to the basics of trial design.

Acupuncture RCT design has its controversies. To avoid bias in patient selection, patient source and inclusion and exclusion criteria should be carefully considered and determined based on research purpose. Some researchers suggest that patient selection should include the complementary and alternative medicine diagnosis, not just the Western medicine diagnosis, if trials are to demonstrate more specific and direct treatment effects. Controversies regarding placebo control and effect in acupuncture RCTs are also widespread. Some researchers recommend that if a non-insertion sham is used, thorough control procedures and methods should be applied to ensure that the sham is really inert [122]. They insist that acupuncture trials should not be labeled “placebo-controlled” unless the investigator can present evidence that the sham procedure is in fact inert. Since we see an increasing use of placebo-controlled studies for acupuncture, better and verifiable controls are needed, as are more sophisticated and rigorous research models and methods for trials of complex acupuncture interventions.

In order to improve the research quality and obtain convincing and scientific results, comprehensive and rigorous methodological designs are not enough. Appropriate and well-considered methods of statistical analysis are also necessary. With the rapid developments in statistical methodology, the linear mixed model is now thought to be more efficient than ANOVA/MANOVA methods for handling repeated-measure and longitudinal data. Especially in a clinical trial in which data is unbalanced due to mistimed measurement or missing data, or which includes data with time-varying and time-invariant covariates, the linear mixed-effect model is much more flexible and parsimonious [134]. Moreover, when applying sophisticated statistical methods, we must keep scope of application in mind and be aware that not all methods are suitable in a given trial. For example, Vickers [127] mentions that ANCOVA should be used only if regression slopes are parallel and data are either normal and homoscedastic or sample sizes are large (n=25 or more per group), because ANCOVA depends on assumptions of normality of distribution and homoscedasticity, or constant variance across observations. The linear mixed-effect model is recommended for handling longitudinal data with continuous responses that are at least approximately normally distributed. However, when responses are discrete, linear models are not appropriate for relating changes in mean response to covariates. Statisticians have developed other methods based on the extended generalized linear model to handle discrete longitudinal data. These include the GEE model, the generalized mixed-effect model, and transition or response-conditioned models. Multiple imputation (MI) is considered one of the best ways to handle missing data and has been widely used. But researchers should also be aware that such factors as missing mechanism (ignorable missing or non-ignorable missing), and data type are key when choosing a suitable MI method.

We would like to emphasize that the quality of acupuncture trials has greatly improved during the last 30 years. Because of rapid developments in biostatistics, many statistical methods discussed in this paper were probably adequate at the time of these trials, and more recent advancements in statistical methods may not have been available to the researchers.

The purpose of our paper is not to criticize past work, but to introduce newer and more powerful analytic methods in the hope that these will be useful for future studies.

## Conclusion

As under current conditions there are no better or more suitable methods for testing the effects of acupuncture than RCTs, we must conduct rigorous RCTs that yield credible results. Although the quality of acupuncture research has greatly improved, due to the specificity and complexity of acupuncture research not all standards generally applied in the RCT are suitable for acupuncture trials, and not all researchers apply the basic standards of the RCT to their acupuncture studies. Issues such as how to choose a suitable control group and whether a valid placebo effect can be achieved are controversial. Ambiguity and uncertainty in such issues are obstacles for researchers who want to improve the quality of their research. Thus there is an urgent need to develop more valid and precise designs. However, with the development of sophisticated statistical methods, appropriate and efficient analysis can now not only assure accurate results, but also may help researchers overcome difficulties and dilemmas in design.

Hopefully, shedding light on the issues discussed here can inspire researchers to improve the design, implementation, analysis, and reporting of acupuncture RCTs. Only improvements in quality and credibility in acupuncture research can lead to better and clearer interpretations and conclusions regarding acupuncture's effects, thus providing good studies for systematic reviews that can present "the best available evidence" to clinicians, health policymakers, and consumers on this important modality.

## Acknowledgments

Dr. Lixing Lao was supported by NIH, NCCAM Grant #: R24 AT001293. Dr. Xiaohua Zhou was supported by NIH, NCCAM Grant #: R13AT005399. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official views of NIH. We wish to thank Dr. Lyn Lowry for editorial assistance.

