



# HHS Public Access

Author manuscript

*Acupunct Med.* Author manuscript; available in PMC 2022 October 26.

Published in final edited form as:

*Acupunct Med.* 2021 August ; 39(4): 309–317. doi:10.1177/0964528420959089.

## Do the effects of acupuncture vary between acupuncturists? Analysis of the Acupuncture Trialists' Collaboration individual patient data meta-analysis

**Andrew J. Vickers, PhD, DPhil,**

Memorial Sloan Kettering Cancer Center, New York, NY

**Emily A. Vertosick, MPH,**

Memorial Sloan Kettering Cancer Center, New York, NY

**George Lewith, MD,**

University of Southampton, Southampton, UK (deceased)

**Hugh MacPherson, PhD,**

University of York, York, UK

**Nadine E. Foster, DPhil,**

Keele University, Staffordshire, UK

**Karen J. Sherman, PhD,**

Kaiser Permanente Washington Health Research Institute, Seattle, WA

**Dominik Irnich, MD,**

Ludwig-Maximilians-Universität München (LMU Munich), Munich, Germany

**Claudia M. Witt, MD,**

University of Zurich and University Hospital Zurich, Zurich, Switzerland; Charite-Universitätsmedizin, Berlin, Germany; University of Maryland School of Medicine, Baltimore, Maryland

**Klaus Linde, MD,**

Technical University Munich, Germany

**Acupuncture Trialists' Collaboration**

### Abstract

---

**Corresponding author:** Andrew J. Vickers, Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, New York, NY 10017 USA, vickersa@mskcc.org, (P): 646-888-8233.

Authors' contributions

The study was conceived by AV, GL, CW and KL. AV was responsible for the overall study design; EV for the systematic review; GL and HM with respect to acupuncture analyses; CW, NF, KS and KL with respect to clinical trial methodology and meta-analysis. Statistical analyses were conducted by EV. The first draft of the manuscript was written by AV and EV. All authors gave comments on early drafts and all (other than GL, see below) approved the final version of the manuscript. AV had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. GL is deceased.

**Disclosures:** The authors have no competing interests and nothing to disclose.

**Objectives**—The degree to which the effects of acupuncture treatment vary between acupuncturists is unknown. We used a large individual patient dataset of trials of acupuncture for chronic pain to assess practitioner heterogeneity.

**Methods**—Individual patient data linked to identifiable acupuncturists were drawn from a dataset of 39 high-quality trials of acupuncture, where the comparators were either sham acupuncture or non-acupuncture control, such as standard care or waitlist. Heterogeneity among acupuncturists was assessed by meta-analysis.

**Results**—A total of 1,206 acupuncturists in 13 trials were included. Statistically significant heterogeneity was found in trials with sham-control groups ( $p = 0.003$ ) and non-acupuncture control groups ( $p = 0.001$ ). However, the degree of heterogeneity was very small, with the observed distribution of treatment effects virtually overlapping that expected by chance. For instance, for non-acupuncture-controlled trials, the proportion of acupuncturists with effect sizes half a standard deviation greater or less than average was expected to be 34%, but was observed to be 37%. A limitation is that the trials included a relatively limited range of acupuncturists, mainly physician acupuncturists.

**Discussion**—Although differences in effects between acupuncturists were greater than expected by chance, the degree of variation was small. This suggests that most chronic pain patients in clinical practice would have similar results to those reported in high-quality trials; comparably, we did not find evidence to suggest that greater standardization of acupuncture practice would improve outcomes. Further research needs to be conducted exploring variability using a sample of acupuncturists with a broader range of practice styles, training and experience.

### Keywords

acupuncture; chronic pain; practitioner variability; acupuncturist variability

---

## INTRODUCTION

Some medical treatments require skill to administer, leading to the possibility that treatment effects vary between practitioners. Heterogeneity among practitioners is likely to involve a range of characteristics, including different educational and learning experiences, unique repertoires of clinical experience, and the utilisation of different interpersonal and technical skills during treatment.

Heterogeneity of outcomes among practitioners, defined as variability over and above that expected by chance, has a number of implications for research and clinical practice. It affects the interpretation of research results because, in the presence of heterogeneity, a referring clinician may be sceptical that the average results reported in the relevant literature will be obtained by a local practitioner. Heterogeneity would also suggest that changes in practice or regulation might be required in order to ensure more consistent outcomes.

Analysis of heterogeneity of outcomes among practitioners is commonly ignored in the analysis of clinical trials.(1) Methodologists are starting to explore the implications of heterogeneity in the context of surgical procedures,(2) primary care services,(3) physician-led diabetes care,(4) the teaching of behavioural interventions,(5) and psychotherapists in

clinical practice.(6) In a retrospective analysis using data from four randomised controlled trial datasets, small but significant variations in outcome between health professionals were found for nurses treating leg ulcers, acupuncturists treating low back pain, physiotherapists providing pre- and post-operative care associated with total knee replacement, and homeopaths treating chronic fatigue syndrome.(7) Therapists have also been identified as a source of variability within psychotherapy outcome studies.(8)

We have recently updated an individual patient data meta-analysis of acupuncture for chronic pain.(9) We found acupuncture to lead to statistically significant improvements in pain when compared to either sham acupuncture or non-acupuncture controls. However, it is not clear to what extent acupuncture effect sizes are affected by practitioner variability. Acupuncture requires skill to administer, but it is unclear whether effects vary greatly between acupuncturists, or whether acupuncturists have approximately similar results. This is of particular interest because acupuncture is provided by a wide range of professionals, including physicians, nurses and physiotherapists with additional training in acupuncture, and specialist acupuncturists with no other healthcare professional qualifications. In addition, there is wide variation in the length of training of these practitioners in acupuncture from short courses of several days, to full Masters' degrees. An additional reason to suspect variation in acupuncture outcomes is that there are several different styles of acupuncture based on different underlying theories.(10) Differences in the way that practitioners manipulate acupuncture needles have also been documented.(11) Here, we take the opportunity of our large individual patient dataset to further explore heterogeneity in treatment effects among different acupuncturists.

## METHODS

### Included Trials

Trials included in these analyses were identified through a systematic literature review that has been previously described.(12) The search included trials of acupuncture for chronic pain published prior to December 2015 and included only high-quality trials where allocation concealment was determined unambiguously to be adequate. Eligible pain types were non-specific back or neck pain, shoulder pain, chronic headache or osteoarthritis – with the additional criterion that the current episode of pain must be of at least four weeks duration for musculoskeletal disorders. This search resulted in the identification of 44 trials.

### Data Acquisition

Individual patient data were obtained from 39 trials. Data on the trial-level characteristics of the acupuncture intervention were obtained directly from trialists. Twenty-six trials had a sham acupuncture control group, and twenty-five trials had a non-acupuncture control group, twelve trials being three arm trials. This study was approved by the Institutional Review Board at Memorial Sloan Kettering Cancer Center (14–113). Informed consent was not necessary as this was a secondary use of deidentified data.

## Outcome

The primary outcome used for this analysis was the primary outcome defined by the study authors for each study. Where multiple criteria were considered in the primary outcome (e.g. a response defined as either a 33% reduction in pain or a 50% reduction in pain medication) or if the primary outcome was inherently categorical, we used a continuous measure of pain measured at the same time point as the original primary outcome. To make the various outcome measurements comparable between different trials, the primary endpoint of each was standardized by dividing by pooled standard deviation.

## Eligible Trials

To be included in the analysis, trialists had to identify the acupuncturist who treated each patient in the acupuncture group, and in sham-controlled trials, identify the acupuncturist who treated each patient in the sham acupuncture group. Most non-acupuncture control trials identified an acupuncturist for control group patients, that is, the acupuncturist whom a patient would have seen had he or she been allocated to acupuncture. This was typically a function of geographic location, with the nearest acupuncturist identified for each patient before randomization. For two trials,(13, 14) acupuncturist information was not provided for the non-acupuncture group, so acupuncturists were randomly assigned to each of the control group patients from the same trial. This approach was deemed reasonable as there was no contact between acupuncturists and control group patients in these two trials and therefore the distribution of control group patient outcomes will be random with respect to acupuncturist. Five sham-controlled trials had practitioner data available and were included in the analysis.(15–19) The other 21 trials with sham control arms did not provide practitioner data and were excluded. (13, 20–39). The Hinman 2014 trial was excluded from the analysis of sham-controlled trials due to a sham arm that was not known to be credible. (13) Twelve trials were included in the analysis of non-acupuncture-controlled trials;(13–17, 19, 40–45) 13 trials were excluded.(20, 23, 25, 26, 33–35, 46–51)

## Statistical Methods

In these analyses, we examined whether the effects of acupuncture for chronic pain varied among acupuncturists. We took two separate but complementary statistical approaches. In the first, a trial-level analysis, we examined whether effect sizes differed between acupuncturists within each trial separately and combined the results on acupuncturist heterogeneity from each trial into a meta-analysis. In the second approach, an acupuncturist-level analysis, data from acupuncturists in all trials was combined into a single analysis.

**Trial-level analysis**—Acupuncturists were eligible for inclusion in the trial-level analysis if they treated two or more patients in the acupuncture group. To determine whether the effects of acupuncture varied between acupuncturists within each trial separately, a mixed effects model with both a random intercept and a random slope was created for each trial to predict pain after treatment, with baseline pain and treatment group as fixed effects. Acupuncturists were included as a random effect, which assumes that the acupuncturists included in the analysis were drawn from a larger theoretical population of acupuncturists.

By estimating the distribution of this theoretical population, we can investigate between-acupuncturist variability.

For these models, the coefficient for treatment is the difference in pain change from baseline between the control group and the acupuncture group, with an increase in slope representing a greater reduction in pain in the acupuncture group compared to the control group. Further modeling details are available in the supplementary material.

We meta-analyzed the random-effects coefficient from these models to test whether there was a significant random effect of acupuncturist among all trials. In some trials, the models did not converge. These models were not included in the meta-analysis, but were analyzed separately (see supplementary materials). We compared the effect of acupuncturist in both sham and non-acupuncture-controlled trials to our updated meta-analysis, to determine whether the overall acupuncturist effect was consistent with our estimate of the effect of acupuncture. Among sham-controlled trials, we also compared these results to the results from the original meta-analysis when three outlying trials(36–38) were excluded, since these three trials were also excluded from the acupuncturist-specific analysis.

**Acupuncturist-level analysis**—In the second approach, we then combined all data into a single analysis. A linear regression model was created for each acupuncturist that predicted outcome after treatment, and was adjusted for baseline pain, treatment group, and any variables used to stratify randomization in the original trial. For the linear regression models to converge we excluded acupuncturists who treated fewer than 3 patients in either group. The regression coefficient for treatment group and its standard error were then entered into a meta-analysis.

Since some variation among acupuncturists is expected by chance, we compared that expected variation with the variation we observed among the acupuncturists in our sample. Since it is difficult to interpret random effects coefficients and the forest plot for the meta-analysis by acupuncturist contains too many acupuncturists to be useful for visualizing the distribution of acupuncturist effects, we instead created histograms comparing the observed distribution of effect sizes to an expected distribution of effect sizes. We overlaid two histograms – a histogram of the distribution of acupuncturist effect size in our sample, and a histogram of the distribution of acupuncturist effect size that would be expected by chance, adapting the method used by Bianco et al to create a similar graph of surgical outcomes by surgeon (see Figure 1).(52) Further details of the permutation analysis are available in the supplementary materials.

As a sensitivity analysis, we repeated these analyses after combining trials that shared acupuncturists. There were two groups of trials conducted in Germany (the ARC trials and the ART trials) in which the same acupuncturists may have treated patients with different pain types in multiple trials. The ARC trials studied the effects of acupuncture on neck pain, headache, knee and hip osteoarthritis, and lower back pain.(40–43) All ARC trials reported a clinical relevant improvement in pain for acupuncture patients compared to control patients that received only usual care. The ART trials studied acupuncture for knee osteoarthritis and lower back pain(15, 19). Both trials reported a significant improvement in

pain in acupuncture patients at 8 weeks after baseline compared to non-acupuncture control; differences between acupuncture and sham were significant for osteoarthritis but not for low back pain. Since each group of trials shared acupuncturists and used the same basic principles (see supplementary materials), we performed a sensitivity analysis in which we combined the data from all four ARC trials and from both ART trials into one ARC trial dataset and one ART trial dataset, so that each acupuncturist was associated with all patients that he or she had treated, regardless of pain type.

In the German trials, including the ART and ARC trials, acupuncture was provided by physicians, whereas most other included trials involved non-physician acupuncturists. As a sensitivity analysis, we repeated the analyses excluding 8 trials where acupuncture was provided by physicians.(13, 15, 16, 19, 40–43)

All analyses were conducted using Stata 13.0 (StataCorp, College Station, TX).