## References

1. Birch S, Hesselink JK, Jonkman F, et al. Clinical research on acupuncture: Part 1. What have reviews of the efficacy and safety of acupuncture told us so far? J Altern Complement Med. 2004; 10:468–480. [PubMed: 15253851]
2. Shea, Jeanne L. Applying evidence-based medicine to traditional Chinese medicine: Debate and strategy. J Altern Complement Med. 2006; 12:255–263. [PubMed: 16646724]
3. Bernie Carter. Methodological issues and complementary therapies: researching intangibles? Complementary Therapies in Nursing & Midwifery. 2003; 9:133–139. [PubMed: 12852929]
4. Birch S. Clinical Research on Acupuncture: Part 2. Controlled Clinical Trials, an Overview of Their Methods. J Altern Complement Med. 2004; 10:481–498. [PubMed: 15253852]
5. Linde K, Jonas WB, Melchart D, et al. The methodological quality of randomized controlled trials of homeopathy, herbal medicines and acupuncture. Int J Epidemiol. 2001; 30:526–531. [PubMed: 11416076]
6. Sherman KJ, Cherkin DC. Challenges of acupuncture research: study design considerations. Clinical Acupuncture and Oriental Medicine. 2003; 3:200–206.
7. Witt C, Brinkhaus B, Jena S, et al. Acupuncture in patients with osteoarthritis of the knee: a randomized trial. Lancet. 2005; 366(9480):136–143. [PubMed: 16005336]
8. Scharf HP, Mansmann U, Streitberger K, et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. Ann Intern Med. 2006; 145:12–20. [PubMed: 16818924]
9. Williamson L, Wyatt MR, Yein K, et al. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. Rheumatology. 2007; 46(9):1445–1449. [PubMed: 17604311]

10. Witt CM, Jena S, Brinkhaus B, et al. Acupuncture in patients with osteoarthritis of the knee or hip. Arthritis & Rheumatism. 2006; 11:3485–3493. [PubMed: 17075849]
11. Foster NE, Thomas E, Barlas P, et al. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomized controlled trial. BMJ. 2007; 335(7617):436. [PubMed: 17699546]
12. Coeytaux RR, Kaufman JS, Kaptchuk tj, et al. A randomized, controlled trial of acupuncture for chronic daily headache. Headache. 2005; 45:1113–1123. [PubMed: 16178942]
13. Melchart D, Streng A, Hoppe A, et al. Acupuncture in patients with tension-type headache: randomized controlled trial. BMJ. 2005; 331(7513):376–382. [PubMed: 16055451]
14. Soderberg E, Carlsson J, Stener-Victorin E. Chronic tension-type headache treated with acupuncture, physical training and relaxation training. Between-group differences. Cephalalgia. 2006 Nov; 26(11):1320–1329. [PubMed: 17059439]
15. Endres HG, Bowing G, Diener HC, et al. Acupuncture for tension-type headache a multicentre, sham-controlled, patient-and observer-blinded, randomized trial. J Headache Pain. 2007 Oct 8.(5): 306–314. [PubMed: 17955168]
16. Vickers AJ, Rees RW, Zollman CE, et al. Acupuncture for chronic headache in primary care: large, pragmatic, randomized trial. BMJ. 2004; 328(7442):744. [PubMed: 15023828]
17. Alecrim-Andrade J, Maciel-Junior JA, Carne X, et al. Acupuncture in migraine prevention: a randomized sham controlled study with 6-months post-treatment follow-up. Clin J Pain. 2008; 24(2):98–105. [PubMed: 18209514]
18. Streng A, Linde K, Hoppe A, et al. Effectiveness and tolerability of acupuncture compared with metoprolol in migraine prophylaxis. Headache. 2006; 46:1492–1502. [PubMed: 17115982]
19. Facco E, Liguori A, Petti F, et al. Traditional acupuncture in migraine: a controlled, randomized study. Headache. 2008; 48(3):398–407. [PubMed: 17868354]
20. Linde K, Streng A, Jurgens S, et al. Acupuncture for Patients with Migraine: a randomized controlled trial. JAMA. 2005; 293(17):2118–2125. [PubMed: 15870415]
21. Diener HC, Konfeld K, Boweing G, et al. Efficacy of acupuncture for the prophylaxis of migraine a multicentre randomized controlled clinical trial. Lancet Neurol. 2006; 5(4):310–316. [PubMed: 16545747]
22. Alecrim-Andrade J, Maciel-Junior JA, Cladellas XC, et al. Acupuncture in migraine prophylaxis: a randomized sham-controlled trial. Cephalagia. 2006; 26(5):520–529.
23. Elden H, Hagberg H, Olsen MF, et al. Regression of pelvic girdle pain after delivery: follow-up of a randomised single blind controlled trial with different treatment modalities. Acta Obstet Gynecol Scand. 2008; 87(2):201–208. [PubMed: 18231889]
24. Lee SW, Liong ML, Yewn KH, et al. Acupuncture versus sham acupuncture for chronic prostatitis/chronic pelvic pain. Am J Med. 2008; 121(1):79.e1–79.e7. [PubMed: 18187077]
25. Lund I, Lundeberg T, Lonnberg L, et al. Decrease of pregnant women's pelvic pain after acupuncture a randomized controlled single-blind study. Acta Obstet Gynecol Scand. 2006; 85(1): 12–19. [PubMed: 16521674]
26. Kvorning N, Holmberg C, Grennert L, et al. Acupuncture relieves pelvic and low-back pain in late pregnancy. Acta Obstet Gynecol Scand. 2004; 83:246–250. [PubMed: 14995919]
27. Selmer-olsen T, Lydersen S, Morkved S. Does acupuncture used in nulliparous women reduce time from prelabour rupture of membranes at term to active phase of labour?. A randomised controlled trial. Acta Obstet Gynecol Scand. 2007; 86(12):1447–1452. [PubMed: 17963050]
28. Hantoushzadeh S, Alhussenini N, Lebaschi AH. The effects of acupuncture during labour on nulliparous women: a randomised controlled trial. Aust N Z J Obstet Gynaecol. 2007; 47(1):26–30. [PubMed: 17261096]
29. Gaudernack LC, Forbord S, Hole E, et al. Acupuncture administered after spontaneous rupture of membranes at term significantly reduces the length of birth and use of oxytocin. A randomized controlled trial. Acta Obstet Gynecol Scand. 2006; 85(11):1348–1353. [PubMed: 17091416]
30. Ziaei S, Hajipour L. Effect of acupuncture on labor. Int J Gynaecol Obstet. 2006; 92(1):71–72. [PubMed: 16253255]