## RESULTS

### Included trials

There were 25 trials that had eligible sham acupuncture control groups; however, only five of these trials reported acupuncturist data (Witt 2005 (19), Suarez-Almazor 2010(17) and White 2012(18) for osteoarthritis, Brinkhaus 2006 for lower back pain(15) and Diener 2006 for migraine(16)), with 213 total acupuncturists, including 9 non-physician acupuncturists from two trials (17, 18) (Table 1). The differences in the number of acupuncturists between the trial- and acupuncturist-level analysis is due to the more stringent criterion for the minimum number of patients in the latter analysis. For the sensitivity analysis, the ART(19) osteoarthritis trial was combined with the ART(15) low back pain trial. The analyses were not repeated in sham-controlled trials when excluding physician-acupuncturists as there were only two eligible trials with 9 acupuncturists.

There were 25 trials with non-acupuncture control groups, and 12 of these trials reported data on acupuncturists (Thomas 2006(14), Witt 2006(43) and Brinkhaus 2006(15) for lower back pain, Witt 2005(19), Witt 2006(42), Suarez-Almazor 2010 (17) and Hinman 2014(13) for osteoarthritis, Diener 2006(16) for migraine, Vickers 2004(44) and Jena 2008(40) for headache, and MacPherson 2015(45) and Witt 2006(41) for neck pain), with 4,961 total acupuncturists, including, 43 non-physician acupuncturists from four trials(14, 17, 44, 45). For the sensitivity analysis, the German-based ARC trials of neck pain(41), osteoarthritis(42), headache(40) and low back pain(43) trials were combined. The ART osteoarthritis(19) and ART low back pain(15) trials were also combined. When combining trials for the sensitivity analysis, more acupuncturists became eligible for inclusion in the acupuncturist-level analysis, since some acupuncturists treated patients in more than one trial (Table 1). Those acupuncturists who did not treat at least three acupuncture patients and three control patients in one individual trial may have treated at least three patients in each group when including patients from multiple trials treated by the same acupuncturist. When excluding trials where acupuncture was provided by non-physicians, there were four trials with 43 acupuncturists and 943 patients included in the analysis.(14, 17, 44, 45).

## Trial-level analysis

For the meta-analysis of random effect statistics from each trial calculated separately, there was a significant random effect of acupuncturist, that is, statistically significant acupuncturist heterogeneity, in both the sham-controlled trials (N = 4 trials, 126 acupuncturists and 1,136 patients, overall random effects statistic 0.17, 95% CI 0.07, 0.26,  $p = 0.001$ ) and those with non-acupuncture control (N = 9 trials, 1,175 acupuncturists, and 5,662 patients, overall random effects statistic 0.08, 95% CI 0.03, 0.12,  $p = 0.001$ ). The Suarez-Almazor 2010(17) osteoarthritis trial was excluded from the meta-analysis of random effects statistics in sham-controlled trials and the Thomas 2006(14) low back pain, Vickers 2004(44) headache and Hinman 2014(13) osteoarthritis trials were excluded from the meta-analysis in non-acupuncture-controlled trials since the random effects models did not converge.

We then combined trials that shared acupuncturists so that all patients treated by the same acupuncturist were analyzed together regardless of the trial in which they originally participated. There were three sham-controlled trials included in this sensitivity analysis: the Diener 2006(16) migraine trial, the White 2012(18) osteoarthritis trial and the combined ART trial (Witt 2005(19) and Brinkhaus 2006(15)). After repeating the meta-analysis for sham-controlled trials, we found similar results (N = 3 trials, 108 acupuncturists and 1,146 patients, overall random effects statistic 0.13, 95% CI 0.05, 0.22,  $p = 0.003$ ). Out of the original nine trials included in the random effects meta-analysis for trials without sham control, the four ARC trials(40–43) were combined into one joint ARC trial, and the two ART trials(15, 19) were combined into one joint ART trial. The ARC trial and the ART trial were included in a meta-analysis with the Diener 2006(16) migraine trial, the MacPherson 2015 neck pain trial(45) and the Suarez-Almazor 2010(17) osteoarthritis trial. Results showed a greater variation among acupuncturists in trials without acupuncture control when the ART and ARC trials were combined, compared to kept separate (N = 5 trials, 1,148 acupuncturists and 7,733 patients, overall random effects statistic 0.12, 95% CI 0.08, 0.16,  $p < 0.0001$ ). This analysis was not performed when excluding physician-acupuncturists as there were too few eligible trials for which the random effects model converged.

## Acupuncturist-level analysis

We first compared the effect size from this meta-analysis, which included only a subset of patients, to the effect size of acupuncture from the updated meta-analysis.(9) For the sham-controlled trials, effect sizes were slightly smaller in both analyses (N = 65 acupuncturists, effect size 0.13; combining trials that shared acupuncturists, N = 63 acupuncturists, effect size 0.12) compared to the effect sizes in the original meta-analysis comparing acupuncture to sham (between 0.16 and 0.30, depending on pain type). For trials with a non-acupuncture control, effect sizes in the full meta-analysis and in the subset of trials that provided individual acupuncturist data were similar: N = 226 acupuncturists, effect size 0.45 and N = 380 acupuncturists, effect size 0.50 when combining trials that shared acupuncturists, versus effect sizes of 0.44 to 0.63 from the updated analysis depending on pain type. As the effect size for the subset included in this analysis is slightly smaller than in the main meta-analysis, this suggests that our estimate of variability between acupuncturists may be a slight underestimate.

In both sham-controlled and non-acupuncture-controlled trials, there was significant heterogeneity among acupuncturists, meaning that there was a significant difference in effect between acupuncturists that was not due solely to chance ( $p < 0.0001$  for both sham-controlled and non-acupuncture-controlled trials). After combining acupuncturists who participated in multiple trials, heterogeneity remained highly significant ( $p < 0.0001$  for both sham-controlled and non-acupuncture-controlled trials).

There were 226 total acupuncturists who treated patients in the non-acupuncture-controlled trials included in the meta-analysis. As seen in Figure 2, the dispersion of effect sizes in non-acupuncture-controlled trials was statistically greater than expected; differences are small, with the observed and expected distributions mostly overlapping. When excluding physician-acupuncturists, there were 30 acupuncturists in this analysis of non-acupuncture-controlled trials. Significant heterogeneity was seen in this subset of acupuncturists ( $p=0.047$ ), although the differences between acupuncturists were small and results consistent with the main analysis (Supplementary Figure 1).

Among sham-controlled trials (Figure 3), there appear to be a wider variation in values and less overlap between our observed distribution and the expected distribution. However, the number of acupuncturists treating patients in sham-controlled trials was small ( $N = 65$ ). As a result, some of the bars in Figure 3, such as that at  $-2$  SD, represent only one acupuncturist.

A visual comparison with Figure 1, which represents variation in complication rates between surgeons(52), indicates that while there is significant variation among acupuncturists, the magnitude of this variation is relatively minor. For instance, in the analysis of trials with non-acupuncture controls, we expected by chance that 15% of acupuncturists would have an effect of zero or less, half a standard deviation less than average; the observed number was only very slightly higher at 18%. For an effect size of 1 or more, half a standard deviation more than average, the number of acupuncturists both expected and observed was 19%. Combining these, the proportion of acupuncturists with effect sizes half a standard deviation greater or lesser than average was expected to be 34%, but was observed to be 37%. For effect sizes different from the average by one standard deviation or more, rates of expected and observed were 7% vs. 10%. For the sham-controlled trials, the expected and observed proportion of acupuncturists with effect sizes more than half a standard deviation greater or lesser than that expected was 34% vs. 48%; for effect sizes different from the average by one standard deviation or more, rates of expected and observed were 9% vs. 15%.

## DISCUSSION

We used individual patient data from a meta-analysis of high-quality trials of acupuncture for chronic pain to conduct two separate analyses addressing the question of practitioner heterogeneity of outcome. We found that differences in effect sizes among acupuncturists were greater than would be expected by chance. This effect was found for trials with sham control, as well as those with non-acupuncture control groups such as usual care. As such, differences in outcome likely include differences specific to acupuncture technique, rather than just good “bedside manner” and other non-specific aspects of clinical care.



Importantly, however, the degree of variation is relatively small. The distribution of observed acupuncturist-specific effect sizes was virtually overlapping with the distribution that would be expected by chance. This is in some distinction with, for instance, studies on variation among surgeons, where a large proportion of surgeons have outcomes very much better or worse than expected.(53, 54) Our results complement a prior study from the Acupuncture Trialists' Collaboration that did not find important modifying effects on acupuncture outcomes for characteristics such as the number or duration of acupuncture sessions, the number of needles used, the age or sex of the practitioner, acupuncture style (Western vs. traditional Chinese) or point prescription approach (fixed, flexible or individualized).(55) The major clinical implication is that chronic pain patients should be expected to have similar results to those reported in high-quality trials. A second implication is that our findings do not provide support for greater standardization of the methods and techniques that are used within routine acupuncture provided to patients with chronic pain.

The relatively limited degree of outcome heterogeneity we found in the sham-controlled trials might be seen as predictable from consideration of the effect size of acupuncture. In the meta-analysis, we reported an effect size of acupuncture compared to sham of approximately 0.2 standard deviations.(9) Sham acupuncture involves techniques such as needles inserted to the wrong depth, and/or away from acupuncture points. If the difference between good acupuncture technique and a grossly inadequate acupuncture technique - sham acupuncture - is fairly modest, then we would not expect large differences in outcome between qualified acupuncturists based on slight variations in their technique.

The authors of the Thomas 2006 trial on acupuncture for low back pain, which was excluded from the comparison in this sub-study of acupuncture vs. non-acupuncture controls because of lack of convergence of the statistical model, have also explored outcome heterogeneity among practitioners.(56) This was assessed by comparing outcomes for the six acupuncturists in the trial, each of whom treated a minimum of 15 patients. The analysis was a nested ANCOVA, whereby acupuncturists were nested within the acupuncture group and baseline pain score was the covariate. In a result that was broadly consistent with the findings of this sub-study, it was concluded that there was no evidence of any important difference in patient outcomes between acupuncturists. Previous regression analyses of the above mentioned ARC trials that evaluated acupuncture in addition to usual care included 9990 patients treated by 2781 physicians and came to the conclusion that physician characteristics such as training and experience did not influence patients' outcome after acupuncture.(57)

The major limitation of our study is that we included a relatively limited range of acupuncturists, and our sample is unlikely to be fully representative of those found in routine acupuncture practice. Our data-set is dominated by the trials from Germany, in which provision of acupuncture was by physicians with a minimum of 180 hours training, although most had far more training. It is not clear the extent to which our results reflect other types of acupuncturists (non-physician), with longer periods of training, from other countries (e.g. from China), practicing other styles of acupuncture (e.g. both in terms of technique and in terms of spending more time with patients, and recommending more self-care). At

least theoretically, these variations in practice could be associated with more heterogeneous outcomes.

Moreover, trialists often actively selected acupuncturists based on reputation or clinical experience. For instance, in the UK National Health Service headache trial(44), acupuncturists were nominated by a professional organization. Given the high profile of the trial, it is likely that only acupuncturists who are better known within the acupuncture community would be nominated. In the UK Thomas trial, acupuncturists were chosen on the basis of experience and proximity to the centre of York, the city where the trial was conducted.(14) However, although it sounds plausible that there is less heterogeneity in outcomes between better-known or more experienced acupuncturists, as mentioned above, prior research in a routine care setting with regular acupuncturists has not found an association between experience and outcome of acupuncture in comparison to non-acupuncture control(57).

## CONCLUSION

Although differences in effect sizes among acupuncturists were greater than would be expected by chance, the degree of variation was small. This suggests that most chronic pain patients in clinical practice would have similar results to those reported in high-quality trials. With respect to pain outcomes, we found no reason to suggest a need for greater standardization of acupuncture practice. However, a major limitation is that the trials included a relatively limited range of acupuncturists, which are not fully representative of those found in routine acupuncture practice. As such, further research needs to be conducted exploring practitioner variability using a wider sample of acupuncturists.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments.

The Acupuncture Trialists' Collaboration is funded by an R21 (AT004189I and an R01 (AT006794) from the National Center for Complementary and Alternative Medicine at the National Institutes of Health to Dr. Vickers) and by a grant from the Samueli Institute. Prof. MacPherson's work on this project was funded in part by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research scheme (RP-PG-0707-10186). Prof. Foster, an NIHR Senior Investigator, was supported through an NIHR Research Professorship (RP-011-015). The views expressed in this publication are those of the author(s) and not necessarily those of the National Center for Complementary and Alternative Medicine, NHS, NIHR, or the Department of Health in England. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the report.