31. Martensson L, Stener-Victorin E, Wallin G. Acupuncture versus subcutaneous injections of sterile water as treatment for labour pain. Acta Obstet Gynecol Scand. 2007; 87(2):171–177. [PubMed: 18231884]
32. Gioia L, Cabrini L, Gemma M, et al. Sedative effect of acupuncture during cataract surgery prospective randomized double-blind study. J Cataract Refract Surg. 2006; 32(11):1951–1954. [PubMed: 17081902]
33. Pohodenko-Chudakova IO. Acupuncture analgesia and its application in cranio-maxillofacial surgical procedures. J Craniomaxillofac Surg. 2005; 33(2):118–122. [PubMed: 15804591]
34. Smith MJ, Tong HC. Manual acupuncture for analgesia during electromyography: a pilot study. Arch Phys Med Rehabil. 2005 Sep; 86(9):1741–1744. [PubMed: 16181936]
35. 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(7569):623. [PubMed: 16980316]
36. Kvist LJ, Larsson BW, Hall-Lord ML, et al. Effects of acupuncture and care interventions on the outcome of inflammatory symptoms of the breast in lactating women. Int Nurs Rev. 2004; 51(1): 56–64. [PubMed: 14764015]
37. Brinkhaus B, Witt CM, Jena S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. Arch Intern Med. 2006; 166(4):450–457. [PubMed: 16505266]
38. Haake M, Muller HH, Schade-Brittinger C, et al. German acupuncture trials (GERAC) for chronic low back pain. Arch Intern Med. 2007; 167(17):1892–1898. [PubMed: 17893311]
39. Witt CM, Jena S, Selim D, et al. Pragmatic randomized trial evaluating the clinical and economic effectiveness of acupuncture for chronic low back pain. Am J Epidemiol. 2006; 164(5):487–496. [PubMed: 16798792]
40. Inoue M, Kitakoji H, Ishizaki N, et al. Relief of low back pain immediately after acupuncture treatment – a randomized, placebo controlled trial. Acupunct Med. 2006; 24(3):103–108. [PubMed: 17013356]
41. Guerreiro da Silva JB, Nakamura MU, Cordeiro JA, et al. Acupuncture for low back pain in pregnancy – a prospective, quasi- randomized, controlled study. Acupunct Med. 2004; 22(2):60–67. [PubMed: 15253580]
42. Kerr DP, Walsh DM, Baxter D. Acupuncture in the management of chronic low back pain: a blinded randomized controlled trial. Clin J Pain. 2003; 19(6):364–370. [PubMed: 14600536]
43. Vas J, Ortega C, Olmo V, et al. Single-point acupuncture and physiotherapy for the treatment of painful shoulder: a multicentre randomized controlled trial. Rheumatology. 2008; 47:887–893. [PubMed: 18403402]
44. Witt CM, Jena S, Brinkhaus B, et al. Acupuncture for patients with chronic neck pain. Pain. 2006; 125(1–2):98–106. [PubMed: 16781068]
45. Dyson-Hudson TA, Kadar P, LaFountaine M, et al. Acupuncture for chronic shoulder pain in persons with spinal cord injury: a small-scale clinical trial. Arch Phys Med Rehabil. 2007; 88(10): 1276–1283. [PubMed: 17908569]
46. Goldman RH, Stason WB, Park SK, et al. Acupuncture for treatment of persistent arm pain due to repetitive use: a randomized controlled clinical trial. Clin J Pain. 2008; 24:211–218. [PubMed: 18287826]
47. Vas J, Perea-Milla E, Mendez C, et al. Efficacy and safety of acupuncture for chronic uncomplicated neck pain: a randomized controlled study. Pain. 2006; 126(1–2):245–255. [PubMed: 16934402]
48. Vincent A, Barton DL, Mandrekar JN, et al. Acupuncture for hot flashes: a randomized, sham-controlled clinical study. Menopause. 2007; 14(1):45–52. [PubMed: 17019380]
49. Deng J, Vichers AJ, Yeung KS, et al. Randomized, controlled trial of acupuncture for the treatment of hot flashes in breast cancer patients. J Clin Oncol. 2007; 25(35):5584–5590. [PubMed: 18065731]
50. Huang MI, Yael Nir, Bertha Chen, et al. A randomized controlled pilot study of acupuncture for postmenopausal hot flashes: effect on nocturnal hot flashes and sleep quality. Fertil Steril. 2006; 86(3):700–710. [PubMed: 16952511]