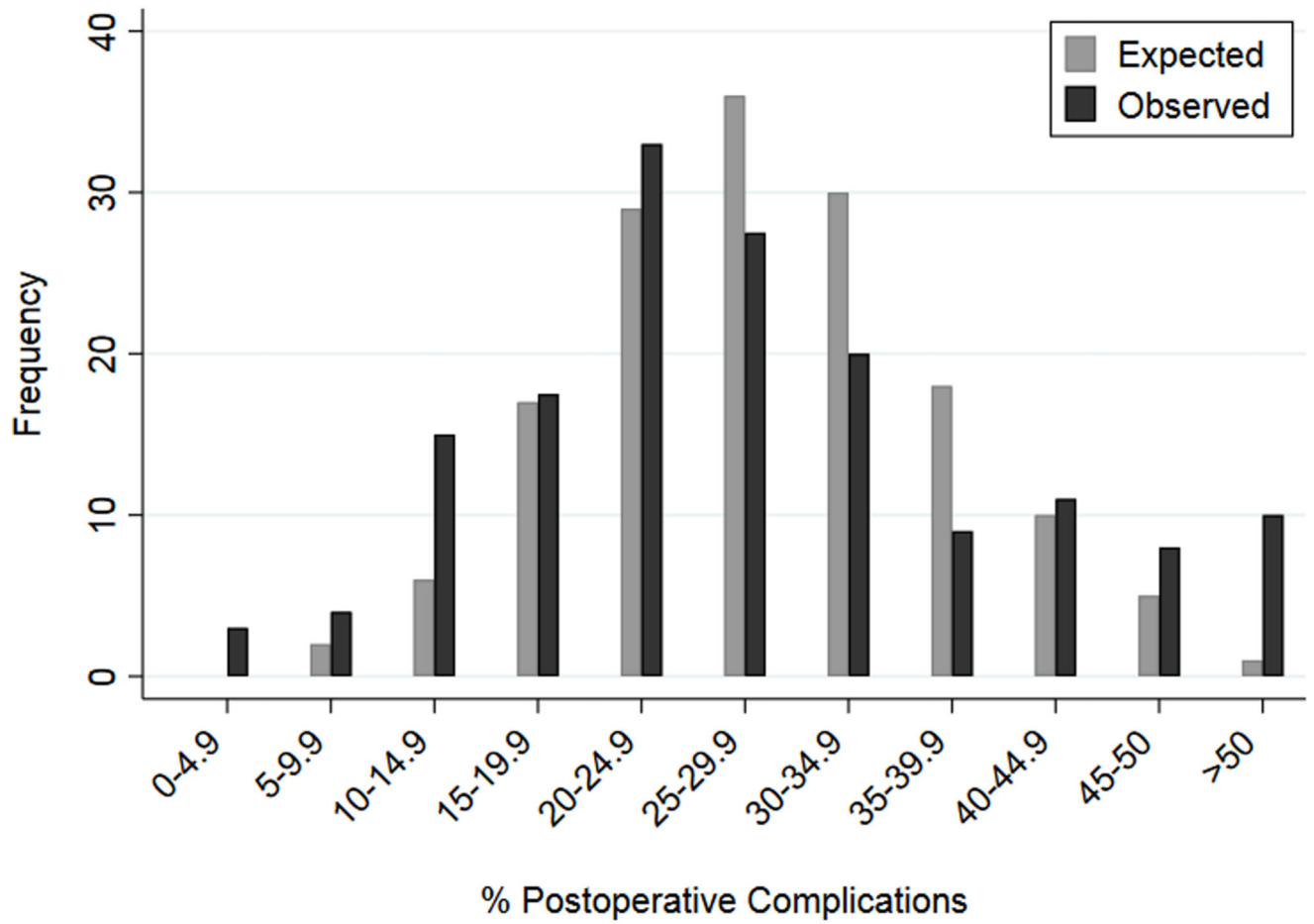
## References

1. Lee KJ, Thompson SG. Clustering by health professional in individually randomised trials. *Bmj*. 2005;330(7483):142–4. [PubMed: 15649931]
2. Stirrat GM, Farrow SC, Fardon J, Dwyer N. The challenge of evaluating surgical procedures. *Annals of the Royal College of Surgeons of England*. 1992;74(2):80–4. [PubMed: 1567147]
3. Roberts C. The implications of variation in outcome between health professionals for the design and analysis of randomized controlled trials. *Stat Med*. 1999;18(19):2605–15. [PubMed: 10495459]