51. Cohen SM, Rousseau ME, Carey BL. Can acupuncture ease the symptoms of menopause? HolisT Nurs Pract. 2003; 17(6):295–299. [PubMed: 14650571]
52. Yael, Nir; Huang, MI.; Schnyer, R., et al. Acupuncture for postmenopausal hot flashes. Maturita. 2007; 56(4):383–395.
53. Paraskeva A, Melemeni A, Petropoulos G, et al. Needling of the extra 1 point decreases BIS values and preoperative anxiety. Am J Chin Med. 2004; 32(5):789–794. [PubMed: 15633813]
54. Westergaard LG, Mao Q, Krogslund M, et al. acupuncture on the day of embryo transfer significantly improves the reproductive outcome in infertile women: a prospective, randomized trial. Fertil Seril. 2006; 85(5):1341–1346.
55. Dieterle S, Ying G, Hatzmann W, et al. Effect of acupuncture on the outcome of in vitro fertilization and intracytoplasmic sperm injection: a randomized, prospective, controlled clinical study. Fertil Steril. 2006; 85(5):1347–1351. [PubMed: 16616748]
56. Smith C, Coyle M, Norman RJ. Influence of acupuncture stimulation on pregnancy rates for women undergoing embryo transfer. Fertil Steril. 2006; 85(5):1352–1358. [PubMed: 16600225]
57. Allen JJ, Schnyer RN, Chambers AS, et al. Acupuncture for depression: a randomized controlled trial. J Clin Psychiatry. 2006; 67(11):1665–1673. [PubMed: 17196044]
58. Manber R, Schnyer RN, Allen JJ, et al. Acupuncture: a promising treatment for depression during pregnancy. J Affect Disord. 2004; 83(1):89–95. [PubMed: 15546651]
59. Xue CC, An X, Cheung TP, et al. Acupuncture for persistent allergic rhinitis: a randomized, sham-controlled trial. Med J Aust. 2007; 187(6):337–341. [PubMed: 17874980]
60. Ng DK, Chow P, Ming S, et al. A double-blind, randomized, placebo-controlled trial of acupuncture for the treatment of childhood persistent allergic rhinitis. Pediatrics. 2004; 114:1242–1247. [PubMed: 15520102]
61. Magnusson AL, Svensson RE, Leirvik C, et al. The effect of acupuncture on allergic rhinitis: a randomized controlled clinical trial. Am J Chin Med. 2004; 32(1):105–115. [PubMed: 15154290]
62. Reynolds JA, Bland JM, MacPherson H. Acupuncture for irritable bowel syndrome – an exploratory randomized controlled trial. Acupunct Med. 2008; 26(1):8–16. [PubMed: 18356794]
63. Schneider A, Enck P, Streitberger K, et al. Acupuncture treatment in irritable bowel syndrome. Gut. 2006; 55(5):649–654. [PubMed: 16150852]
64. Forbes A, Jackson S, Walter C, et al. Acupuncture for irritable bowel syndrome: a blinded placebo – controlled trial. World J Gastroenterol. 2005; 11(26):4040–4044. [PubMed: 15996029]
65. Assefi NP, Sherman KJ, Jacobsen C, et al. A randomized clinical trial of acupuncture compared with sham acupuncture in fibromyalgia. Ann Intern Med. 2005; 143:10–19. [PubMed: 15998750]
66. Harris RE, Tian X, Williams DA, et al. Treatment of fibromyalgia with formula acupuncture: investigation of needle placement, needle stimulation, and treatment frequency. J Altern Complement Med. 2005; 11(4):663–671. [PubMed: 16131290]
67. Guerreiro da Silva JB, Nakamura MU, Cordeiro JA, et al. Acupuncture for insomnia in pregnancy – a prospective, quasi-randomised, controlled study. Acupunct Med. 2005; 23(2):47–51. [PubMed: 16025784]
68. Smith P, Mosscrop D, Davies S, et al. The efficacy of acupuncture in the treatment of tempromandibular joint myofascial pain: a randomized controlled trial. J Dent. 2007; 35(3):259–267. [PubMed: 17095133]
69. Flachskampf FA, Gallasch J, Gefeller O, et al. Randomized trial of acupuncture to lower blood pressure. Circulation. 2007; (115):3121–3129. [PubMed: 17548730]
70. Macklin EA, Wayne PM, Kalish LA, et al. Stop hypertension with the acupuncture research program (SHARP) results of a randomized, controlled clinical trial. Hypertension. 2006; 48(5): 838–845. [PubMed: 17015784]
71. Gronlund MA, Stenevi U, Lundeberg T. Acupuncture treatment in patients with keratoconjunctivitis sicca: a pilot study. Acta Ophthalmol Scand. 2004; 82(3):283–290. [PubMed: 15115449]
72. Tseng KL, Liu HJ, Tso KY, et al. A clinical study of acupuncture and SSP (Silver Spike Point) electro-therapy for dry eye syndrome. Am J Chin Med. 2006; 34(2):197–206. [PubMed: 16552832]