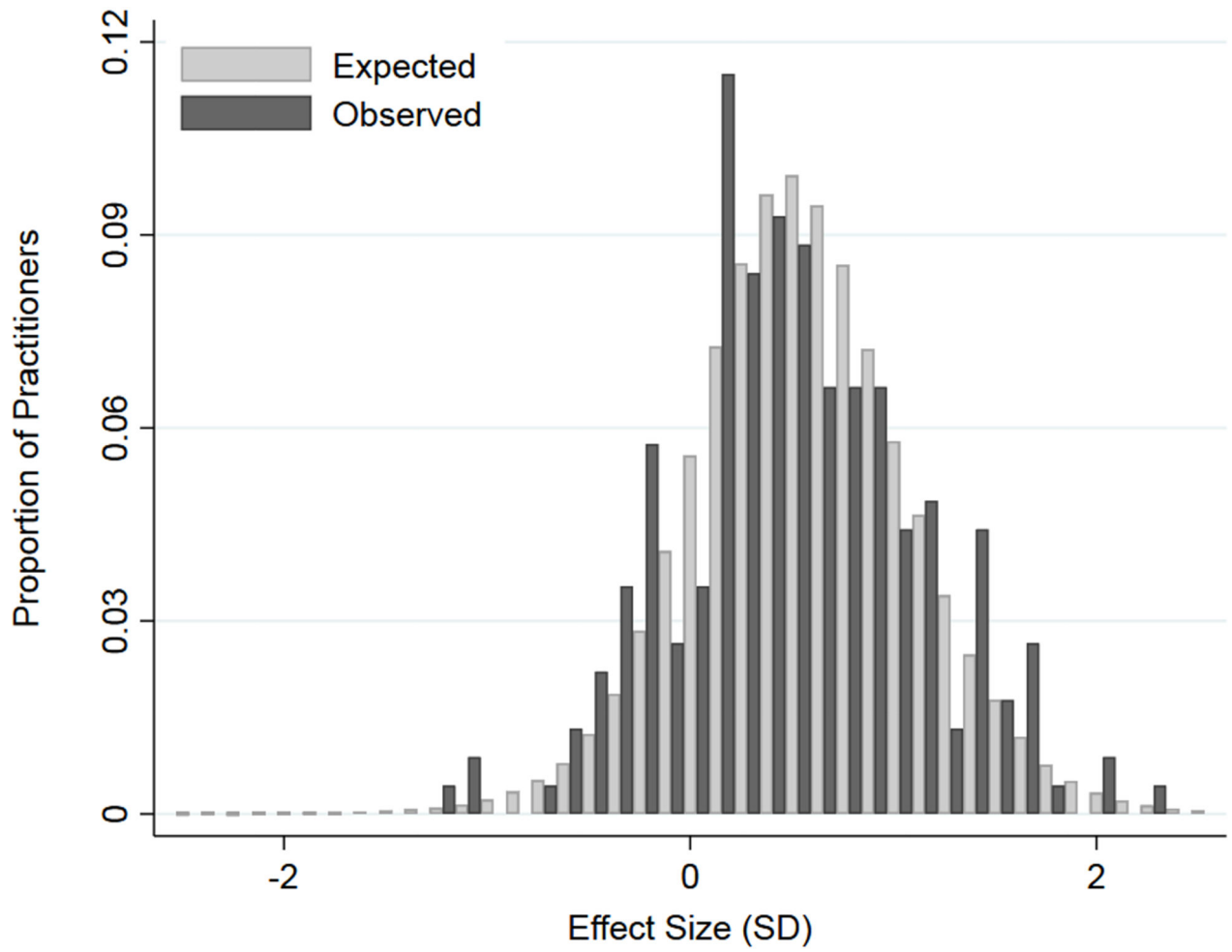
4. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med.* 2002;136(2):111–21. [PubMed: 11790062]
5. Hoover DR. Clinical trials of behavioural interventions with heterogeneous teaching subgroup effects. *Stat Med.* 2002;21(10):1351–64. [PubMed: 12185889]
6. Wampold BE, Brown GS. Estimating variability in outcomes attributable to therapists: a naturalistic study of outcomes in managed care. *Journal of consulting and clinical psychology.* 2005;73(5):914–23. [PubMed: 16287391]
7. Walters SJ. Therapist effects in randomised controlled trials: what to do about them. *Journal of clinical nursing.* 2010;19(7–8):1102–12. [PubMed: 20492055]
8. Walwyn R, Roberts C. Therapist variation within randomised trials of psychotherapy: implications for precision, internal and external validity. *Statistical methods in medical research.* 2010;19(3):291–315. [PubMed: 19608603]
9. Vickers AJ, Vertosick EA, Lewith G, MacPherson H, Foster NE, Sherman KJ, et al. Acupuncture for Chronic Pain: Update of an Individual Patient Data Meta-Analysis. *The journal of pain : official journal of the American Pain Society.* 2017.
10. Vickers A, Zollman C. ABC of complementary medicine. Acupuncture. *Bmj.* 1999;319(7215):973–6. [PubMed: 10514163]
11. Seo Y, Lee IS, Jung WM, Ryu HS, Lim J, Ryu YH, et al. Motion patterns in acupuncture needle manipulation. *Acupunct Med.* 2014;32(5):394–9. [PubMed: 24938530]
12. Vickers AJ, Cronin AM, Maschino AC, Lewith G, MacPherson H, Foster NE, et al. Acupuncture for chronic pain: individual patient data meta-analysis. *Arch Intern Med.* 2012;172(19):1444–53. [PubMed: 22965186]
13. Hinman RS, McCrory P, Pirodda M, Relf I, Forbes A, Crossley KM, et al. Acupuncture for chronic knee pain: a randomized clinical trial. *Jama.* 2014;312(13):1313–22. [PubMed: 25268438]
14. Thomas KJ, MacPherson H, Thorpe L, Brazier J, Fitter M, Campbell MJ, 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]
15. Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Arch Intern Med.* 2006;166(4):450–7. [PubMed: 16505266]
16. Diener HC, Kronfeld K, Boewing G, Lungenhausen M, Maier C, Molsberger A, et al. Efficacy of acupuncture for the prophylaxis of migraine: a multicentre randomised controlled clinical trial. *Lancet Neurol.* 2006;5(4):310–6. [PubMed: 16545747]
17. Suarez-Almazor ME, Looney C, Liu Y, Cox V, Pietz K, Marcus DM, et al. A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication. *Arthritis Care Res (Hoboken).* 2010;62(9):1229–36. [PubMed: 20506122]
18. White P, Bishop FL, Prescott P, Scott C, Little P, Lewith G. Practice, practitioner, or placebo? A multifactorial, mixed-methods randomized controlled trial of acupuncture. *Pain.* 2012;153(2):455–62. [PubMed: 22169359]
19. Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S, et al. Acupuncture in patients with osteoarthritis of the knee: a randomised trial. *Lancet.* 2005;366(9480):136–43. [PubMed: 16005336]
20. 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(12):901–10. [PubMed: 15611487]
21. Carlsson CP, Sjolund BH. Acupuncture for chronic low back pain: a randomized placebo-controlled study with long-term follow-up. *Clin J Pain.* 2001;17(4):296–305. [PubMed: 11783809]
22. Chen LX, Mao JJ, Fernandes S, Galantino ML, Guo W, Lariccia P, et al. Integrating acupuncture with exercise-based physical therapy for knee osteoarthritis: a randomized controlled trial. *Journal of clinical rheumatology : practical reports on rheumatic & musculoskeletal diseases.* 2013;19(6):308–16. [PubMed: 23965480]