73. Alexander DN, Cen S, Sullivan KJ, et al. Effects of acupuncture treatment on poststroke motor recovery and physical function: a pilot study. Neurorehabil Neural Repair. 2004; 18(4):259–267. [PubMed: 15537996]
74. Park J, White AR, James MA, et al. Acupuncture for subacute stroke rehabilitation: a sham-controlled, subject- and assessor-blind, randomized trial. Arch Intern Med. 2005; 165:2026–2031. [PubMed: 16186474]
75. Streitberger K, Diefenbacher M, Bauer A, et al. Acupuncture compared to placebo-acupuncture for postoperative nausea and vomiting prophylaxis: a randomized placebo-controlled patient and observer blind trial. Anaesthesia. 2004; 59:142–149. [PubMed: 14725517]
76. Johansson KM, Adolfsson LE, Foldevi MO. Effects of acupuncture versus ultrasound in patients with impingement syndrome: randomized clinical trial. Physical Therapy. 2005; 85(6):490–498. [PubMed: 15921471]
77. Lewith GT, Prescott P, Davis CL. Can a standardized acupuncture technique palliate disabling breathlessness: a single-blind, placebo-controlled crossover study. Chest. 2004; 125:1783–1790. [PubMed: 15136391]
78. Tam LS, Leung PC, Li TK, et al. Acupuncture in the treatment of rheumatoid arthritis: a double-blind controlled pilot study. BMC Complement Altern Med. 2007; 7:35. [PubMed: 17980044]
79. Dickman R, Schiff E, Holland A, et al. Clinical trial: acupuncture vs doubling the proton pump inhibitor dose in refractory heartburn. Aliment Pharmacol Ther. 2007; 26(10):1333–1344. [PubMed: 17875198]
80. Rosted P, Bundgaard M. Can acupuncture reduce the induction time of a local anaesthetic? – A pilot study. Acupunct Med. 2003; 21(30):92–99. [PubMed: 14620304]
81. Emmons SL, Otto L. Acupuncture for overactive bladder: a randomized controlled trial. Obstet Gynecol. 2005; 106(1):138–143. [PubMed: 15994629]
82. Alraek T, Baerheim A. The effect of prophylactic acupuncture treatment in women with recurrent cystitis: kidney patients fare better. J Altern Complement Med. 2003; 9(5):651–658. [PubMed: 14629843]
83. Elden H, Ladfors L, Olsen MF, et al. Effects of acupuncture and stabilising exercises as adjunct to standard treatment in pregnant women with pelvic girdle pain randomised single blind controlled trial. BMJ. 2005; 330(7494):761. [PubMed: 15778231]
84. Fink M, Rollnik JD, Bijak M, et al. Needle acupuncture in chronic poststoke leg spasticity. Arch Phys Med Rehabil. 2004; 85(4):667–672. [PubMed: 15083445]
85. Engelhardt PF, Daha LK, Zils T, et al. Acupuncture in the treatment of psychogenic erectile dysfunction: first results of a prospective randomized placebo-controlled study. Int J Impot Res. 2003; 15(5):343–346. [PubMed: 14562135]
86. Hollifield M, Sinclair-Lian N, Warner TD, et al. Acupuncture for posttraumatic stress disorder: a randomized controlled pilot trial. J Nerv Ment Dis. 2007; 195:504–513. [PubMed: 17568299]
87. Witt CM, Reinhold T, Brinkhaus B, et al. Acupuncture in patients with dysmenorrhea: a randomized study on clinical effectiveness and cost-effectiveness in usual care. Am J Obstet Gynecol. 2008; 198(2):166.e1–166.e8. [PubMed: 18226614]
88. Fu WB, Liang ZH, Zhu XP, et al. Analysis on the effect of acupuncture in treating cervical spondylosis with different syndrome types. Chin J Integr Med. 2009 Dec; 15(6):426–430. [PubMed: 20082247]
89. Walker EM, Rodriguez AI, Kohn B, et al. Acupuncture versus venlafaxine for the management of vasomotor symptoms in patients with hormone receptor-positive breast cancer: a randomized controlled trial. J Clin Oncol. 2010 Feb 1; 28(4):634–640. [PubMed: 20038728]
90. Shen YF, Younger J, Goddard G, et al. Randomized clinical trial of acupuncture for myofascial pain of the jaw muscles. J Orofac Pain. 2009 Fall;23(4):353–359. [PubMed: 19888488]
91. Tong Y, Jia Q, Sun Y, et al. Acupuncture in the treatment of diabetic bladder dysfunction. J Altern Complement Med. 2009 Aug; 15(8):905–909. [PubMed: 19678782]
92. Korinenko Y, Vincent A, Cutshall SM, et al. Efficacy of acupuncture in prevention of postoperative nausea in cardiac surgery patients. Ann Thorac Surg. 2009 Aug; 88(2):537–542. [PubMed: 19632409]

93. Liu SY, Hsieh CL, Wei TS, et al. Acupuncture stimulation improves balance function in stroke patients: a single-blinded controlled, randomized study. Am J Chin Med. 2009; 37(3):483–494. [PubMed: 19606509]
94. Lin ZP, Lan LW, He TY, et al. Effects of acupuncture stimulation on recovery ability of male elite basketball athletes. Am J Chin Med. 2009; 37(3):471–481. [PubMed: 19606508]
95. Asher GN, Coeytaux RR, Chen W, et al. Acupuncture to initiate labor (Acumoms 2): a randomized, sham-controlled clinical trial. J Matern Fetal Neonatal Med. 2009 Oct; 22(10):843–848. [PubMed: 19526433]
96. Lathia AT, Jung SM, Chen LX. Efficacy of acupuncture as a treatment for chronic shoulder pain. J Altern Complement Med. 2009 Jun; 15(6):613–618. [PubMed: 19489707]
97. Lembo AJ, Conboy L, Kelley JM, et al. A treatment trial of acupuncture in IBS patients. Am J Gastroenterol. 2009 Jun; 104(6):1489–1497. [PubMed: 19455132]
98. Li Y, Liang F, Yang X, et al. Acupuncture for treating acute attacks of migraine: a randomized controlled trial. Headache. 2009 Jun; 49(6):805–816. [PubMed: 19438740]
99. Cherkin DC, Sherman KJ, Avins AL, et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. Arch Intern Med. 2009 May 11; 169(9): 858–866. [PubMed: 19433697]
100. Borud EK, Alraek T, White A, et al. The Acupuncture on Hot Flushes among Menopausal Women (ACUFLASH) study, a randomized controlled trial. Menopause. 2009 May-Jun;16(3): 484–493. [PubMed: 19423996]
101. Dieterle S, Li C, Greb R, et al. A prospective randomized placebo-controlled study of the effect of acupuncture in infertile patients with severe oligoasthenozoospermia. Fertil Steril. 2009 Oct; 92(4):1340–1343. [PubMed: 19394002]
102. Ling L, Jiang XM, Xue JW, et al. Clinical study on the visceral differentiation-based acupuncture therapy for insomnia. J Tradit Chin Med. 2008 Dec; 28(4):270–273. [PubMed: 19226897]
103. Deng G, Rusch V, Vickers A, et al. Randomized controlled trial of a special acupuncture technique for pain after thoracotomy. J Thorac Cardiovasc Surg. 2008 Dec; 136(6):1464–1469. [PubMed: 19114190]
104. Suzuki M, Namura K, Ohno Y, et al. The effect of acupuncture in the treatment of chronic obstructive pulmonary disease. J Altern Complement Med. 2008 Nov; 14(9):1097–1105. [PubMed: 19055335]
105. Brinkhaus B, Witt CM, Jena S, et al. Acupuncture in patients with allergic rhinitis: a pragmatic randomized trial. Ann Allergy Asthma Immunol. 2008 Nov; 101(5):535–543. [PubMed: 19055209]
106. Hübscher M, Vogt L, Bernhörster M, et al. Effects of acupuncture on symptoms and muscle function in delayed-onset muscle soreness. J Altern Complement Med. 2008 Oct; 14(8):1011–1016. [PubMed: 18990049]
107. Smith CA, Crowther CA, Collins CT, et al. Acupuncture to induce labor: a randomized controlled trial. Obstet Gynecol. 2008 Nov; 112(5):1067–1074. [PubMed: 18978107]
108. Gottschling S, Reindl TK, Meyer S, et al. Acupuncture to alleviate chemotherapy-induced nausea and vomiting in pediatric oncology - a randomized multicenter crossover pilot trial. Klin Padiatr. 2008 Nov-Dec;220(6):365–370. [PubMed: 18949672]
109. Elden H, Fagevik-Olsen M, Ostgaard HC, et al. Acupuncture as an adjunct to standard treatment for pelvic girdle pain in pregnant women: randomised double-blinded controlled trial comparing acupuncture with non-penetrating sham acupuncture. BJOG. 2008 Dec; 115(13):1655–1668. [PubMed: 18947338]
110. So EW, Ng EH, Wong YY, et al. A randomized double blind comparison of real and placebo acupuncture in IVF treatment. Hum Reprod. 2009 Feb; 24(2):341–348. [PubMed: 18940896]
111. Hervik J, Mjåland O. Acupuncture for the treatment of hot flashes in breast cancer patients, a randomized, controlled trial. Breast Cancer Res Treat. 2009 Jul; 116(2):311–316. [PubMed: 18839306]
112. Jena S, Witt CM, Brinkhaus B, et al. Acupuncture in patients with headache. Cephalalgia. 2008 Sep; 28(9):969–979. [PubMed: 18624803]