23. Cherkin DC, Sherman KJ, Avins AL, Erro JH, Ichikawa L, Barlow WE, et al. A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Arch Intern Med.* 2009;169(9):858–66. [PubMed: 19433697]
24. Endres HG, Bowing G, Diener HC, Lange S, Maier C, Molsberger A, et al. Acupuncture for tension-type headache: a multicentre, sham-controlled, patient-and observer-blinded, randomised trial. *J Headache Pain.* 2007;8(5):306–14. [PubMed: 17955168]
25. Haake M, Muller HH, Schade-Brittinger C, Basler HD, Schafer H, Maier 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(17):1892–8. [PubMed: 17893311]
26. Foster NE, Thomas E, Barlas P, Hill JC, Young J, Mason E, et al. Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the knee: randomised controlled trial. *Bmj.* 2007;335(7617):436. [PubMed: 17699546]
27. Guerra de Hoyos JA, Andres Martin Mdel C, Bassas y Baena de Leon E, Vigara Lopez M, Molina Lopez T, Verdugo Morilla FA, et al. Randomised trial of long term effect of acupuncture for shoulder pain. *Pain.* 2004;112(3):289–98. [PubMed: 15561384]
28. Irnich D, Behrens N, Molzen H, Konig A, Gleditsch J, Krauss M, et al. Randomised trial of acupuncture compared with conventional massage and “sham” laser acupuncture for treatment of chronic neck pain. *Bmj.* 2001;322(7302):1574–8. [PubMed: 11431299]
29. Kennedy S, Baxter GD, Kerr DP, Bradbury I, Park J, McDonough SM. Acupuncture for acute non-specific low back pain: a pilot randomised non-penetrating sham controlled trial. *Complement Ther Med.* 2008;16(3):139–46. [PubMed: 18534326]
30. 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):364–70. [PubMed: 14600536]
31. Kleinhenz J, Streitberger K, Windeler J, Gussbacher A, Mavridis G, Martin E. Randomised clinical trial comparing the effects of acupuncture and a newly designed placebo needle in rotator cuff tendinitis. *Pain.* 1999;83(2):235–41. [PubMed: 10534595]
32. Li Y, Zheng H, Witt CM, Roll S, Yu SG, Yan J, et al. Acupuncture for migraine prophylaxis: a randomized controlled trial. *Cmaj.* 2012;184(4):401–10. [PubMed: 22231691]
33. Linde K, Streng A, Jurgens S, Hoppe A, Brinkhaus B, Witt C, et al. Acupuncture for patients with migraine: a randomized controlled trial. *Jama.* 2005;293(17):2118–25. [PubMed: 15870415]
34. Melchart D, Streng A, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, et al. Acupuncture in patients with tension-type headache: randomised controlled trial. *Bmj.* 2005;331(7513):376–82. [PubMed: 16055451]
35. Scharf HP, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C, et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. *Ann Intern Med.* 2006;145(1):12–20. [PubMed: 16818924]
36. Vas J, Mendez C, Perea-Milla E, Vega E, Panadero MD, Leon JM, et al. Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: randomised controlled trial. *Bmj.* 2004;329(7476):1216. [PubMed: 15494348]
37. Vas J, Ortega C, Olmo V, Perez-Fernandez F, Hernandez L, Medina I, et al. Single-point acupuncture and physiotherapy for the treatment of painful shoulder: a multicentre randomized controlled trial. *Rheumatology (Oxford).* 2008;47(6):887–93. [PubMed: 18403402]
38. Vas J, Perea-Milla E, Mendez C, Sanchez Navarro C, Leon Rubio JM, Brioso M, et al. Efficacy and safety of acupuncture for chronic uncomplicated neck pain: a randomised controlled study. *Pain.* 2006;126(1–3):245–55. [PubMed: 16934402]
39. White P, Lewith G, Prescott P, Conway J. Acupuncture versus placebo for the treatment of chronic mechanical neck pain: a randomized, controlled trial. *Ann Intern Med.* 2004;141(12):911–9. [PubMed: 15611488]
40. Jena S, Witt CM, Brinkhaus B, Wegscheider K, Willich SN. Acupuncture in patients with headache. *Cephalalgia.* 2008;28(9):969–79. [PubMed: 18624803]
41. Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN. Acupuncture for patients with chronic neck pain. *Pain.* 2006.