113. Avis NE, Legault C, Coeytaux RR, et al. A randomized, controlled pilot study of acupuncture treatment for menopausal hot flashes. Menopause. 2008 Nov-Dec;15(6):1070–1078. [PubMed: 18528313]

114. Domar AD, Meshay I, Kelliher J, et al. The impact of acupuncture on in vitro fertilization outcome. Fertil Steril. 2009 Mar; 91(3):723–726. [PubMed: 18314118]

115. Brian R, Flay. Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs. Preventive Medicine. 1986:451–474. [PubMed: 3534875]

116. MacPherson H, White A, Cummings M, et al. Standards for reporting interventions in controlled trials of acupuncture: The STRICTA recommendations. Acupunct Med. 2002; 20(1):22–25. [PubMed: 11926601]

117. Schneyer RN, Allen JJB. Bridging the Gap in Complementary and Alternative Medicine Research: Manualization as a Means of Promoting Standardization and Flexibility of Treatment in Clinical Trials of Acupuncture. J Altern Complement Med. 2002; 8:623–634. [PubMed: 12470444]

118. Hedden SL, Woolson RF, Malcolm RJ. Randomization in substance abuse clinical trials. Subst Abuse Treat Prev Policy. 2006; 1(1):6. [PubMed: 16722591]

119. Lao, L.; Ezzo, J.; Hammerschlag, R.; Berman, B. Assessing clinical efficacy of acupuncture: Considerations for designing future acupuncture trials. In: Stux, G.; Hammerschlag, R., editors. Clinical Acupuncture: Scientific Basis. Berlin: Springer; 2001. p. 187-209.

120. Jadad, AR. Randomized controlled trials: A user's guide. London: BMJ Publishing Group; 1998.

121. Chow, SC.; Liu, JP. Design and Analysis of Clinical Trials: Concept and Methodologies. New York: Wiley; 1998.

122. Birch S. A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert. J Altern Complement Med. 2006; 12:303–310. [PubMed: 16646730]

123. Langevin HM, Hammerschlag R, Lao LX, et al. Controversies in acupuncture research: selection of controls and outcome measures in acupuncture clinical trials. J Altern Complement Med. 2006; 12:943–953. [PubMed: 17212566]

124. Lewith, G.; Jonas, WB.; Walach, H. Clinical Research in complementary therapies: Principles, problems and solutions. Edinburgh: Churchill Livingstone; 2002.

125. Piantadosi, S. Clinical Trials: A methodologic perspective. New York: John Wiley&Sons, Inc.; 1997.

126. Moher D, Schulz K, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomised trials. Lancet. 2001; 357:1191–1194. [PubMed: 11323066]

127. Vickers AJ. Statistical reanalysis of four recent randomized trials of acupuncture for pain using analysis of covariance. Clin J Pain. 2004; 20:319–323. [PubMed: 15322438]

128. Fitzmaurice, G.; Davidian, M.; Verbeke, G., et al. Longitudinal Data Analysis: Handbooks of Modern Statistical Methods. Taylor & Francis Group, LLC; 2009.

129. Zyzanski SJ, Flocke SA, Dickinson LM. On the nature and analysis of clustered data. Ann Fam Med. 2004; 2:199–200. [PubMed: 15209193]

130. Hrobjartsson A?Forfang E, Haahr MT, et al. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. Int J Epidemiol. 2007 Jun; 36(3):654–663. [PubMed: 17440024]

131. Lund I, Naslund J, Lundeberg T. minimal acupuncture is not a valid placebo control in randomised controlled trials of acupuncture: a physiologist's perspective. Chin Med. 2009 Jan 30.4:1. [PubMed: 19183454]

132. Lao L, Bergman S, Hamilton G, Langenberg P, et al. Evaluation of acupuncture for pain control after oral surgery. Archives of Otolaryngology: Head and Neck Surgery. 1999; 125(5):567–572. [PubMed: 10326816]

133. Berman BM, Lao L, Langenberg P, et al. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee. Ann Intern Med. 2004; 141(12):901–910. [PubMed: 15611487]

134. Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G. Longitudinal Data Analysis. Chapman & Hall/CRC; 2009.