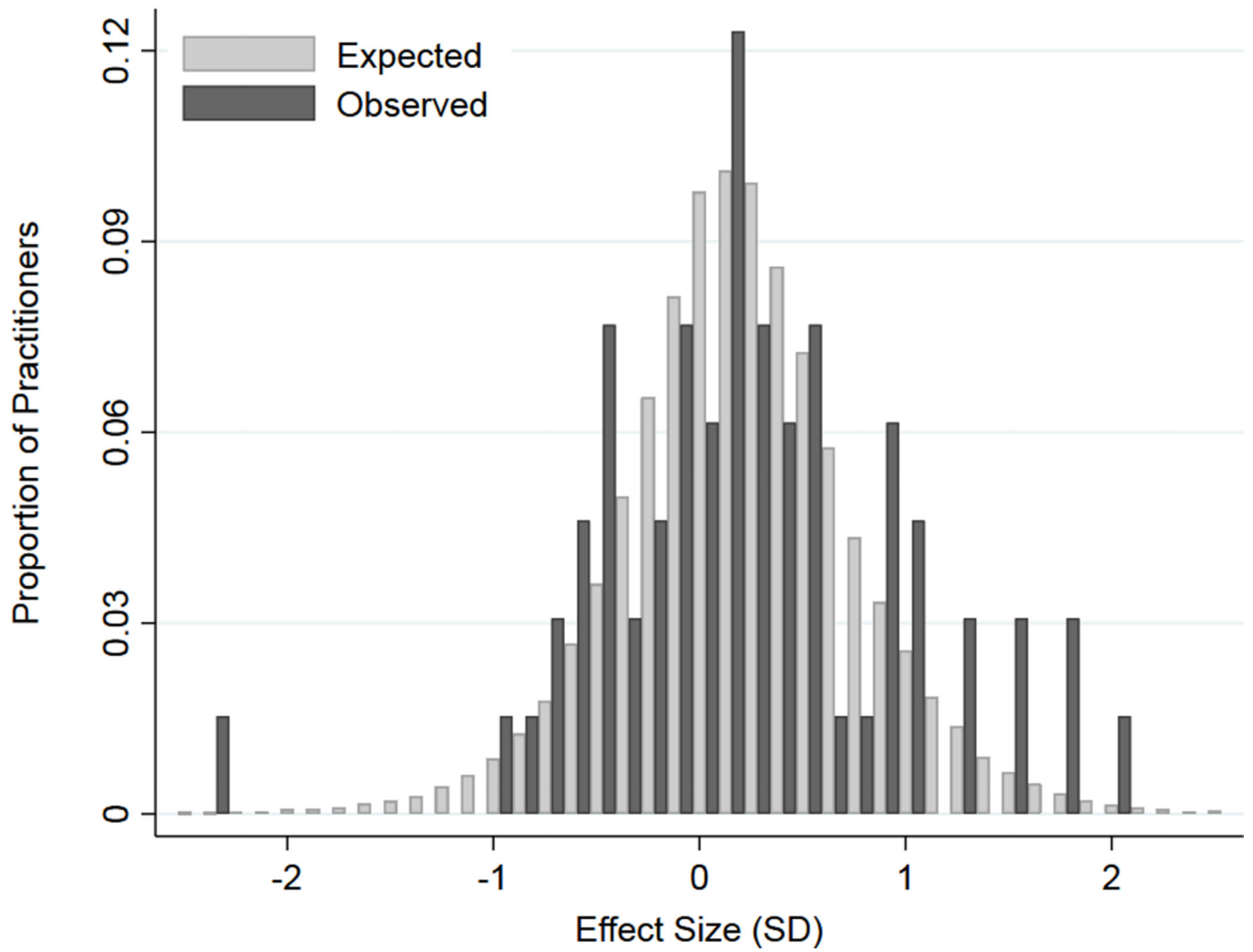
42. Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN. Acupuncture in patients with osteoarthritis of the knee or hip: a randomized, controlled trial with an additional nonrandomized arm. *Arthritis Rheum.* 2006;54(11):3485–93. [PubMed: 17075849]
43. Witt CM, Jena S, Selim D, Brinkhaus B, Reinhold T, Wruck K, et al. Pragmatic Randomized Trial Evaluating the Clinical and Economic Effectiveness of Acupuncture for Chronic Low Back Pain. *Am J Epidemiol.* 2006;164(5):487–96. [PubMed: 16798792]
44. Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith CM, Ellis N, et al. Acupuncture for chronic headache in primary care: large, pragmatic, randomised trial. *Bmj.* 2004;328(7442):744. [PubMed: 15023828]
45. MacPherson H, Tilbrook H, Richmond S, Woodman J, Ballard K, Atkin K, et al. Alexander Technique Lessons or Acupuncture Sessions for Persons With Chronic Neck Pain: A Randomized Trial. *Ann Intern Med.* 2015;163(9):653–62. [PubMed: 26524571]
46. Cherkin DC, Eisenberg D, Sherman KJ, Barlow W, Kaptchuk TJ, Street J, 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(8):1081–8. [PubMed: 11322842]
47. Hunter RF, McDonough SM, Bradbury I, Liddle SD, Walsh DM, Dhamija S, et al. Exercise and Auricular Acupuncture for Chronic Low-back Pain: A Feasibility Randomized-controlled Trial. *Clin J Pain.* 2012;28(3):259–67. [PubMed: 21753728]
48. Lansdown H, Howard K, Brealey S, MacPherson H. Acupuncture for pain and osteoarthritis of the knee: a pilot study for an open parallel-arm randomised controlled trial. *BMC Musculoskelet Disord.* 2009;10:130. [PubMed: 19852841]
49. Salter GC, Roman M, Bland MJ, MacPherson H. Acupuncture for chronic neck pain: a pilot for a randomised controlled trial. *BMC Musculoskelet Disord.* 2006;7:99. [PubMed: 17156464]
50. Weiss J, Quante S, Xue F, Muche R, Reuss-Borst M. Effectiveness and acceptance of acupuncture in patients with chronic low back pain: results of a prospective, randomized, controlled trial. *J Altern Complement Med.* 2013;19(12):935–41. [PubMed: 23738680]
51. Williamson L, Wyatt MR, Yein K, Melton JT. Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement. *Rheumatology (Oxford).* 2007;46(9):1445–9. [PubMed: 17604311]
52. Bianco FJ Jr., Riedel ER, Begg CB, Kattan MW, Scardino PT. Variations among high volume surgeons in the rate of complications after radical prostatectomy: further evidence that technique matters. *The Journal of urology.* 2005;173(6):2099–103. [PubMed: 15879851]
53. Bianco FJ Jr., Vickers AJ, Cronin AM, Klein EA, Eastham JA, Pontes JE, et al. Variations among experienced surgeons in cancer control after open radical prostatectomy. *The Journal of urology.* 2010;183(3):977–82. [PubMed: 20083278]
54. Begg CB, Riedel ER, Bach PB, Kattan MW, Schrag D, Warren JL, et al. Variations in morbidity after radical prostatectomy. *The New England journal of medicine.* 2002;346(15):1138–44. [PubMed: 11948274]
55. MacPherson H, Maschino AC, Lewith G, Foster NE, Witt CM, Vickers AJ. Characteristics of acupuncture treatment associated with outcome: an individual patient meta-analysis of 17,922 patients with chronic pain in randomised controlled trials. *PloS one.* 2013;8(10):e77438.
56. Thomas KJ, MacPherson H, Ratcliffe J, Thorpe L, Brazier J, Campbell M, et al. Longer term clinical and economic benefits of offering acupuncture care to patients with chronic low back pain. *Health Technol Assess.* 2005;9(32):iii-iv, ix-x, 1–109.
57. Witt CM, Ludtke 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? *The journal of pain : official journal of the American Pain Society.* 2010;11(5):431–5.



**Figure 1.** Difference between expected and observed complication rates between surgeons, adapted from Bianco et al.(52)



**Figure 2.** Distribution of observed and expected acupuncturist effect sizes in non-acupuncture-controlled trials, N=226.



**Figure 3.** Distribution of observed and expected acupuncturist effect sizes in sham-controlled trials, N=65.



Number of acupuncturists included in eligible trials, as well as included in trial-level and acupuncturist-level analysis.

**Table 1.**

Control Group	Analysis	Number of trials	Number of acupuncturists	Number in trial-level analysis	Number in acupuncturist-level analysis	Non-physician acupuncturists in trial-level analysis	Non-physician acupuncturists in acupuncturist-level analysis
Sham Acupuncture	Primary	5	213	132	67	9	9
	Combining ARC and ART trials	4	189	114	65	9	9
Non-Acupuncture	Primary	12	4,961	1,203	280	41	31
	Combining ARC and ART trials	8	3,006	1,172	438	41	31
	Excluding non-physician acupuncturists	4	43	41	31	41	31

Note that the differences in the number of acupuncturists included results from differences between analytic approaches in the minimum number of patients an acupuncturist needs to have treated to be included.