**Table 1**

Participants' previous acupuncture experience (non-insertion sham control studies)

| Acupuncture Experience | Efficacy Studies | Effectiveness Studies | Unspecified Study | Total |
|---|---|---|---|---|
| Acupuncture naïve | 0 | 3 (27%) | 1 (25%) | 4 (15%) |
| No acupuncture in the past 6 or 12 months | 4 (36%) | 3 (27%) | 1 (25%) | 8 (31%) |
| Not reported | 7 (64%) | 5 (46%) | 2 (50%) | 14 (54%) |
| Total | 11 (100%) | 11 (100%) | 4 (100%) | 26 (100%) |

**Table 2**

Masking/blinding methods used

| Masking/Blinding Methods | Efficacy Study | Effectiveness Study | Unspecified Study | Total |
|---|---|---|---|---|
| Patients only | 15 (33%) | 9 (20%) | 7 (36%) | 31 (29%) |
| Patients + Evaluators | 14 (31%) | 9 (20%) | 2 (11%) | 25 (23%) |
| Patients + Practitioners (Double Blinding) | 3 (7%) | 2 (5%) | 0 | 5 (5%) |
| Evaluators only | 4 (9%) | 5 (11%) | 0 | 9 (8%) |
| No Masking/Blinding | 1 (2%) | 2 (5%) | 0 | 3 (3%) |
| Unclear | 8 (18%) | 17 (39%) | 10 (53%) | 35 (32%) |
| Total | 45 (100%) | 44 (100%) | 19 (100%) | 108 (100%) |

**Table 3**

Different types of sham acupuncture based on needling and point selection

| | Efficacy Studies | Effectiveness Studies | Other studies | Total |
|---|---|---|---|---|
| Non-insertion at non-acupuncture points | 7 (20%) | 6 (29%) | 2 (15%) | 15 (22%) |
| Non-insertion at true acupuncture points | 4 (11%) | 5 (24%) | 2 (15%) | 11 (16%) |
| Insertion (superficial) at non-acupuncture points | 20 (58%) | 10 (47%) | 7 (55%) | 37 (53%) |
| Insertion (superficial) at true acupuncture points | 4 (11%) | 0 | 2 (15%) | 6 (9%) |
| Total | 35 (100%) | 21 (100%) | 13 (100%) | 69 (100%) |

**Table 4**

Different Types of Primary Outcomes Reported in Results

| Primary outcomes | Efficacy studies | Effectiveness studies | Unspecified studies | Total |
|---|---|---|---|---|
| Subjective | 27 (60%) | 26 (60%) | 8 (42%) | 61 (57%) |
| Objective | 9 (20%) | 9 (20%) | 7 (37%) | 25 (23%) |
| Both Subjective and Objective | 9 (20%) | 9 (20%) | 4 (21%) | 22 (20%) |
| Total | 45 (100%) | 44 (100%) | 19 (100%) | 108 (100%) |

**Table 5**

Methods for dealing with missing data

| Methods | Efficacy Study | Effectiveness Study | Other Study | Total |
|---|---|---|---|---|
| Last observation carried forward | 11 (65%) | 13 (42%) | 2 (67%) | 26 (51%) |
| Replaced with average available data on either side of missing value | 0 | 1 (3%) | 0 | 1 (2%) |
| Replaced with mean values obtained | 1 (6%) | 2 (7%) | 0 | 3 (6%) |
| Replaced with baseline data | 1 (6%) | 5 (16%) | 1 (33%) | 7 (13%) |
| Patients with missing data counted as failures | 1 (6%) | 1 (3%) | 0 | 2 (4%) |
| Worst case method for experimental group and best case method for control group | 1 (6%) | 1 (3%) | 0 | 2 (4%) |
| Multiple imputation | 2 (11%) | 8 (26%) | 0 | 10 (20%) |
| Total | 17 (100%) | 31 (100%) | 3 (100%) | 51 (100%) |

**Table 6**

Multiple comparisons

| | Efficacy study | Effectiveness study | Other study | Total |
|---|---|---|---|---|
| More than two groups being compared | 16 | 10 | 5 | 31 |
| Multiple comparison used | 8 (50%) | 8 (80%) | 1 (25%) | 17 (55%) |
| Multiple comparison not used | 8 (50%) | 2 (20%) | 4 (75%) | 14 (45%) |
| Bonferroni test used | 3 (19%) | 4 (40%) | 0 | 7 (23%) |